Image splitting techniques for a dual layer high dynamic range LCD display

Gabriele Guarnieri\textsuperscript{a}, Luigi Albani\textsuperscript{b} and Giovanni Ramponi\textsuperscript{a}

\textsuperscript{a}Image Processing Laboratory, DEEI, University of Trieste, Trieste, Italy
\textsuperscript{b}FIMI-Philips, Saronno, Italy

ABSTRACT

Liquid crystal displays (LCD) are replacing analog film in radiology and permit to reduce diagnosis times. Their typical dynamic range, however, can be too low for some applications, and their poor ability to reproduce low luminance areas represents a critical drawback. The black level of an LCD can be drastically improved by stacking two liquid crystal panels in series. In this way the global transmittance is the pointwise product of the transmittances of the two panels and the theoretical dynamic range is squared. Such a high dynamic range (HDR) display also permits the reproduction of a larger number of gray levels, increasing the bit depth of the device. The two panels, however, are placed at a small distance one from each other due to mechanical constraints, and this introduces a parallax error when the display is observed off-axis. A complex, spatially-adaptive algorithm is therefore necessary to generate the images used to drive the two panels.

In this paper, we describe the characteristics of a prototype dual-layer HDR display and discuss the issues involved in the image splitting algorithms. We propose some solutions and analyze their performance, giving a measure of the capabilities and limitations of the device.

Keywords: Liquid crystal displays (LCD), High dynamic range (HDR), Medical imaging, Image processing

1. DESCRIPTION OF PURPOSE

Consumer displays sold nowadays are almost totally of the LCD type. They offer flat surface and small physical dimensions, a high brightness and resolution, no flicker, no geometric distortion, good longevity, low power consumption and low electromagnetic emissions. LCDs are also widespread in medical imaging applications and are gradually replacing analog film in radiology. LCDs, however, still do not match the performance of traditional radiographic film. A film-based x-ray viewed on the traditional light box can reach a peak brightness of around 4000 cd/m\textsuperscript{2} and a dynamic range (defined as peak brightness divided by black level) of around 3000:1, whereas an LCD display can reach a few hundred cd/m\textsuperscript{2} and a dynamic range below 3 orders of magnitude. In particular, the black level of an LCD display still has a nonzero luminance, because the panel is not able to completely block the light coming from the backlight unit. This may be not acceptable in some medical applications where a high dynamic range is necessary in order to discriminate a larger number of luminance levels and detect fine details with very small luminance differences.

The black level of an LCD monitor can be drastically reduced by stacking two panels in series.\textsuperscript{1} In this way, the global transmittance is the pointwise product of the transmittances of the two panels and the theoretical dynamic range is squared. The grayscale reproduction accuracy is also increased, since the light coming from the backlight unit is modulated twice, although the resulting bit depth is not doubled because different combinations of the two panels can produce the same output level. Unlike a conventional display, a dual layer display can not be directly connected to the output of a PC graphics card or a compatible image source; some processing is necessary in order to generate the two images which drive the two panels. For this purpose, one major difficulty comes from the fact that the two panels are placed at a small distance one from each other due to mechanical

\textsuperscript{1} Further author information: (Send correspondence to Gabriele Guarnieri)
Gabriele Guarnieri: E-mail: gguarnieri@units.it
Luigi Albani: E-mail: luigi.albani@philips.com, Telephone: +39 (02) 96175348
Giovanni Ramponi: E-mail: ramponi@units.it, Telephone: +39 (040) 558 7853
constraints, and therefore some parallax error is introduced if the position of the user is not orthogonal to the panel. The error can be reduced by means of appropriate image processing techniques.

Alternative designs exist for HDR displays which use a single LCD panel combined with a spatially-varying backlight unit, typically consisting of an array of individually modulated LEDs which are dimmed locally in correspondence of the dark areas of the image. This solution also requires the use of complex image processing techniques, since the LED array has a significantly lower resolution than the LCD panel, and is critical with regard to the high power consumption and consequent thermal problems.

This paper is structured as follows. In Section 2 we describe the hardware characteristics of the prototypes and their operating principles. In Section 3 we show the effects of the parallax error, which is intrinsically present in a dual layer display, and discuss the general issues involved in an image splitting algorithm. In Sections 4 and 5 we propose two image splitting techniques which guarantee a perfect reconstruction of the image when the two panels are aligned, and try to minimize the parallax error introduced by off-axis view. The first proposed method computes the backpanel by means of an appropriate nonlinear lowpass filter; the second proposed method seeks an optimal solution to the splitting problem by performing a constrained minimization of an appropriate functional. In Section 6 we discuss the results obtained using the second proposed method, and some open issues. Since the method is based on a constrained optimization procedure, in which the objective function measures the parallax error and the constraints are strictly derived from the device specifications, the results give a measure of the capabilities and physical limitations of the proposed dual layer display. Finally, in Section 7 we conclude and give an overview of the future work.

2. CHARACTERISTICS OF THE PROTOTYPES

Several prototypes of dual layer LCD displays for medical imaging applications were built at FIMI-Philips and are currently under test. Each prototype was built using two grayscale IDTech 18 inch LCD panels. After removing the original backlight unit, the individual LCD glass panels were mechanically matched one on top of the other, taking care to align the active areas at best, and finally combined to a single high-brightness backlight unit. The panel facing the backlight unit is referred to as the backpanel, and the one seen by the viewer is referred to as the frontpanel. The driving electronics is properly arranged in order to allow each panel to be driven independently from a standard DVI input. To operate the display a standard PC is used, on which a graphics card with dual DVI output is installed. Each panel of the dual layer monitor is attached to one output. A specific software package has been developed for the proper generation of the two images which drive the frontpanel and the backpanel of the prototype.

A picture of a prototype is visible in Figure 1. On the left half of the screen, the dual layer display is fully exploited. For comparison, on the right half of the screen a conventional single layer LCD is simulated by displaying the whole image on the frontpanel and a white background on the backpanel. The improved black level of the dual layer display is clearly visible. Besides the high dynamic range and the improved bit depth, the grayscale dual layer LCD prototypes also features an enhanced performance in terms of viewing angle, in the sense that a reduced luminance drop-off with angle compared to a conventional display was measured.

3. SPLITTING ALGORITHMS

The processing used to generate the two images to be displayed on the two panels plays a fundamental role in the performance of the device. The simplest possible technique is to perform the splitting on a pixel-by-pixel basis. More precisely, if we indicate with \( L_{in}(x, y) \) the luminance of the input image at pixel location \((x, y)\) and with \( L_b(x, y) \) and \( L_f(x, y) \) the luminance of the backpanel and frontpanel respectively, the splitting algorithm takes the form

\[
L_b(x, y) = F(L_{in}(x, y)) \quad L_f(x, y) = \frac{L_{in}(x, y)}{L_b(x, y)}.
\]

In other words, the backpanel is computed by mapping the input luminance with a suitable nonlinear function \( F(\cdot) \), and the frontpanel is subsequently computed by division in order to guarantee that the product of the two images reproduces the input. The luminance values in Equation (1) and the following are expressed in normalized units; this allows us to simplify the notation by dropping any scaling factors and to consider without
distinction the output luminance or the panel transmittance. The splitting is computed on linear data; the nonlinear encoding of the source image (if present) and the response of the liquid crystal panels are compensated appropriately by mapping the data before and after the processing. An intuitive choice for the function $F(\cdot)$, suggested by the symmetry of the system, is a square root. In this way, each panel displays the same image. However, in a real device the liquid crystals are enclosed between two glass plates which introduce a finite distance between the planes on which the images are formed even if the two panels are in contact. Instead of seeing the correct image $L_{\text{out}}(x,y) \triangleq L_b(x,y) L_f(x,y)$, an observer looking at the display from an off-axis position sees a distorted image $L_{\text{out}}(x,y) \triangleq L_b(x + \Delta x, y + \Delta y) L_f(x,y)$, where the displacements $\Delta x$ and $\Delta y$ depend on his viewing angle. This form of distortion is referred to as *parallax error* and gives rise to artifacts such as those shown in Figure 2.

In order to reduce the distortion, the backpanel image should be blurred, so that a small displacement does not alter the pixel values excessively. The frontpanel image is then sharpened to compensate for the blurring. Designing the corresponding algorithm is however a nontrivial task which should meet different and conflicting requirements. One desirable property for a splitting algorithm is that the image resulting from the combination of the backpanel and the frontpanel be equal to the input image when the two panels are aligned. We shall call this property *perfect reconstruction*:

$$L_b(x,y) L_f(x,y) = L_{\text{in}}(x,y) \quad \forall (x,y).$$

(2)
This simplifies the problem because only the backpanel needs to be computed; the frontpanel is then generated automatically by division. It is also possible to compute the frontpanel first, but the former approach is more natural to follow because the design objectives are easier to express for the backpanel. After the splitting, the frontpanel and backpanel luminances are converted to digital driving levels (DDLS) and quantized to 8 bits. As previously noticed, a nonlinear mapping must be performed in order to compensate the distortion introduced by the panels; this operation is commonly known as *gamma correction*. For greater accuracy, we measured the actual response of the panels used in the prototypes, rather than using an analytical curve defined in standard recommendations for display or television devices.

4. METHOD 1: CONSTRAINED FILTERING OF THE BACKPANEL

A simple improvement on the square root technique described in the previous section consists in blurring the square root of the input image with a lowpass filter in order to obtain a smooth backpanel. The frontpanel is then computed by division:

\[
L_b(x, y) = \text{BLUR} \left( \sqrt{L_{in}(x, y)} \right) \quad L_f(x, y) = \frac{L_{in}(x, y)}{L_b(x, y)}.
\]

(3)

Blurring is however a nontrivial task, because it must take into account the limited dynamic range of the panels. If a linear lowpass filter is used, approximately half the pixels in the backpanel will be darker than \(\sqrt{L_{in}}\), and the corresponding pixels in the frontpanel will be brighter. If no precaution is taken, some pixels in the frontpanel might exceed the white level, and the consequent clipping will introduce a distortion in the reconstructed image. In order to prevent the frontpanel from clipping, some sort of nonlinear filter must be used.

A problem which often arises in signal and image processing is the computation of *constrained lowpass filters*, or *envelopes*. This operation basically consists in approximating a signal with a smoother function which is bounded from below by the signal itself. The effect can be visually interpreted as an elastic rope or membrane resting on top of the signal graph. If a higher smoothness is required, it is possible to simulate the behavior of an elastic beam or plate, typically by means of spline models. A comparison of the two formulations on a 1D signal is illustrated in Figure 3; the main difference is that a membrane-based envelope can have a discontinuous first derivative in the points where the constraint is active, whereas spline-based envelopes are smooth also in this case.

![Figure 3. Membrane-based and spline-based envelope filters with different “bandwidths”](image)

A possible technique which avoids the clipping of the frontpanel consists in filtering the square root of the input image with a constrained lowpass filter rather than a simple linear lowpass filter. The backpanel computed in this way meets by definition the constraint \(L_b(x, y) \geq \sqrt{L_{in}(x, y)}\). Consequently, the frontpanel computed by division will be upper bounded by the square root, which is lower than the white level provided that the input image is suitably scaled, and clipping is avoided:

\[
L_b(x, y) \geq \sqrt{L_{in}(x, y)} \quad \Rightarrow \quad L_f(x, y) = \frac{L_{in}(x, y)}{L_b(x, y)} \leq \sqrt{L_{in}(x, y)}
\]

(4)

This property is visualized in Figure 4 using a simple 1D test signal. An application to a real image is shown in Figure 5.
A constrained lowpass filter can be implemented in several ways. A heuristic approach may consist in blurring the image with a linear lowpass filter and then adding an offset to the output or parts of it. More advanced methods may formulate the filtering operation as a constrained optimization problem. A possible objective function is

$$\int \int \left\{ \left| \nabla f(x, y) \right|^2 + \lambda \left[ f(x, y) - u(x, y) \right]^2 \right\} dx dy = \text{Min} \quad f(x, y) \geq u(x, y),$$

where $u(x, y)$ is the input image, $f(x, y)$ is the filtered output and the integral is computed over the whole image area. The first term privileges a smooth output, and the second term privileges an output which closely approximates the input; the scalar $\lambda$ allows to set the tradeoff between these two objectives and adjust the “bandwidth” of the filter. A spline-based filter can be obtained by replacing the first term in (5) with a thin plate spline energy functional; this approach involves higher-order derivatives and has a remarkably higher computational complexity. Once (5) is discretized, a quadratic programming problem is obtained, which can be solved by means of appropriate iterative methods that exploit its structure and sparsity.

The images shown in this paper were generated using the quadratic programming formulation (5). We used a modified conjugate gradient method to solve the constrained optimization problem, because its behavior is well known theoretically. An efficient implementation is currently under development, and will likely involve the use of multigrid methods, which are among the fastest known algorithms for the solution of the sparse linear systems obtained from the discretization of partial differential equations or variational problems. Good results were also obtained using the algorithm introduced by Frankle and McCann. In this method, a narrowband lowpass filter is computed efficiently using a cascade of convolutions with large but very sparse kernels; the constraint is incorporated by means of a threshold operation after each convolution. A multiresolution implementation which further reduces the computational cost was later introduced.

5. METHOD 2: BACKPANEL GENERATION BY CONSTRAINED OPTIMIZATION

In the algorithm described above, the constraints have an asymmetric nature, since they only include a lower bound for the backpanel. In presence of dark details on a light background, the constrained lowpass filter tends to “fill the hole” producing a uniform backpanel and the detail is reproduced on the frontpanel only. In case
of light details on a dark background, instead, the backpanel must follow the transition as shown previously in Figure 4, and on some critical images this variation in the backpanel can give rise to visible parallax error even if a narrowband lowpass filter is used. Artifacts can appear especially around synthetic parts of the image such as text or frames over a uniform dark background. An example is shown in Figure 6. Improved implementations of the constrained lowpass filter can reduce this problem, but not avoid it completely, since it is intrinsically present in the method.

Moreover, the constraints themselves can be chosen in a more clever way. By using a constrained lowpass filter, the backpanel and the frontpanel are bounded, from below and from above respectively, by the square root of the input image as shown in (4). These constraints can be relaxed in order to exploit some further degrees of freedom without compromising the feasibility of the solution.

Finally, by using a constrained lowpass filter to blur the backpanel, it is guaranteed that the frontpanel image will not exceed the white level. An examination of Figure 4, however, shows that the frontpanel computed in this way presents some undershoots near sharp edges; in some cases, these pixels may fall below the black level and dark details will be lost. In order to prevent also this artifact, some sort of double constraint must be used.

In order to derive the exact constraints, we shall suppose that each pixel of the two panels can have a normalized transmittance between $1/d$ and 1, where $d$ is the dynamic range. The maximum transmittance of an LCD panel is actually much lower (0.33 for the IDTech grayscale panels used in the prototypes), but we will use normalized units for simplicity of notation. By definition, the backpanel does not saturate if the displayed image $L_b(x,y)$ is within the limits:

$$ \frac{1}{d} \leq L_b(x,y) \leq 1 \quad \forall (x,y) \quad (6) $$

If perfect reconstruction is requested, the frontpanel is computed by division. The constraints in this case are

$$ \frac{1}{d} \leq \frac{L_{in}(x,y)}{L_b(x,y)} \leq 1 \quad \Rightarrow \quad L_{in}(x,y) \leq L_b(x,y) \leq d L_{in}(x,y) \quad \forall (x,y) \quad (7) $$

Therefore, merging the two sets of constraints (6) and (7), we obtain that both panels do not saturate if the backpanel satisfies

$$ \max\{L_{in}(x,y), 1/d\} \leq L_b(x,y) \leq \min\{d L_{in}(x,y), 1\} \quad \forall (x,y) \quad (8) $$

If the dynamic range of the input image is greater than $d^2$, the constraints (8) become incompatible, therefore the input image should be pre-processed in order to remove any exceedingly dark pixels. This processing is automatically accomplished if the digital data of the source image are decoded following a DICOM grayscale display function which matches the dynamic range of the display device.

Based on the assumption that the parallax error is small if the backpanel image is smooth, the proposed method generates, by means on an optimization procedure, the smoothest possible backpanel image subject
to the constraints (8). Since the eye response approximately follows Weber’s law, we measure the perceived distortion introduced by a displacement in the backpanel image using the relative error:

$$E(x, y) \triangleq \frac{L_b(x + \Delta x, y + \Delta y) - L_b(x, y)}{L_b(x, y)}.$$  

(9)

The backpanel we are seeking is a smooth function and the displacements $\Delta x$ and $\Delta y$ are small; therefore, we approximate the expression (9) with a first-order Taylor expansion:

$$E(x, y) \approx \frac{1}{L_b(x, y)} \left[ \frac{\partial L_b(x, y)}{\partial x} \Delta x + \frac{\partial L_b(x, y)}{\partial y} \Delta y \right] = \nabla L_b(x, y) \cdot \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix},$$  

(10)

where $\cdot$ indicates a scalar product. In order to obtain a single scalar which can be minimized by an optimization procedure, we compute the mean square norm of the gradient term in the right hand side of Equation (10), thus obtaining the following expression in which the identity follows from simple calculus:

$$E_{\text{mean}} \triangleq \int \int |\nabla L_b(x, y)|^2 dx dy = \int \int |\nabla \log L_b(x, y)|^2 dx dy.$$  

(11)

In this approach the backpanel is generated \textit{ex novo} by an optimization procedure. The proposed method theoretically produces an optimal result, since the objective function (11) measures the parallax error and the constraints (8) are strictly derived from the device specifications. The method is also easy to use, because it does not contain any free parameter which requires a manual adjustment. An example of its behavior on a 1D test signal is shown in Figure 7. The algorithm produces a constant backpanel wherever possible; if the dynamic range of the input signal is less than $d$, the signal is displayed on the frontpanel only and the parallax error is completely avoided. In the bright areas, the lower bound is active and the backpanel takes a value $L_b(x, y) \geq L_{\text{in}}(x, y)$; in the dark areas, instead, the upper bound is active and the backpanel takes a value $L_b(x, y) \leq d L_{\text{in}}(x, y)$. The step in the backpanel is therefore approximately $d$ times smaller than the step in the input signal; it can also be noticed that the lower bound in (8) is lower than the square root of the input image, so there is a better chance that the algorithm produces a smooth backpanel. If no constraint is active, the backpanel forms a linear slope (in logarithmic scale) in order to minimize the functional (11).

![Figure 7. Second proposed method applied to a 1D test signal. Left to right: input signal, backpanel and constraints (shaded), frontpanel. The vertical axis is in logarithmic scale for a better legibility](image)

On 2D images, the algorithm performs in a similar way. The backpanel computed with the proposed method is similar to a membrane-based envelope, with the difference that the membrane is not attracted to an input signal (mathematically, this means setting $\lambda = 0$ in (5)) and its shape is only determined by the constraints. In the areas where no constraint is active, the backpanel satisfies the Laplace equation.

6. RESULTS AND OPEN PROBLEMS

The algorithm described in Section 5 can not eliminate completely the parallax effects if the input image contains sharp edges which have a greater magnitude than the dynamic range $d$ of the panels. In this case, due to the
presence of the upper bound in (8), the frontpanel alone is not able to completely reproduce the edge, and a fraction of its magnitude must be transferred onto the backpanel. An edge in the backpanel becomes visible in case of parallax error. An example is visible in Figure 8. The source image contains sharp edges and isolated black pixels; if perfect reconstruction is requested, a very dark pixel can only be reproduced when both panels are dark and the degrees of freedom allowed by the constraints are strongly reduced. On the other hand, the first proposed method does not consider an upper bound; in some cases, perfect reconstruction can be lost.

Figure 8. Example of splitting. Left to right: original image, backpanel, frontpanel.

Since the upper bound was introduced to guarantee perfect reconstruction and detail preservation, we must deduce that these requests are incompatible with that of parallax error reduction. A possible solution to this problem is to relax the perfect reconstruction constraint (2) and allow some distortion in the visualized image. By observing Figure 8, it can be noticed that the dark pixels typically carry little information content; if this were true in general, one could filter the input image before the splitting in order to remove isolated black pixels and limit sharp edges. It is also known\textsuperscript{10,11} that the high dynamic range of the human visual system is mainly due to local adaptation, and if a scene contains high contrast boundaries the details near the edge appear blurred and indistinct. Because of this limitation, in some cases a reconstruction error near a sharp edge might not be visible. Research must be done to verify if this is acceptable with medical images, and if it can give rise to temporal artifacts when the algorithm is applied to video sequences.

7. CONCLUSIONS AND FURTHER WORK

In this paper we presented a prototype of a high dynamic range display based on dual layer LCD technology. The prototype exhibits a dynamic range of over 4 orders of magnitude, surpassing the performance of conventional radiographic film and standard displays. The two panels, however, have physical and optical constraints which require a complex algorithm for the creation of the images used to drive each of them. In this paper we proposed two methods which guarantee a perfect reconstruction of the image when the two panels are aligned, and reduce in most cases the parallax error which is intrinsically present in the dual layer technology, thus allowing an off-axis view where possible. Due to the limited dynamic range of the single panels, the algorithm output must satisfy appropriate constraints in order to prevent any unwanted clipping; we derived the exact constraints which achieve this goal without any loss of degrees of freedom. The overall approach we propose is original, since the backpanel is not computed by filtering the input image, but is generated \textit{ex novo} by an optimization procedure.

The prototypes and the image processing algorithms are still under development. We are currently working on an improved algorithm which attempts to overcome the physical limitation of HDR displays described at the end of Section 3 by taking into account the limited sensitivity of the human visual system. Also the problem of DICOM compliance\textsuperscript{12} in high dynamic range displays is under study.

ACKNOWLEDGMENTS

This research was partially supported by grants of the University of Trieste and of the Regione Friuli-Venezia Giulia within the “Eladin” project.
The authors thank Mr. Silvio Bonfiglio of FIMI-PHILIPS for the helpful discussions and encouragement throughout this work.

REFERENCES