State of Flat-Panel Display Technology and Future Trends

DAVID E. MENTLEY

Invited Paper

The rivalry between the cathode-ray tube and flat-panel displays (FPDs) has intensified as performance of some FPDs now exceeds that of that entrenched leader in many cases. Besides the well-known active-matrix-addressed liquid-crystal display, plasma, organic light-emitting diodes, and liquid-crystal-on-silicon displays are now finding new applications as the manufacturing, process engineering, materials, and cost structures become standardized and suitable for large markets.

Keywords—Cathode-ray tube (CRT), CRT market, electroluminescent displays, flat-panel display (FPD), LCD market, liquid-crystal display (LCD), microdisplay, near-eye systems, organic electroluminescence (OEL), organic light-emitting diode (OLED), plasma display panel (PDP), projection systems, Self-scanTM, supertwisted nematic technology, television, thin-film transistor (TFT).

I. INTRODUCTION

The lure of the electronic display has drawn the attention of scientists, engineers, inventors, and consumers since the early 1900s. As information technology moved from painstakingly slow generation and permanent storage on bulk materials such as stone and paper to nearly instantaneous production and transfer through pervasive electronic and optical networks, an electronic display has become a requirement at billions of termination points. Consider the installed base of computers, television (TV) receivers, and cell phones. It is at these terminals that the human/machine connection is made and of the two main senses—audio and visual—the visual mode is clearly able to provide the more robust information transfer. This explains in large part the tremendous amount of effort that has been applied to the pursuit of the perfect electronic information display and, more specifically, the perfect flat-panel display (FPD). The perfect electronic display has yet to be invented and may not be if history is any guide, for we will continue to employ different display types for applications according to the price and feature requirements.

In retrospect, there have been very few singular display inventions that have punctuated the evolutionary path of all of the current major electronic display types. For the most part, display technology has pushed ahead as a result of many clever engineers and scientists studying the art and initializing new solutions in the form of materials, process, circuit design, architecture, or optics to solve problems and increase manufacturability.

Displays, being the complex psychoelectrovisual interfaces that they are, have many dimensions in the form of features, characteristics, and properties. Examining these properties for the various types of displays is a lengthy task. In conclusion, the important primary characteristics of an electronic display are:

1) contrast ratio;
2) brightness (luminance);
3) speed;
4) resolution (dots/inch);
5) color depth and gamut;
6) freedom from artifacts.

Other important characteristics and, in fact, shortcomings in these areas may preclude a particular type of display from being used in some applications are:

1) manufacturing cost;
2) price;
3) lifetime;
4) power consumption;
5) operating temperature range;
6) viewability in sunlight;
7) weight;
8) volume.

The list can be extended even further to account for areas that may not be obvious to an initial observer:

1) voltage requirements;
2) efficiency;
3) storage temperature range;
4) electromagnetic emissions;
5) electromagnetic sensitivity;
6) toxic material content.

The search for the ultimate FPD has always been influenced heavily by the existence of a nearly perfect role model—the cathode-ray tube (CRT). Except for the fact that the CRT is voluminous, which typically increases geometrically with the screen diagonal size, the CRT exhibits many of the ideal characteristics of an electronic display.

II. CATHODE-RAY TUBE

The development of electronic displays began with the CRT, which was first explored in the 19th century, although the observation of a glow from the electrical excitation of a gas in an evacuated tube may go back as early as the 17th century. The invention of the CRT device itself is generally attributed to Braun. The “Braun tube,” reportedly first built in Strasbourg, Germany, perhaps as early as 1896 [1], used both deflection and a fluorescent material for the display screen. It was probably the first application of the Crookes tube concept to a display, since Braun apparently was attempting to use the device as an indicator of the alternating current (ac) output of an electrical generator.

Real progress began when large-scale production of monochrome CRT-based TV commenced sometime around 1945. Color versions of the CRT were not necessary until color broadcasting was authorized in 1950. In a classic battle of wills and a wager on the future of display technology, the path chosen by Sarnoff at RCA for color TV relied on the shadow mask CRT, rather the electromechanical spinning of a synchronized color wheel proposed by CBS. Eventually, of course, the CRT became the standard technology for TV during the second half of the 20th century [2].

The first successful shadow mask color picture tube was developed in 1951 by a team of researchers at RCA's Research Center under the direction of Edward W. Herold [3]. This research resulted in a high-quality low-cost color display that launched the TV industry complete with the accompanying social changes and the development of mass production techniques for color CRTs during the late 1950s and throughout the 1960s (see Fig. 1). This manufacturing expertise was eventually applied to the computer display industry. This kind of application-technology matching has been a hallmark of every successful display type. More specifically, there has been a rigid matching of appropriate technology and an application driven feature-set for every high-volume display. The TV, digital watch, calculator, computer monitor, and portable computer are prime examples.

Before exploring the display characteristics, it is helpful to step back and compare the basic philosophical differences between the CRT and the generic FPD. The CRT has evolved from the triode structure of the Braun tube, which was the first scanning version of the device in the late 1890s. It then became a highly complex and robust design capable today of high-definition TV or better. The CRT took advantage of a gift of nature that enabled it to achieve high performance at low cost at a time when most electronic systems were controlled by switches, relays, and vacuum tubes and were built by macroscopic manufacturing techniques. This gift is the third dimension. By employing the third dimension as a means to control both addressing and writing of information on a two-dimensional (2-D) viewing screen, the CRT is able to use relatively simple architecture and control electronics to display high-performance video and data in a low-cost and highly manufacturable component. A single electron beam (or three in the case of a color tube) is scanned across the interior surface of the display (addressing), controlled by two inputs (X and Y electrodes) while information is sent over the modulated beam current (writing).

The fact that this was possible 40 years before similar performance was available from a FPD is testament to the ingeniously simple design of the CRT. The flaw is that the performance is intimately tied to its three-dimensional (3-D)
construction. Various attempts to squeeze the traditional 3-D CRT into two dimensions or at least reduce the depth substantially have been unsuccessful to date. This is because the mechanical complexity increases far beyond the manufacturability allowed by current plant and equipment standards.

The quest to replicate CRT performance in a 2-D device has cost tens of billions of dollars in both research and development and plant construction. The best hope so far is the active-matrix-addressed liquid-crystal display (LCD). It is not widely known, but thin-film transistor (TFT) technology was first applied to a display as a control matrix to an electroluminescent display, not to an LCD. A thorough first-person review of the early days of the TFT was published by Brody [4]. FPDs are expensive, relative to CRTs, because they perform both important functions (addressing and writing) in two dimensions, while the CRT uses the third dimension essentially for free. Addressing a matrix of millions of points from within a 2-D structure demands a sophisticated selection scheme that avoids crosstalk and delivers high contrast at high speed. The writing of information likewise demands sophisticated electronics that can accurately place precise quantities of electrons in the right place at the right time. This task implicitly demands a device or material with a fast nonlinear switching characteristic. A diode can perform the job, but a transistor performs much better.

The writing and addressing functions can be accomplished in a FPD by several nonelectronic means: thermal, chemical, acoustic, mechanical, temporal, optical, and various combinations thereof. Thermal and chemical addressing and writing have not been successful to date, typically due to reversibility or lifetime issues. Acoustically addressed displays have never become commercially widespread due to complexity and efficiency problems. Mechanical methods were employed early in the development of displays and lost favor until the recent arrival of microelectromechanical modulators that have been coupled with a color wheel. Electronic addressing and writing is intuitively the best hope for taking advantage of the same phenomena that drive the semiconductor business, but the present state of materials development demands that some other optical writing phenomenon be added to the electronic addressing function.

Within the universe of the electrooptical systems, other broad divisions must be drawn. First is the split between projection displays and direct view or near-to-the-eye displays. The CRT is the ultimate example of an emissive display. Light is generated only where it is needed, yielding extremely high contrast and it can even be concentrated in small areas to give extra “punch” when desired. This feature, along with full analog control of the intensity makes the CRT, match the human visual system very closely, so that it provides a highly realistic and film-like image. Plasma displays, inorganic and organic light-emitting diode (OLED) displays, and electroluminescent displays also are emissive and share some of the features of the CRT. The drawbacks are that emissive displays usually require a high voltage or high current and are not often viewable in outdoor locations. OLEDs have challenged the voltage problem by reducing driving voltage to around 10 V and inorganic light-emitting diodes are now visible in direct sunlight in large video displays.

LCDs as well as electromechanical and chemical display types are light reflecting or modulating systems. Color, contrast, viewing angle, and overall appearance are typically sacrificed in a reflective display. These displays can be used in reflective designs to operate with a few tenths of a microampere per pixel. Transmissive color LCD designs provide a much higher level of quality since they include an internal light source and a color filter subsystem tuned for brightness, color, and contrast. However, new emissive displays, such as the OLED display, may be a natural evolutionary step beyond such transmissive and reflective systems.

Looking back at the history of electronic displays, it can be seen that the CRT became the benchmark for onscreen performance. Various FPD technologies have since evolved to challenge and in some cases exceed the goals of the past. It has always been clear that no single type of display can address every requirement—form, function, economics, availability—and, therefore, the display industry still produce many different options. This situation is not likely to change until the perfect technology comes along, which is not on the horizon at this time.

The next section will review the current status of some of the main display technologies and predict the evolution of these technologies over the next five years. Since each display technology is covered in extreme technical detail in this special issue, the emphasis will be upon comparing technical features, strengths, weaknesses, and market fit.

III. LIQUID-CRYSTAL DISPLAYS

Although liquid crystallinity was first observed in 1888 by Reinitzer, it was more than 30 years before Mauguin [5] discovered and described the twisted-nematic structure that later became the basis for LCD technology. During the 1920s and 1930s, work on liquid-crystal materials and the electrooptic effects that they produced was conducted in France, Germany, Russia, and Great Britain. The first patent on a light-valve device, which used liquid crystals, was awarded to the Marconi Wireless Telegraph Company in 1936 [6]. It was not until the 1960s, however, that serious studies of the materials and the effects of electric fields on these devices were carried out. The first use of a liquid-crystal device as a display dates back to Williams and Heilmeier at the David Sarnoff Research Center, Princeton, NJ, in 1963 [7], [15]. The group at RCA discovered important effects such as dynamic scattering, dichroic-dye LCDs, and phase-change LCDs. In the late 1960s, both Ferguson and Schadt/Helfrich developed displays based on the twisted-nematic LCD effect, which is still employed in most of the active-matrix and direct-drive LCDs made today. It was a long road from the simple watch displays of the early 1970s to the popular-selling full-color LCD desktop monitors of today. While the first LCDs looked exciting and promising, the limitations soon became apparent. Alt and Pleshko pointed out that root-mean-square voltage-responding displays dependent upon an
electrooptic curve that was not very steep would not be able to deliver useful contrast with a large number of addressable rows [8].

One by one, every technical challenge posed by the CRT standard has been met and in many cases exceeded. Operating lifetime was one of the first problems. This was solved with purified materials and alkaline ion barrier layers on the substrate. Contrast ratio, which was lacking in highly multiplexed displays, was improved markedly through super-twisted nematic technology in 1985 [9]. Response time was the next serious obstacle to attaining CRT performance. Determined efforts on all drive electronics, physics, and material-science fronts eliminated response time as an issue. High power consumption plagued backlit LCDs in the early years and notebook-computer makers complained that the displays drew too much current from limited batteries. Power consumption for a typical notebook computer fell from about 20 W in 1990 to less than 4 W in 2000. This improvement came about by increasing the aperture ratio, fine-tuning the color filter and lamp spectrum, designing better light pipes to distribute light more efficiently, and recycling light that would otherwise be wasted. Viewing-angle limitations fell as remarkable new architectures produced LCDs with nearly paper-like appearance. At the end of 2001, we could find excellent LCD monitors on the shelf in any major retail electronics store and the market exceeds $20 billion (see Fig. 2).

However, despite its impressive performance, the conventional color transmissive LCD architecture employs a convoluted electrooptical path. This path starts with the electrical generation of an ultraviolet-emitting plasma. This is followed by a conversion to red, green, and blue wavelengths and a distribution of white light over a large area. The polarization of the white light and selective polarizing of specific areas provides further filtering to individual red, green, or blue spatial elements. This leads to the final blocking or transmission of the individual color. This is the process employed in every notebook computer or flat-panel desktop LCD monitor. This complexity has stimulated the search for a more intuitively simple approach to making FPDs.

IV. PLASMA DISPLAYS

The plasma display panel (PDP) might be considered as an outgrowth of the neon lamp, invented in 1915 by Claude in France, although the current structure has added a complex wavelength conversion subsystem to provide color. The display is often referred to as a gas-discharge display. The basic mechanism of operation is fairly simple: inert gases, such as helium, neon, argon, xenon, or mixtures among these, hermetically sealed in a glass envelope are subjected to a high-voltage electric field that causes the gas to ionize and, thus, produce plasma. The plasma generates ultraviolet light, which excites a phosphor, which in turn emits visible red, green, and blue wavelengths. The electric-field excitation may be either an ac- or dc-plasma type. Weber offers a good historical account of gas discharge [10].

Although Bell Laboratories used a gas-discharge display to demonstrate the first transmission of TV pictures in 1927, work in the development of PDPs did not receive much attention until the early 1960s. In 1964, researchers at the University of Illinois developed the first ac-driven memory display that they called the “plasma display panel” in a report published in 1966 [11], [16]. Five years later, workers at Owens-Illinois [12], [17] reported on the development of an improved version of the PDP with a longer life. Owens-Illinois eventually commercialized the display, which they dubbed, the DIGIVUE display.

The next major development occurred in 1970 with the introduction of the self-scanTM display, which was described in detail in 1972 [13]. This display, manufactured for several years by Burroughs Corporation, greatly reduced the number of circuit drivers by making the display act as a shift register. Around the same time, Fujitsu developed the self-shift display [14], which also reduced the number of circuit drivers.

Monochrome plasma displays acquired a reputation for reliability and later for low cost, but the market for monochrome displays was generally ceded to LCDs in the 1980s. In the 1990s, color-plasma technology was pushed forward by major research efforts in Japan, notably at Fujitsu, Matsushita, and NHK, as well as, Pioneer, Hitachi, NEC, and...
Mitsubishi. In the latter part of the 1990s, dozens of companies around the world were pursuing every aspect of plasma display research in an effort to bring real products to the market.

Plasma technology has finally made the crossover from development to production stage. PDPs no longer need to be delivered with apologies for weak luminance, poor contrast, or low resolution. Standard products now look stunning and rival direct-view CRTs, rear projection CRTs, and direct-view and projection LCDs. Some direct-view LCD makers continue to threaten intrusion into the 30–70-in diagonal market space that plasma technology claims. The two remaining characteristics that need improvements are luminous efficiency and lifetime.

Plasma-display luminous efficacy, measured in lumens per watt, has been gradually increasing. Most plasma panels operate at a maximum of 1 lm/W. Power increases relatively linearly as the relative luminance increases up to a peak of 100%. This contrasts with the LCD, which operates continuously with nearly the same power consumption level whether a single pixel or all pixels are at maximum luminosity. The level of 1 lm/W is nominally acceptable, but an industry goal is 5 lm/W and early laboratory reports indicate that it has been demonstrated. Increased efficacy translates to longer operating life at reduced power levels with less heat waste and higher luminance.

As an emissive display technology, PDPs possess many of the advantages of the CRT, such as wide viewing angle and excellent color. Furthermore, PDPs are true flat panels, so they do not suffer from the bulkiness of projection or CRT systems. A key limitation is the high price, which remains the main obstacle to market expansion. Prices for plasma displays have already begun to drop dramatically and will continue to do so as the volumes increase and costs decline.

V. ORGANIC LIGHT–EMITTING DIODE DISPLAYS

The OLED is progressing along one of the most rapid development paths seen in recent display history. Building on two key technologies—semiconductor processing and LCDs—OLED design and production is moving from the laboratory to the prototype stage at an impressive rate. The LCD, in comparison, was not produced in active-matrix form within a decade of its first demonstration.

Organic electroluminescent devices operate exclusively in the diode mode. The preferred term in Asia for the light emission phenomenon is organic electroluminescence (OEL) and the devices are called OEL devices. In the United States, the more common term is OLED, although OLED may also stand for organic light-emitting device. OLED displays are pixilated arrays of OLEDs. The light-emitting materials in OLEDs may be small molecules or polymer materials.

Tang and Van Slyke of Eastman Kodak first reported light emission from small-molecule organic systems in 1987. The company went on to patent this materials technology and has since licensed it to other companies, which have continued its development. In the past few years, many new small-molecule OLEDs have been discovered and refined, considerably diversifying the playing field. The first OLED product, a monochrome car stereo display, reached the market in 1997.

Conjugated polymers were found to have similar light-emitting qualities in 1990 by Burroughes, working in the research group of Richard Friend at Cambridge University. Cambridge Display Technology was founded on the basis of the resulting patents. However, as with small-molecule technology, many other polymer materials have come onto the scene recently.

The year 2001 has witnessed advances in all areas of OLED technology. Both small-molecule and polymer materials are available with emissions in all color regions of the spectrum. Lifetimes and efficiencies are several times better than they were a year ago. New materials have been devised to achieve better charge transfer at the interfaces between layers. Device architectures have improved to increase light output and fabrication of thin layers can be done more reliably. The first standardized passive OLED drivers have been introduced and several arrangements of TFTs for active-matrix panels are being explored. Manufacturing development has focused on inkjet printing for polymers and controlled deposition through masks for small molecules.

The number of companies and research institutions working on OLED materials, devices, components, or production techniques has increased during the past year. The total now tops 100, a sign of the continued robustness of the growing OLED industry. Despite the speed of laboratory development and the number of players investing in OLED, the development of a manufacturing infrastructure will require time. The greatest challenges for the near future lie in developing this infrastructure. The OLED display offers many very important advantages over other types of flat panels and in particular, the LCD, its prime competitor. As an emissive technology, it offers superior viewing angles and contrast ratios to the LCD. Its response time is also faster, which may play a role when OLEDs move into active-matrix applications.

OLED displays also offer the potential for manufacturing cost savings. The LCD process is notoriously expensive and complicated, requiring numerous deposition and patterning steps. Furthermore, moving to each larger substrate size requires a complete retooling and costs more than half a billion dollars. Although cleanliness is also very important in the OLED process, fewer layers are required and photolithography may not be necessary. Polymer materials can even be applied by an inkjet printing process. OLEDs are also amenable to advanced manufacturing methods, such as roll-to-roll processing, which cannot be used for LCDs.

However, OLEDs are not ready to displace LCDs from the market in the near future because challenges remain with their commercialization and successful incorporation into mass-market products. Although OLED manufacturing offers big potential gains, achieving those gains requires overcoming some substantial hurdles. No standardized processing and test equipment exists for OLEDs. Machines, sensors, and other equipment are still being designed and
developed. Lines are yet to be built and manufacturing techniques optimized. Much of OLED production is still in an immature state.

Other hurdles include technical development issues such as materials stability, particularly in high-temperature environments. Other problems include achieving full color and integration of appropriate current-based drivers for OLEDs.

In the long term, OLED devices face both technological and business issues. While the former may be solved through sufficient investment in research and manufacturing infrastructure, the latter must be addressed in parallel if the OLED device is to succeed in the marketplace. Displacing firmly entrenched display technologies such as CRTs and LCDs will require strategies above and beyond demonstrating superior technical performance from OLED devices.

VI. MICRODISPLAYS

Microdisplays are loosely defined as sub-1.5-in displays that are viewed indirectly either through magnifying optics near to the eye or through a projection system that puts a large high pixel-count image on a screen. Displays with high information content currently have from several hundred thousand to more than 1 million pixels. The two most compelling features of microdisplay technology are the ability to display large numbers of pixels and to do so in a lightweight package that occupies a small space. Traditional display technologies tend to become larger, heavier, and more expensive as image size and pixel count increase. The fact that microdisplays can be highly compact without sacrificing image quality gives them an advantage when competing with existing products. In addition, these features of microdisplays enable new products that could not be served by existing technologies, such as a wearable monitor in the form of a lightweight headset.

Some microdisplay technologies also have the potential for lower manufacturing costs than other types of displays. The materials cost is low, due to the use of a small area of standard silicon, and the incorporation of drivers, interface, and controllers onto the display backplane. The substitution of optics for a large display area can also conserve material. Microdisplays that employ sequential color enjoy the benefits of using inexpensive LEDs as the light source, requiring no color filters, and having only one-third as many pixels as are needed for a full-color display not using sequential color.

Two methods used to implement microdisplays in products are:
1) projection systems, which magnify a small, real image onto a screen for viewing by one or more users;
2) near-eye systems, which use an optical system to project a virtual image, intended to be viewed by a single person (and, hence, are also called personal viewers). These systems are best suited for compact hand-held or head-worn devices.

Most microdisplays are fabricated on silicon or quartz substrates rather than on glass, as in the production of direct-view FPDs, a process that has both technical and business consequences. By building the display directly on semiconductor substrates, designers can integrate electronic components—such as row and column drivers and digital and video interface and control circuits—directly alongside or under the display on the same substrate. This can lead to higher performance with lower manufacturing costs. On the business side, this has led to production arrangements that more closely resemble those of the semiconductor industry than those typically seen in the display industry. Rather than require new investments in production facilities that use glass substrates with an increasingly large area, microdisplays can be fabricated on existing semiconductor equipment, leading to the possibility of flexible arrangements between display developers and semiconductor firms.

The four types of microdisplays are the following.
1) Reflective, which modulate an external light source by varying the properties of a reflecting surface. Liquid-crystal on silicon and microelectromechanical systems devices are examples of reflective microdisplays.
2) Transmissive, which modulate an external light source as it is transmitted through the device. Polycrystalline silicon LCDs are transmissive designs.
3) Emissive, which produce light internally. OLED on silicon, electroluminescent, and vacuum fluorescent devices are emissive displays that have been put into the microdisplay format.
4) Scanning, which write images directly onto the retina. Some LED and laser light sources coupled with a scanner are also used in microdisplay implementations.

Microdisplays, especially those made on silicon, offer tremendous potential that is not available in other parts of the display industry. Specifically, a silicon backplane microdisplay holds the potential to bring Moore’s law to the display industry. In the future, multi-million pixel displays could be delivered without the scale-up problems associated with active-matrix TFT displays, CRTs or other flat panel types. Much work remains to be done, though, since it is not simply an electronics issue, but a complex interaction of optics, materials, packaging, and chemistry.

VII. CONCLUSION

While active-matrix LCDs are indisputably attractive and functional, the economics of making the displays are not so attractive. The problems stem from the huge investment to install a new factory to make active-matrix LCDs—on the order of one billion dollars for the latest meter diagonal sheet size—in combination with two other important elements. One is the lengthy planning and startup cycle in a quick moving end-product market and the other is the very high material content in the cost of the finished product. These characteristics have led to the situation where manufacturers are often forced to sell at prices that are near or below costs. The approaches to solving the cost problem include reducing the capital equipment needs. The use of flexible substrates can help in this area. Also, simplifying the manufacturing process through new designs and materials and integrating more subsystems, such as polysilicon backplanes, which can reduce driver, interface, memory, and con-
troller costs. OLED displays also have the potential to reduce the display complexity substantially by eliminating the backlight subsystem, color filter, and polarizers, among other elements.

Display technology has evolved impressively in the last 105 years. We can only speculate about what the next ten years will bring. We already have photorealistic FPDs with over 300 dots-per-inch resolutions. Inevitably, the future display is portrayed as a wraparound screen with indiscernible pixels. Certainly we will have the petabit per second or exabit per second conduits to deliver content. Whether the displays will be as evolved is an open question. There are no limits to the creativity of the display developers or to the tools available for discovery and development. The challenge is to find the right combination of display performance and manufacturing cost to make FPDs attractive and affordable to billions of people, which in turn enrich their lives through technology.

ACKNOWLEDGMENT

The author would like to thank Dr. K. Allen and Dr. J. Castellano for help in writing this paper.

REFERENCES


David E. Mentley received the B.S. degree in ceramic science from Alfred University, Alfred, NY, in 1973, and the M.S. degree in materials science and the MBA degree from the University of California at Berkeley in 1976 and 1981, respectively.

He is currently a Senior Vice President with Stanford Resources, San Jose, CA, an iSuppli, Inc., company. His current research interests include developing, marketing, and analyzing electronic displays and display markets.