Far Field Imaging Through Liquid Crystal Displays for Biometrics

Timothy Large*, Neil Emerton*, Sehoon Lim*, Chris Nuesmeyer*

*Microsoft Corp.
One Microsoft Way, Redmond, WA, 98052

Abstract
A camera is placed behind a standard liquid crystal display with a backlight that is designed to retain a near-infra-red image, while simultaneously delivering diffuse illumination in the visible spectrum. The camera image can be used for system log-on or continuous authentication, without occupying valuable bezel space.

Author Keywords
Imaging through display, biometrics, authentication

1. Introduction
Modern computing systems utilize face recognition systems like Windows Hello to enable rapid and easy user authentication. However, there is simultaneously a demand to reduce bezel sizes in computing systems to better utilize display real estate, and at the same time there is increasing pressure to use the bezels for other functions, such as RF antennas. We have demonstrated previously that image recovery through OLED panels is realistic [1 & 2], here we demonstrate recovery of image information through an LCD.

In standard LCD construction, the high opacity of the LC panel and backlight components (surface relief diffusers, prism films, light-guide, and reflector films) prohibits through-screen imaging. We demonstrate that it is possible to construct an LCD device using a standard IPS liquid crystal panel along with a novel backlight, for face authentication through the display.

Face authentication systems fall into three general categories; systems that utilize unstructured illumination and IR images; systems that utilize structured illumination and 3D reconstruction; and systems that use dual cameras and 3D reconstruction. In general, systems that use IR illumination are preferred because they are less dependent upon ambient illumination.

In this case we use unstructured illumination and detect faces using Weiner deconvolution, and a face detection algorithm based on three independent components: face detection, pose measurement and facial recognition. We demonstrate that the image is of sufficient quality to recognize and differentiate faces despite the visual opacity of the display system.

2. Outline Optical System Design
The optical system consists of a standard LCD and camera, with a unique backlight construction to allow an infra-red image in 950nm wavelength region to pass. The experimental system is shown in Figure 1.

A standard sRGB IPS LCD (we used a 201ppi IPS panel manufactured by Sharp Corporation, shipped in Microsoft Surface Laptop) has sufficient transmission in the infra-red spectrum, though the black mask creates Fraunhofer diffraction that is imposed on the image.

To substantially reduce attenuation and scatter in the infra-red resulting from the backlight components, while maintaining screen gain in the visible region, we used a film set from 3M Company called NITS (near infra-red transmission system), that

was originally designed for fingerprint imaging in LCDs. A remaining challenge however is created by the lightguide, as modern light guide plates typically have a complex surface micro-structure intended to optimize the backlight output. In this case, we need the light-guide plate to be substantially transparent and have minimal light extraction features, without inducing pixel level non-uniformity in the displayed image.

Figure 1. Experimental IR transmissive LCD.

3. Liquid Crystal Panel Test Results
The panel pixel structure is shown in Figure 2, where the left-hand image is under visible illumination, and the right hand under 950nm IR.
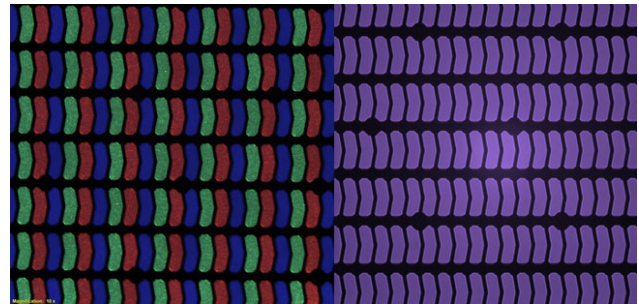

Figure 2. IPS LC panel viewed in transmission in visible (left) and 950nm IR (right).

The low duty cycle of obscuration in the IR is critical. As reported in [3,4] periodic obscuration causes loss of transmitted information. If the obscuration duty cycle exceeds 50%, some spatial frequencies in the captured image are entirely lost.

The measured spectral transmission of the panel over wavelengths of interest is shown in Figure 3. The transmission rises steeply across the 800-900nm region, because this is the absorption band edge of the dyes in the polarizer and color filters. In the 900-1000nm region, the remaining significant loss is the black matrix of the LC panel and the diagonal touch sensor traces. The panel has an overall luminous transmission of 6.8% and an IR transmittance of 47%.

Figure 4 shows the effect of diffraction, calculated from the black matrix shown in Figure 2, and measured using a laser. The structure of the diffraction pattern is very close to that expected from the mask, though the sub-harmonics are slightly different.
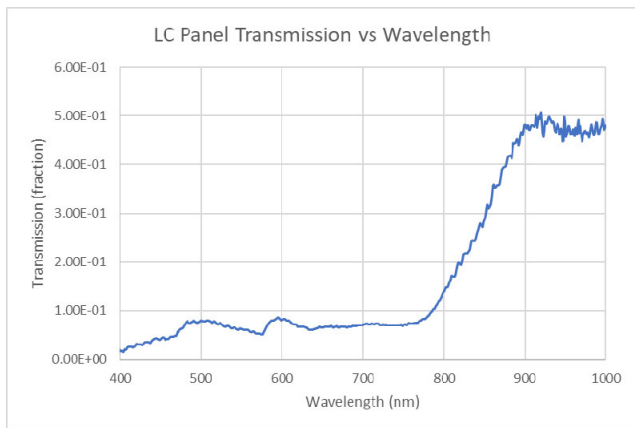

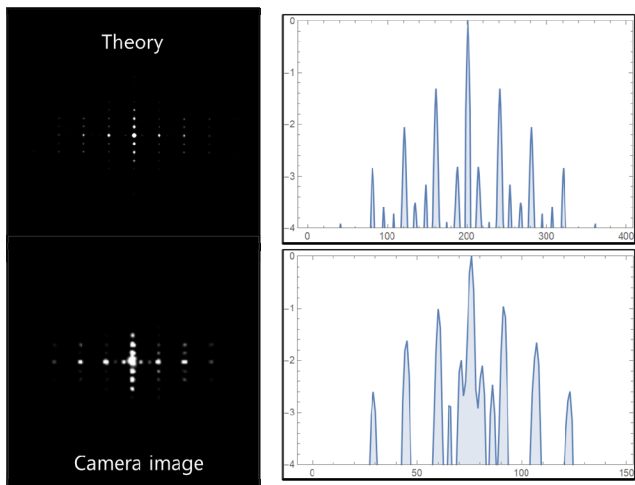
Figure 3. Measured IPS LCD transmission spectrum

Figure 4. Measured IPS LCD diffraction orders.

4. Backlight Structure

The backlight structure is shown in Figure 5. The nomenclature of NITS films is -R for reflector, -D diffuser, and -C for collimator [5].
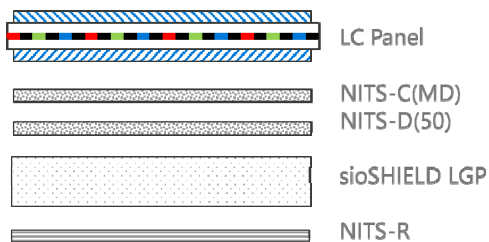

Figure 5. Backlight structure

A transmission image of the sioSHIELD light guide plate (LGP) is shown in Figure 6, the extraction features are highlighted in

black because they scatter light outside of the microscope objective numerical aperture. Note that the light guide plate extraction structures are sparse (which reduces interference with the through image), and the distribution is partially randomized (to reduce intensity fluctuations in the visible displayed image). The lightguide is made from UV embossed PMMA sheet 0.5mm thick.
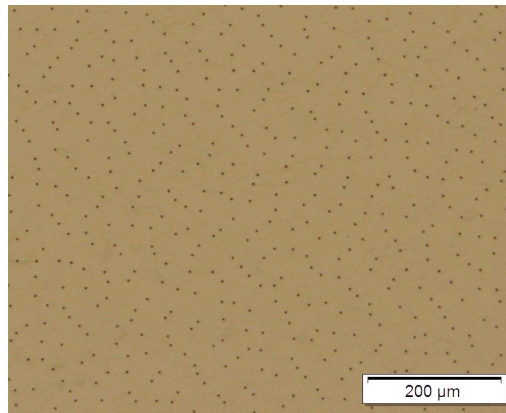

Figure 6. sioSHIELD lightguide extractor structure

The backlight structure contains components that add some diffusion to the through image (lightguide plate and NITS-D, but no structure that produces additional regular diffraction, so the diffraction imposed on the image is substantially from the LC panel itself.

5. System performance

The completed display system has a good balance of IR transmission and screen gain. The performance is shown in the following Figures. Figure 7 shows the measured transmission of the backlight alone, and the backlight and LC panel together.
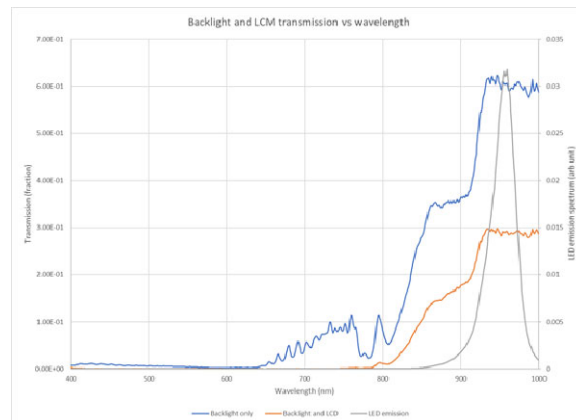

Figure 7. Overall system transmission. LED spectrum is overlaid.

Figure 8 shows the luminance profile of the display. The full width half maximum is ~55 degrees in both horizontal and vertical. Although the light distribution from the lightguide is fairly close to Lambertian, the NITS-C film increases the axial luminance by reflecting some of the off-axis illumination back into the lightguide cavity. The axial luminance is ~80% of the value of a standard backlight module driven with the same LED power.
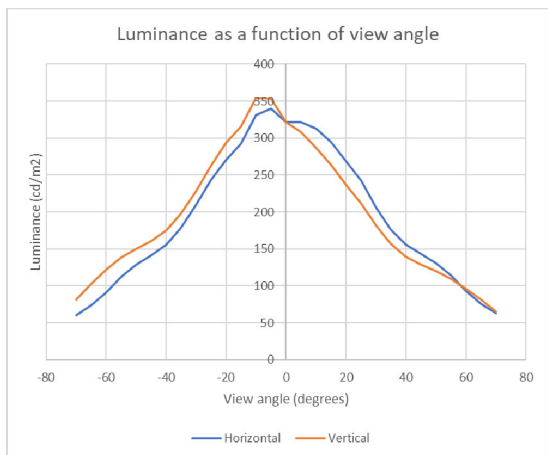
Figure 8. System luminance characteristic.

The luminance distribution of Figure 8 is normalized to the standard LED drive current of 20mA per LED.

The color performance of the display is shown in Figure 9. It is noted that further optimization of white point is required to meet D65 which would be normal for a display of this type.
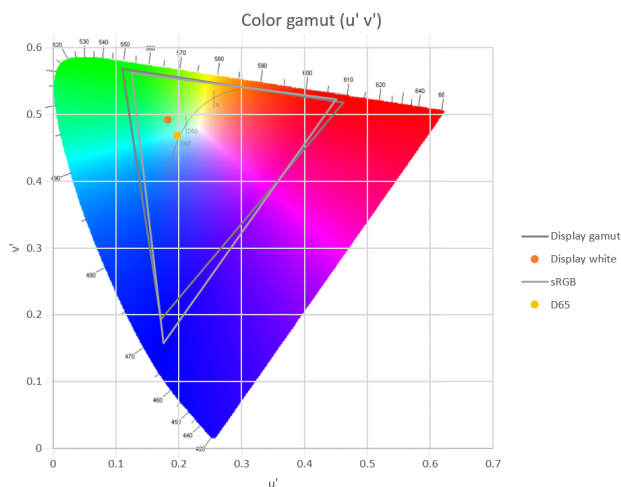

Figure 9. System color characteristic.

6. Camera system power budget

Ideally, the IR LEDs should produce sufficient illumination to fully illuminate the scene at the target distance, such that a white object will produce a full signal at the camera sensor.

In this experiment, a commercial camera (Ximea MOQ042-RG) is used. The camera contains a CMOSIS CMV4000 sensor with 5.5x5.5um pixels. The sensor is oversize for this application and a region of interest with 1024x768 resolution is used for face recognition.

The IR system power budget is shown in Table 1. The illuminance for the LED source is used to calculate the optical power incident on a camera pixel, given various loss factors, and assuming that the target is a Lambertian re-radiator. This is compared to the full-well charge capacity of the sensor. The results of calculating the optical power incident during an LED pulse show a good match.

Table 1. IR system power budget

Calculate from LED Radiance				Calculate from camera response		
Irradiance at 0.5m	3.51	W/A/m2		Photon energy	2.08E-19	J
(measured)	710	uW/A/cm2 two LEDs		Well depth	13500	e
	3.55	W/A/m2 one LED		Full signal energy	2.81E-15	J
Target reflectance	0.8	fraction		QE	0.07	fraction
Target distance	0.47	m		Pulse length	4.00E-03	s
Obliquity factor	0.93			Optical pwr at pixel	1.00E-11	W
Target radiance	0.89	W/A/str/m2				
2 LEDs @1.7A	3.04	W/str/m2				
NA of lens	1.4	number				
Focal length	0.009	m				
Lens pupil diameter	0.0064	m				
Lens pupil area	3.25E-05	m2				
Subtended angle	0.000147	Str				
Pixel area	3.03E-11	m2				
Projected pixel area	8.25E-08	m2				
Display transmission	0.3	fraction				
Optical pwr at pixel	1.11E-11	W				

The LEDs are modulated with 4ms pulse duration, matching the global shutter time of the camera. Although it is feasible to use an optical filter on the camera to eliminate visible light leakage from the backlight entering the camera, in this case the backlight is modulated at 50Hz in anti-phase to the camera.

7. Image Processing

The camera image shows the effect of diffraction from the LCD as shown in Figure 4, and some random diffusion from the backlight films as previously discussed. In order to remove some of these artifacts, the system point spread function (PSF) is measured using a point source. Using the measured PSF, the image is recovered using Weiner deconvolution. The original and deconvolved images are shown in Figure 10, with the measured modulation transfer functions shown in Figure 11. MTF is measured using slant-edge test method.
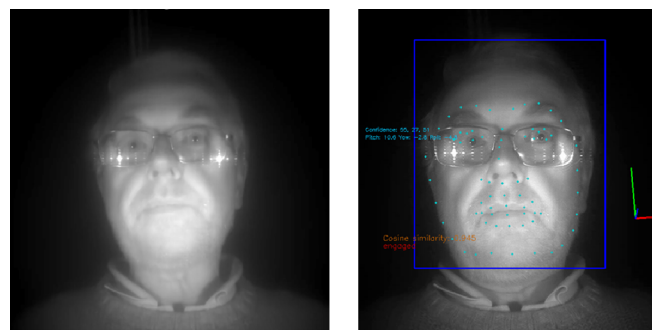

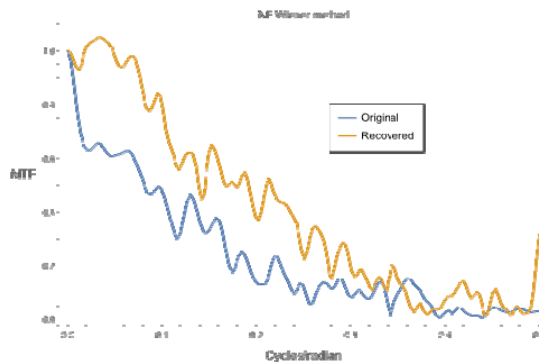**Figure 10.** Original and deconvolved images.

Figure 11. Original and deconvolved measured MTF

The face recognition system consists of three components. Firstly, the face is detected in the scene; secondly, the pose of the face is estimated; thirdly the face is recognized.

The face detection and face pose models are custom, lightweight models based on the Mobilenet architecture [6]. The face pose model jointly estimates rotation angles and landmarks.

The landmarks are used to align the face using a similarity transform before being passed to the face embeddings (recognition) network. The face crop that is input to the face embeddings network is first normalized to zero mean and unit variance which is helpful for low contrast images.

The face embeddings model is based on the DenseNet-161 architecture [7] and trained on tens of thousands of unique identities and millions of images using adaptive margin loss [8].

The face pose and cosine similarity are reported for individual video frames, based on comparison to a previously enrolled face. Cosine similarity is a measure of the difference between the enrollment face and test face in feature space.

In Figure 12, the cosine similarity is plotted against detected face pose, for a number of different LED power levels. We can see the algorithm is somewhat sensitive to pose and the cosine similarity metric falls slightly as the face is rotated. Eliminating pose outside +/-5 degrees, we can see the cosine similarity function of camera signal level in Figure 13.

We can see that the ability of the system to produce reliable detection metrics is a function of signal level, and the IR signal should ideally be maintained to more than 40% of the design value.
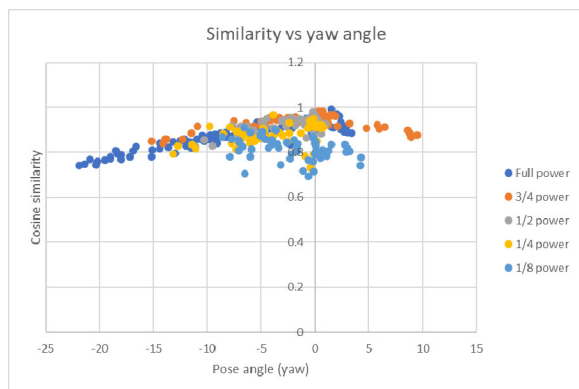

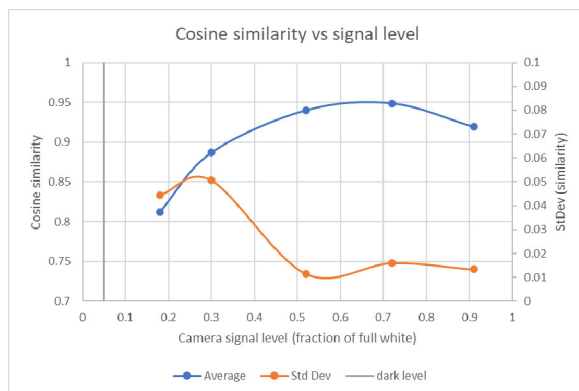**Figure 12.** Cosine similarity vs yaw for different LED illumination levels.

Figure 13. Cosine similarity vs signal for a white target.

Figure 14 shows the cosine similarity that is calculated for two different faces. As we expect, when we test the face of Author 2 against enrolled face of Author 1, the cosine similarity is very low, indicting that the face is not recognized.
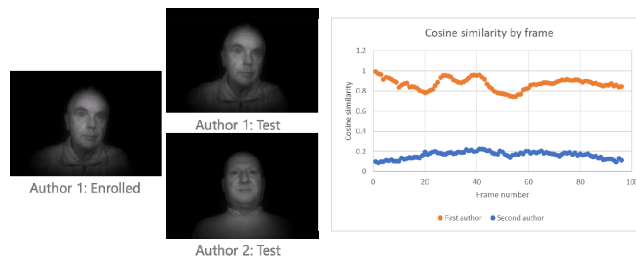

Figure 14. Cosine similarity power level for different faces.

8. Conclusions

We have demonstrated that it is possible to produce far-field images of sufficient quality to allow face authentication through a liquid crystal display, with limited compromise to the displayed image, using a camera behind the LCD operating at 950nm. Face authentication is demonstrated using an image processing pipeline consisting of deconvolution, face detection, pose estimation and finally face recognition.

9. Acknowledgements

10. References

1. S. Lim, Y. Zhou, N. Emerton, T. Large, and S. Bathiche, "Image Restoration for Display Integrated Camera," SID, Display Week, 74-1, (2020).
2. S. Lim, L. Liang, N. Emerton, T. Large and S. Bathiche, "Smart Display Clearly Sees via Deep Learning", Proc. IDW, Vol 27, 2020 977-980
3. Quang Tang et al, "Study of the Image Blur through FFS LCD Panel Caused by Diffraction for Camera Under Panel" SID Displayweek 28-2 (2020)
4. N. Emerton, D. Ren, T. Large, "Image Capture through TFT Arrays" SID Displayweek 28-1 (2020)
5. Hideaki Shirotori, Quinn Sanford, Jennifer Chen; "Near Infrared Transmission System for Under LCD Sensor Displays" International Display Workshop INP1-2 (2020)
6. A. G. Howard, et al, "Mobilenets: Efficient convolutional neural networks for mobile vision applications". arXiv:1704.04861, 2017
7. G Huang, Z Liu, L van der Maaten, K Weinberger; "Densely connected convolutional networks" CVPR, 2017, pp. 4700-4708
8. Hao Liu, Xiangyu Zhu, Zhen Lei, Stan Z. Li; "AdaptiveFace -Adaptive Margin and Sampling for Face Recognition",CVPR, 2019, pp. 11947-11956