Debunking an Urban Legend:
Uniformity in Edge-Lit Frustrated TIR Displays

M. G. Selbrede*

*Unipixel Displays, Inc., Science Department
8708 Technology Forest Place Suite 100, The Woodlands TX 77384, martin.selbrede@unipixel.com

ABSTRACT

Field sequential color display systems that utilize frustrated total internal reflection to activate pixels offer significant advantages over existing display technologies, including display transparency, creating visible and infrared images on the same surface, efficiencies of 25 lumens/watt and higher, and elimination of TFTs. Development of such MEMS-based displays has been stymied by a surprisingly resilient urban legend: the claim that light in a TIR waveguide weakens with distance from the light sources, creating a brightness gradient in such displays. It is assumed that even if the light *were* uniformly distributed in the slab initially, activation of pixels entails "less light downstream" for pixels farther from the source, again creating an unacceptable brightness gradient. The actual preconditions for display uniformity are considered, countering the one persistent obstacle impeding the development of MEMS-based displays that exploit frustrated TIR. Such displays can be designed to exhibit excellent uniformity, initial prejudices notwithstanding.

Keywords: flat panel displays, display uniformity, edge-lit backlights, display optimization

1 EDGE-LIT FTIR DISPLAYS

Flat panel displays built around the principles of field sequential color using light valves that frustrate total internal reflection (FTIR) to provide pulse width modulated gray scales have evident advantages over conventional display technologies. MEMS-based systems that peripherally tether an ultra-thin deformable membrane about a micron above a TIR waveguide using low refractive index standoffs provide the core functionality of a light valve that can controllably frustrate TIR. Such structures exhibit inherent hysteresis, and there can therefore dispense with the need for active matrix drive methods based on TFTs. Color being generated using field sequential color techniques allows for pixel architectures that are unicellular rather than tricellular, while the absence of polarizers and color filters delivers impressive benefits in power efficiency (>25 lumens/watt). Since such displays are transparent by nature, redundancy for mission critical applications can be achieved along the z-axis (vertical stacking) rather than by shrinking the display and its sister backup unit to occupy space in the x-y plane of the instrumentation module of interest. Moreover, such displays can even modulate infrared *and* visible light by way of temporal interleave, can control additional primaries to expand the fraction of the CIE color space being transduced, and can temporally interleave the left and right frames of an autostereoscopic video image in synchronization with a global dynamic redirection layer. The vast reduction in device complexity also bodes well for commercialization.

2 HISTORIC ISSUES IMPEDING R&D

In the 1990s, development of FSC-based FTIR displays suffered due to industry-wide confidence in the massive amount of brute force being applied to AMLCD displays to conquer the large area display challenge. It took more than a decade before it became widely recognized, and all but irrefutable, that this initial confidence was misplaced, igniting renewed interest in elegant solutions that sought to bypass brute force techniques.

Although the unicellular solution using FSC-based FTIR techniques had great promise as one such elegant solution, the approach had two additional hurdles to overcome: color breakup artifacts (which do indeed exist) and allegations of gross brightness nonuniformity (a bias fueled by a common misapplication of physical intuition).

2.1 Motional Color Breakup Artifacts

Even at video frame rates of 60 fps, the likelihood of motional color image breakup is high. The human eye can rotate at angular velocities up to 700 degrees per second. Even below this threshold, the fact that field sequential color requires *consecutive* generation of the red, green, and blue tristimulus components of an image creates problems. As illustrated in Figure 1 (an idealized top-down view), as the observer tracks an object on the display surface, the observer's retina rotates. As it rotates, the red, green, and blue components of the tracked target will fall in different places on the retina. The resulting image exhibits a form of temporal smear, illustrated in Figure 2.

The artifact can be suppressed by making the red, green, and blue components fall on the retina in the same location. This is traditionally done by rewriting the trailing edge frames to anticipate retinal motion [1], but if two or more targets exist, this approach has reduced utility. It is adopted because LCD response doesn't permit an alternative. But an FSC-based FTIR display can emit all three colors for a 60 fps video frame in a time $t < 4$ ms, followed by a quiescent

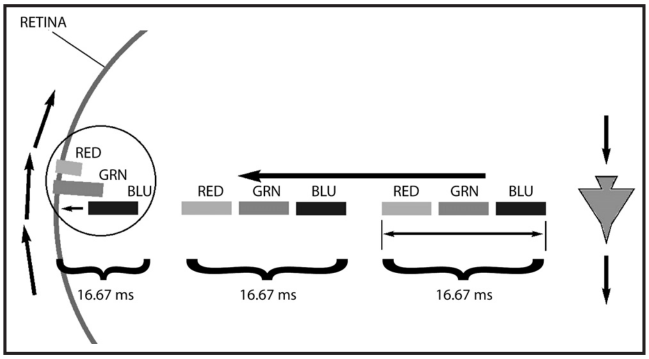


Figure 1: Cause of motional color breakup illustrated.

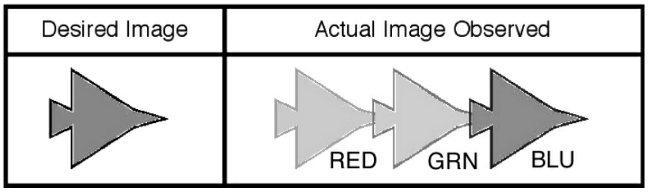


Figure 2: FSC image without artifact suppression

black subframe for the remaining 12.67 ms of the frame. As the global pulse is truncated to one-fourth its normal length, the intensity must be quadrupled to insure that the total flux admitted by the retina remains unchanged [2]. Power consumption is also unchanged due to the intrinsic duty cycle compensation. As a result, a MEMS-based solution to the problem lies within reach.

The second issue raised against such unicellular displays (2.2 below) has served to impede progress on this otherwise promising front. This is all the more unfortunate given the advantages such nanoscale MEMS-based displays could deliver once unleashed.

2.2 Allegedly Poor Brightness Uniformity

Development of FSC/FTIR/MEMS-based displays has been stymied by a surprisingly resilient urban legend: the claim that light in a TIR waveguide weakens with distance from the light sources, creating a brightness gradient in such displays. It is assumed that even if the light *were* uniformly distributed in the slab initially, activation of pixels entails "less light downstream" for pixels farther from the source, again creating an unacceptable brightness gradient.

This is at heart a cognitive problem, inasmuch as human intuition is not so easily reined in, even by sound physical reasoning. Even engineers viewing a prototype exhibiting excellent uniformity will allege that nonuniformities will nonetheless arise, as if intuition even trumps their own observations. This is, in effect, a variation on the sociological observation codified in Blaauw's Law, which states that established technology tends to persist in the face of new technology [3]. In the present case, established perceptions persist in the face of new paradigms. Since early adoption is the key to rapid technological advance, this impediment to progress needs to be overcome, so that the technology afflicted by the bias can be judged on its actual merits.

Human intuition serves many valuable functions, as it can parse complex concepts heuristically to expand our horizons of understanding. On occasion, it becomes mired in its own paradigms (in support of the contention of William James that people believe they're thinking when they're actually only rearranging their prejudices). Intuitions are rooted in rational considerations (as shall be shown), but those considerations are evaluated asymmetrically, thereby crippling our objectivity.

It is difficult to appreciate the fact that the slab wave guide used in a FTIR display is extensively mirrored on at least three of the four smaller surfaces: TIR is supported only on the large surfaces (*top* and *bottom*, by convention). This certainly changes the conception of "downstream" since the manifold of interest undergoes replication folding. The human mind resists yielding to this complicating factor, and tenaciously bases its case on other factors held to create non-uniformity: *rapid attenuation* of energy, and *crosstalk* between different pixels.

With a conventional backlight, each pixel in an LCD display receives light from the region immediately behind it. There is no communal sharing of light (in effect, crosstalk) between pixels due to the one-to-one correspondence between each ray emitted from the backlight and the pixel that ray passes through. This zero crosstalk scenario, given a uniform backlight, is intuitively known to be ideal for good brightness uniformity.

Since there is considerable pixel-to-pixel crosstalk in an FTIR display, intuition tends to accrete around the notion that pixels are being shorted the luminous flux necessary to insure uniformity. Because the pixels are assumed to be efficient light valves, the magnitude of this "theft" tends to compound the perceived problem. Robbing Peter to pay Paul is considered an unsound foundation for display uniformity among FSC/FTIR-based systems. Since this theft is most severe at the first activated pixels encountered, human intuition concludes that non-uniformity is a certainty. In effect, the original basis for the allegation is recovered in the process of sorting through these factors.

The damage done by these unchallenged notions, particularly in lost time and opportunity, is hard to calculate. Once it is demonstrated that these biases are incorrect, it will be possible to roll back the effects of FUD (fear, uncertainty, and doubt) that were once cast upon FSC/FTIR displays by such misplaced appeals to incompletely structured physical intuition.

3 FACTORS AFFECTING UNIFORMITY

Several factors combine to insure uniformity in an FTIR display using an edge-lit slab waveguide. A few of them appear, at first glance, to harm uniformity, but careful *systemic* assessment of all the factors involves will lead compellingly to the conclusion that uniformity is a parameter that can be tuned. It is of interest that ultra-high uniformity, while achievable for FTIR displays, is not necessarily desirable. There is a point beyond which improving uniformity harms signal-to-noise ratio. Therefore, an appropriate engineering trade-off must be made at the design stage.

3.1 Lowered Individual Pixel Efficiency

Detuning individual pixel coupling efficiency (the frustration of TIR) is one of the keys to uniformity. Such detuning is a necessary, but not sufficient, condition for high uniformity to be realized. Peter is allowed to rob Paul, but not allowed to take his entire purse. Such detuning can be done by restricting the coupling area, or directly affecting coupling efficiency by careful selection of refractive indices among the layers involved in the coupling process. Areal restriction includes both aperture restriction and other geometric means to reduce the size of the coupling region.

If an activated (open) pixel only couples 25% of the TIR light that interacts with it, it will take 18 encounters with open pixels for a given photon ensemble to be depleted down to 1% of its original intensity: 99% of the light will have been used, albeit distributed among 18 open pixels. After another 18 encounters, that remaing 1% will drop to 1% of 1%, or about 0.01%. We reach a practical asymptote rather quickly, so it is useful to use the 99% depletion point and not consider the last 1% in our heuristic model.

Note that any given pixel on the display will be receiving ensembles arriving at the pixel on their first encounter, other ensembles on their second, still others on their third, etc. In fact, with so many trillions of ensembles traveling through the waveguide, the results through any given pixel region become statistically mixed to a high degree. The solution of these summations provide two key pieces of evidence: the actual *system* power efficiency of the display, and the *maximum power difference* between any two arbitrarily chosen pixels on the display surface. Both quantities are affected, not only by individual pixel efficiency, but by the interaction of the other parameters that work together to insure high display uniformity.

3.2 Display Thickness to Area Ratio

The thickness of the edge-lit waveguide as compared to its areal extent is another key factor in display uniformity. In conjunction with detuned pixel coupling efficiencies, relative thickness of the display determines how far a photon ensemble will travel through the slab before it reaches 99% depletion (18 bounces). For an ultra-thin slab, trigonometric considerations may entail the depletion of 99% of the TIR-compliant light energy within a short distance of the light insertion point. The thicker slab leads to longer distances between ensemble encounters with open pixels, thereby giving a longer path length for the ensemble during its meaningful lifespan (the release of the first 99% of its energy).

It is significant that a display constructed in violation of either parameter 3.1 and 3.2 will exhibit precisely the kind of non-uniformity that intuition tells us should occur. The acknowledgement of this fact tells us that our intuition isn't faulty so much as it is merely working with incomplete data.

As noted above, an excessively long ensemble path is not necessarily desirable. The noise floor for such edge-lit wave guides is due to (1) deviation from optical purity of the waveguide material, such that while it is largely transparent, some light in it nonetheless undergoes volume scattering; and (2) deviations from orthogonality and parallelism, which cause TIR-compliant light near the critical angle to drift out of compliance by cumulative interaction with optical surfaces deviating from true rectilinearity. For both (1) and (2), longer ensemble paths exacerbate the potential effect, raising the system noise floor and thus reducing the signal-to-noise ratio of the display. Photon ensembles need to travel long enough in the waveguide to cross the slab several times (for uniformity's sake), but not long enough to generate excess noise (which is a function of how long a photon ensembles travels through an imperfect medium).

It is a curious circumstance that the noise floor is significantly affected by the refractive index of the slab. The higher the index, the small the exit cone available for the release of scattered light (i.e., system noise) to the observer.

3.3 Unlimited Pixel Crosstalk

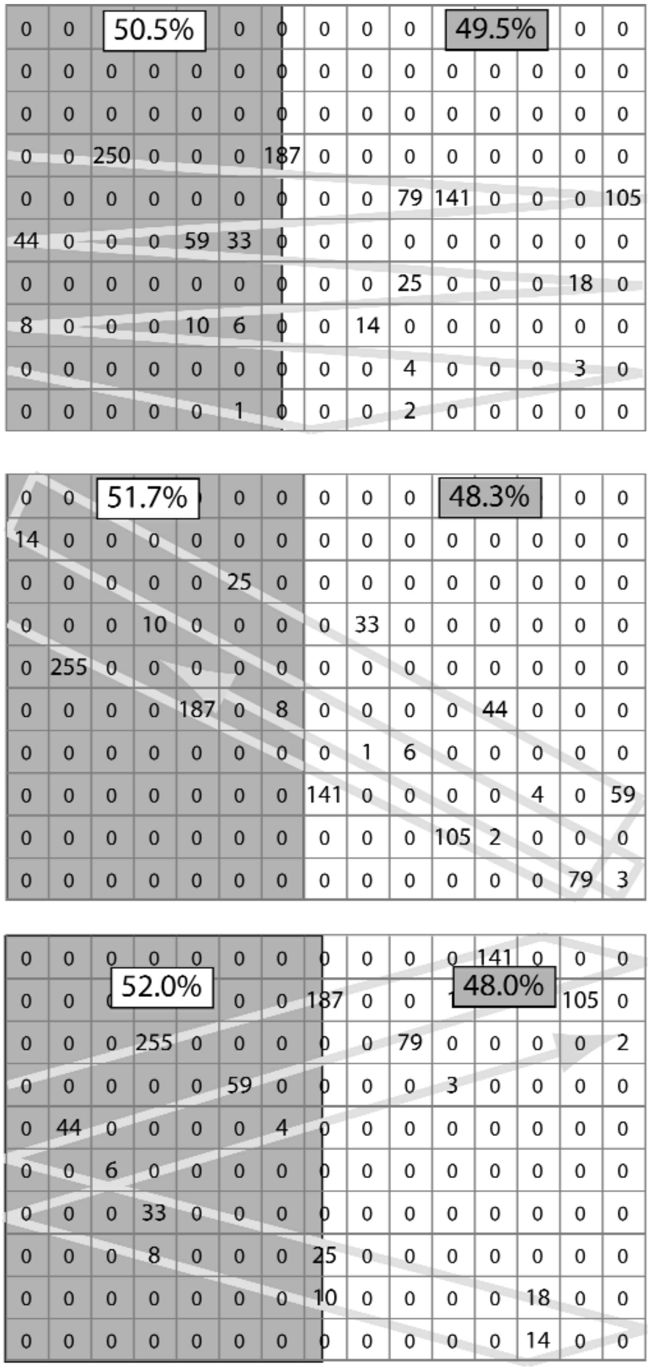
Doubly ironic is the circumstance that for edge-lit waveguides driving an FTIR-based display, unlimited (so-called *infinite*) crosstalk is not only acceptable but crucially important. The first efforts to codify the principles underlying display uniformity for such system occupied Rodabaugh for nine years, beginning in 1997 [4]. The analysis was conducted mathematically, based on a powerful formalism that used a chi-squared best fit statistical criterion to determine pixel-to-pixel power variation in worst-case scenarios. The relationship between individual pixel coupling efficiency and uniformity were discovered. Detuning individual pixels such that overall (global) efficiency was 61% still guaranteed no pixel-to-pixel variation greater than 0.987 dB. By detuning a couple of percentage points lower, the uniformity improves to < 0.3 dB variation across the entire display, regardless of program content.

The fact of unlimited crosstalk means that the activation of any given pixel *affects all other pixels almost equally.* Rodabaugh couched this in terms of predecessor relationships: the likelihood that within the system, any given photon ensemble striking pixel X had previously encountered pixel Y was *very* high, regardless of the positions of X and Y on the display. Here, intuition can be more serviceable. It is clear that if one pixel affects all the others nearly equally, the display will not suffer from a uniformity problem. The *drain* imposed by the opening of pixel X is, therefore, distributed over *all other pixels* nearly evenly in such displays.

One consequence worth noting is that while uniformity is aided by unlimited crosstalk, this effect does cause a predictable nonlinearity in *absolute output*. For example, consider a display with one million pixels. In one instance, only one pixel is lit at full intensity. In another instance, all one million are lit. The million lit pixels do *not* emit one million times the intensity of the single lit pixel: the single pixel is brighter than any of the million lit pixels by a small factor. Since the human eye processes *contrast* more readily than *absolute luminosity*, there would be little benefit in correcting for this imperceptible nonlinearity. If such were desired, however, the adjustment to the illumination module is so straightforward that pure absolute linearity can readily be recovered with a trivial adjustment to the drive system.

4 ENERGY PARTITIONING

The argument based on energy partitioning is illuminating insofar as it actually aids in visualizing the distribution of a photon ensemble's flux during transit through the waveguide as various open pixels are encountered.



0	0	50.5%			0	0	0	0	0	49.5%			0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	250	0	0	0	187	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	79	141	0	0	0	105
44	0	0	0	59	33	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	25	0	0	0	18	0
8	0	0	0	10	6	0	0	14	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	4	0	0	0	3	0
0	0	0	0	0	1	0	0	0	2	0	0	0	0	0

0	0	51.7%			0	0	0	0	0	48.3%			0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	25	0	0	0	0	0	0	0	0	0
0	0	0	10	0	0	0	0	33	0	0	0	0	0	0
0	255	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	187	0	8	0	0	0	0	44	0	0	0
0	0	0	0	0	0	0	0	1	6	0	0	0	0	0
0	0	0	0	0	0	0	141	0	0	0	0	4	0	59
0	0	0	0	0	0	0	0	0	0	105	2	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	79	3

0	0	0	0	0	0	0	0	0	0	0	141	0	0	0
0	0		52.0%		0	0	187	0	0		48.0%		105	0
0	0	0	255	0	0	0	0	0	79	0	0	0	0	2
0	0	0	0	0	59	0	0	0	0	3	0	0	0	0
0	44	0	0	0	0	4	0	0	0	0	0	0	0	0
0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	33	0	0	0	0	0	0	0	0	0	0	0
0	0	0	8	0	0	0	25	0	0	0	0	0	0	0
0	0	0	0	0	0	0	10	0	0	0	0	18	0	0
0	0	0	0	0	0	0	0	0	0	0	0	14	0	0

Figure 3: Energy partition for representative ensembles

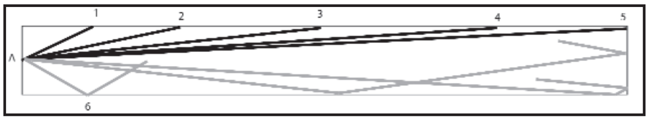


Figure 4: Side view of top bounces and some conjugates

Figure 3 illustrates three of many millions of possible ensemble paths traveling through the waveguide. These paths all start at the same origin, but obviously have different destinies. The ensembles are arbitrarily assigned 1,000 photons each. The first encounter with an open pixel is represented by a loss of 25% of the energy (250 photons released). The next encounter release 25% of the remaining 750 photons, or 187 photons. The depletion series continues 141, 105, 79, down to 1.

One could laboriously assemble all of these snapshots for individual pixels, overlay them, and count up all the numbers (as Rodabaugh did statistically). Or, one can show that another effect is inexorably operating in these illustrations: the energy between the left and right halves of the screen is averaging out at the center of the display.

Before we consider this tactic, note the side view cross-section of the waveguide in Figure 4. For every ensemble emanating from light source A striking the *top* of the waveguide first (Point 1), there is a sister ensemble (the *conjugate* rays) striking the *bottom* of the waveguide first (Point 6). Compared to their top-first counterparts, the conjugate ensembles strike the top of the waveguide *offset by a factor tied to the incidence angle*. The behavior of all ensembles and their conjugate partners is involved in the analysis.

In Figure 3, the equipartition point for light energy is demarcated by the terminator between gray on the left and white on the right half of each display. The large percentages along the top tabulate energy per integral pixel, whereas the exact 50% point (at the gray/white terminator) can cut through fractional pixels.

The *average* position (considering *multiple* ensemble paths) for the terminator, for the top-first ensembles and their conjugates, is dead-center, mid-screen. The more trial ensemble paths are used, the closer to the center the terminator approaches. *A centered terminator evidently lies at the limit of the summation process*. This is significant because a requirement for a uniform display is that the left half of the display exhibit the same intensity as the right half. This simple prerequisite is satisfied by the heuristic model presented. Our intuition conflicts with this fact, but our intuition never bothered to count the actual photons in action.

REFERENCES

[1] Koma, N. and Uchida, T. "A Novel Display Method for Field Sequential Color LCD without Color Breakup," *SID Digest 2001*, pp. 400-403.

[2] *Phase I Final Report: Contract No. F42650-03-P-2751* (SBIR AF03-280) 17NOV03 by Uni-Pixel Displays, Inc., p. 3ff.

[3] Dickson, Paul. *The Official Rules and Explanations* (Springfield: Merriam-Webster, 1978), p. 14.

[4] Rodabaugh, D. *White Paper on TMOS Uniformity*, proprietary in-house document, Unipixel Displays, Inc, Sept. 2004, first issued as *TMOS Simulation Project* for Tralas Technologies, Inc. in 1997.