1

Introduction

1.1 Microdisplays

Advances in several technologies related to projection displays and near-to-eye (NTE) displays have intensified interest in these areas. Development of the microdisplay is a key technology. The term “miniature display” does not distinguish small displays, such as watch displays, from displays designed with magnification in mind as indicated in Figure 1.1.[1] Microdisplays are a natural extension of the familiar microfilm, where magnification is essential to readout. A coarse image is recognizable on a microdisplay, but the full resolution is only discernable with the aid of magnification. The term “microdisplay” now in general use means a compact display designed for use with a magnification system. Microdisplay projectors have taken over the consumer market from CRT projection displays that have long dominated large-screen television. NTE products such as camera viewfinders and head-mounted displays evolve with microdisplay developments.

The direct-view CRT display has grown bigger and better over the years. Competing flat panel technologies such as liquid crystal and plasma displays have followed a similar path. Conversely, microdisplays have become smaller, in step with the shrinkage in microelectronics. NTE displays have progressed from miniature CRTs to lightweight microdisplays with superior characteristics. Attractive image quality and improvements in viewing comfort have opened up a commercial market for NTE displays. Figure 1.2 shows an early head-mounted display (HMD) designed for the consumer market, employing dual liquid crystal microdisplays. The unit weighs 150 gm and simulates a 52-inch diagonal color image viewed from a distance of 7 ft with resolution 260 × 346 × 3 RGB pixels.

The vacuum tube nature of the CRT has made it difficult to expand in scale over the years, and now appears to be saturating at a cumbersome 40-inch diagonal. Projection displays powered by special-purpose compact CRTs overcame the size limitation. In comparison, the simple slide projector provided a much better image, and obviously with the invention of a suitable “electronic slide” (microdisplay) would be the basis of a new video projection technology. A wide variety of electro-optic devices has struggled to fill the microdisplay role in projectors. The current choices,

Introduction to Microdisplays D. Armitage, I. Underwood and S-T. Wu

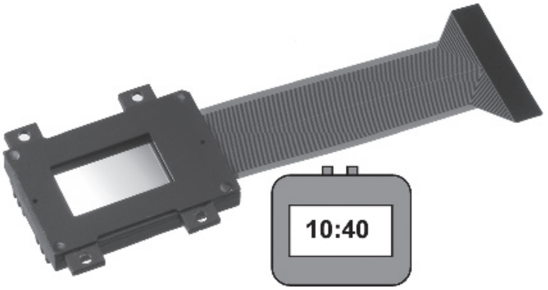

Figure 1.1 Miniature watch display compared with microdisplay having 9 µm pixel pitch, 1920 × 1080 resolution giving active area 0.78-inch diameter. Reprinted courtesy of Sony Corporation

liquid crystal and micromechanical devices, have proved commercially viable and will be difficult to displace. Price continues to exert a downward pressure on microdisplay area, due to the device cost and the influence of area on system cost. Figure 1.3 shows a portable business projector weighing 2.4 lb, with dimensions 2.6 × 6.1 × 7.8 inch, employing a single micromechanical microdisplay of 1024 × 768 resolution, diagonal 0.7 inches (DMD™), projecting full-color 1500 ANSI lumens at contrast ratio 1100:1.

Rear projection TV is a huge consumer market that drives the development of microdisplays. The latest products employ folded optic systems that reduce the unit depth to about 8 inches eroding the

Figure 1.2 Head-mounted microdisplay GT270. Reprinted courtesy of Canon Inc.

Figure 1.3 Microdisplay portable business projector LP70+. Reprinted courtesy of InFocus Corp.

space advantage of flat panel displays. The highest quality products produce the best images of any display technology. Figure 1.4 shows a 65-inch diagonal rear projection TV, employing three XGA LCOS microdisplays: luminance 450 cd/m^2, contrast ratio 2000, weight 95 lb.

1.2 Human Factors

The display engineer needs some knowledge of human vision to understand display performance and comfort for a given application. An engineering specification of human vision provides details of resolution, sensitivity, response time, and wavelength dependence. Luminance measured in candela/m^2 (lumen/steradian/m^2) is a measure related to our sensation of brightness. The lumen is a photometric unit of light flux derived from the product of the eye's wavelength sensitivity and radiant power,

Figure 1.4 Microdisplay rear-projection TV model BR768HC. Reprinted courtesy of Syntax-Brillian Corp.

giving the visual response to different wavelengths directly. At peak wavelength sensitivity 555 nm, flux in lumens = 683(radiant flux in watts); the same radiant flux at any other wavelength is worth fewer lumens, falling to zero lumens outside the visible wavelength range. Ambient light, contrast, and color saturation also influence our sense of brightness.

1.2.1 Color

Perception of color depends on wavelength in a complicated way that allows three primary wavelengths to represent a wide color gamut according to the luminance values of the primaries. Red, green, and blue primaries (RGB) optimize the color gamut. An even wider gamut of colors follows from the addition of more primaries. The narrower the spectral range of each primary color, the more saturated the primary color becomes.

Reduction to three primaries is of great utility in electronic displays, where picture elements (pixels) grouped in RGB triads below visual resolution merge to provide a full-color display. Color pixels are standard for direct-view CRT, flat panel displays, and some microdisplays. One of the options in projector design is to superimpose primary color images on the screen, to achieve full color. Each primary has a dedicated microdisplay, requiring three microdisplays for this parallel color system. Alternatively, color sequential display systems exploit the eye's response time by presenting the primary colors in rapid succession to give the perception of a single color represented by the primaries. One microdisplay can handle the sequence of colors, but it must operate at high frame rate. Eye movement during the color sequence will cause some separation of primary colors on the retina, perceived as color breakup. Eye movement sets a lower limit on the color field rate needed to suppress color breakup.

Color temperature is a measure of the spectral distribution of a light source, by comparison with the temperature of equivalent black body radiation. A color temperature of 6504 K represents average daylight. An arc-lamp source can be adjusted to a given color temperature by spectral filtering, with some sacrifice in output. Primary color separation requires filtering, where strong color saturation incurs further loss. Some compromise between color temperature, color saturation, and luminance is required. Color vision is lost at low light level $<10^{-4}\,\mathrm{cd/m^2}$; display engineering is essentially concerned with substantially higher luminance levels to provide good color vision.

1.2.2 Resolution

Display costs generally increase with resolution, making it wasteful to provide display resolution beyond viewer requirements. Light level, modulation depth, and wavelength of the signal influence eye resolution, peaking at maximum sensitivity 555 nm. The response to one-dimensional sinusoidal luminance patterns characterizes eye resolution as a function of spatial frequency (u), luminance (L), and contrast (C).

$$\text{Contrast ratio} \quad CR = \frac{L_{\max}}{L_{\min}} \tag{1.1}$$

$$\text{Contrast} \quad C = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} = \frac{CR-1}{CR+1} \to 1 - \frac{1}{CR} \tag{1.2}$$

$$\text{Perception boundary} \quad L = L_{\mathrm{av}}\left[1 + C_{\min}\sin u\right] \quad L_{\mathrm{av}} = \frac{L_{\max} + L_{\min}}{2} \tag{1.3}$$

The contrast sensitivity function $S(u, L_{\mathrm{av}}) = 1/C_{\min}$, where $C_{\min}$ is the minimum contrast required to perceive the sinusoidal fringes at spatial frequency u and average luminance L_{av}. Contrast sensitivity plots reveal the most sensitive spatial frequency and upper frequency cutoff. Figure 1.5 shows contrast

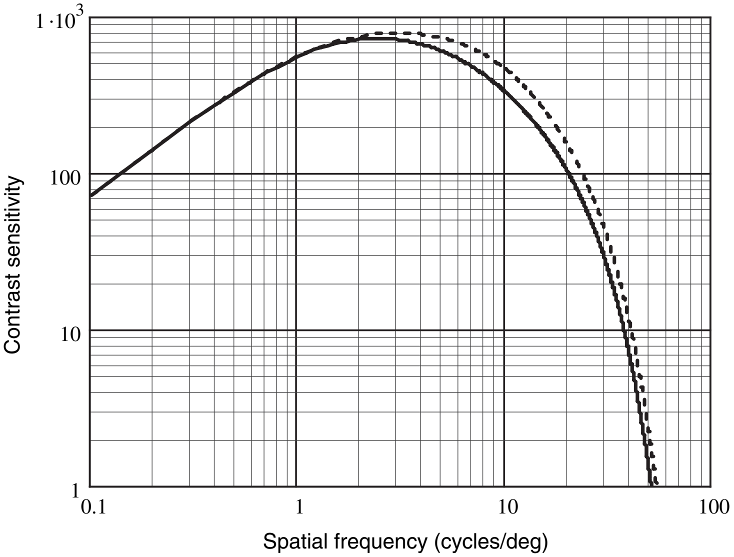

Figure 1.5 Contrast sensitivity for luminance: solid line 100 cd/m^2; dotted line 900 cd/m^2

sensitivity plots appropriate to displays. The plots are generated from formulae derived from a signal processing description of human vision, with parameters fitted to reported data.[2] Peak contrast sensitivity is about 4 cycles/deg, and cutoff about 60 cycles/deg, consistent with a limiting angular resolution of 0.5 minute of arc applicable to edge detection, and comparable with an accepted average value of 1 minute for human visual acuity. Increase in luminance provides a small increase in resolution at display luminance levels. The frequency cutoff corresponds to foveal vision, which is limited to about 2 degrees field of view. Eye and head movement compensate the limited field of view of the eye at high resolution, creating the impression of high resolution over a wide angle. Restriction of natural eye and head movement has a disturbing effect on vision that can result in eyestrain and discomfort.

The lower sensitivity of non-foveal vision contributes to the decline in sensitivity at low spatial frequency, where the spatial wavelength extends beyond the foveal range. Further low-frequency attenuation occurs at the neural processing level. Gradual variation in luminance of order 50% from the center to boundary of a display may not be noticeable, while shifts less than 1% are obvious at sensitive spatial frequencies. Color uniformity is similarly sensitive to spatial frequency. Moreover, we are more sensitive to color change than luminance change at low spatial frequency, making color uniformity more difficult to achieve in display systems.

Correlation in the visual process, known as hyperacuity, allows perception of an object's positional accuracy well beyond visual acuity. Vernier acuity is the ability to detect misalignment in lines and objects, which can be an order of magnitude higher than visual acuity. Spatial aliasing associated with pixelation gives rise to jaggy diagonal lines, detected by vernier acuity before perception of pixelation in general. Reducing pixel size below visual acuity may not suppress all the effects of pixelation.[3]

1.2.3 Flicker

Fluctuations in display luminance give rise to an annoying flicker. Frame-by-frame writing of electronic displays requires sufficient frame rate to avoid flicker. The early movies picked up the name flicks due to inadequate frame rate. With increase of flicker frequency, flicker attenuates, approaching

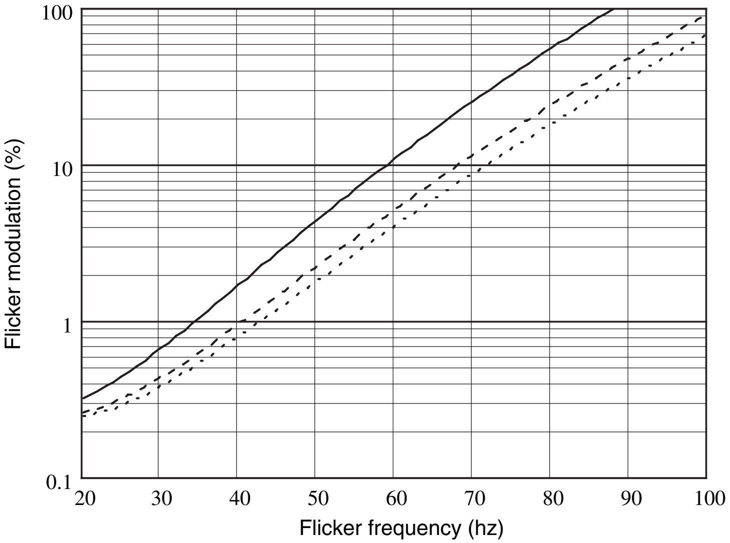

Figure 1.6 Threshold modulation for 50% probability of seeing flicker on a full white screen. Average luminance: solid line 100 cd/m^2; broken line 500 cd/m^2; dotted line 900 cd/m^2

threshold perception at a critical flicker frequency (CFF). Threshold perception is the level at which 50% of the population will detect flicker. Experimental studies of the dependence of flicker perception on luminance, modulation level, and frequency enable the elimination of display flicker.

The model for human vision referred to earlier generates Figure 1.6, on inserting typical parameters for threshold perception.[2] The flicker modulation m is of sinusoidal form $[1 + m \cdot \cos(2\pi wt)]$, at frequency w hertz. For non-sinusoidal flicker, m is the value of the fundamental Fourier component. The flicker parameters generating Figure 1.6 assume the worst case of a completely white screen at full brightness, a severe test of flicker compared with typical video projection. Perceptible flicker modulation decreases with increase of luminance, demanding a higher flicker frequency for suppression. An adjustment of the model parameters will give plots for a lower probability of flicker perception, predicting display conditions to eliminate flicker for essentially the entire population.

1.2.4 Contrast Ratio

Contrast is an important characteristic of image quality; low-contrast images have a washed-out appearance. Contrast sensitivity plots such as Figure 1.5 show that high contrast is required to achieve limiting eye resolution. High contrast ratio in all primary color channels promotes strong color saturation, and eliminates low-level color distortion, e.g. true black rather than dark purple.

The eye adapts to changes in average luminance, enhancing the visibility of dark scenes in a movie. To maintain image quality in dark scenes requires adequate contrast ratio at low luminance, implying high contrast at full luminance. Cinema-quality imaging requires $\mathrm{CR} > 1000$ to reveal detail in dark scenes.

Ambient light reflecting from the display surface sets a minimum luminance level, requiring higher display luminance for a given CR. Screen luminance <60 cd/m^2 is accommodated by low lighting in a cinema supporting $\mathrm{CR} > 1000$. Monitor displays designed for typical office lighting favor luminance of 150 cd/m^2 or higher, and light reflected from the screen may degrade $\mathrm{CR} < 50$. Outdoor displays exposed to sunlight demand extreme luminance, prompting the development of reflective displays that exploit the high luminance environment. Shielded viewing inherent to NTE displays makes them insensitive to ambient light.

1.2.5 Grayscale

Analog displays such as CRT and LCD are continuously variable, making grayscale accuracy an issue of stability and noise. The advantages of digital systems result in a digital video signal converted to analog form to drive an analog display. Digital look-up tables (LUTs) assure the correct analog drive level, but digital bit-depth limits the number of gray levels. Making any display part of a digital system imposes grayscale quantization. Displays with only on/off luminance capability are inherently digital and achieve grayscale by pulse width modulation (PWM), averaging to the desired gray level over a frame period. Binary pulse code modulation generates equally spaced luminance levels over the luminance range; a poor match to the eye with an approximately logarithmic response to luminance, requiring progressively increasing steps in luminance for a perceptually uniform grayscale.

An inadequate number of gray levels gives rise to a contouring artifact in regions of near uniform luminance (e.g. image of sky), where the minimum quantized step is visible as an edge depicting the luminance contour.[4] A similar effect in color quantization appears as color contour boundaries, described as posterizing. At a given light level, fewer than 100 gray levels are required to avoid luminance contouring, provided the gray level steps are perceptually uniform, i.e. logarithmic. To provide adequate grayscale in dark scenes requires enhanced gray level count; cinema quality demands 1000 perceptually uniform gray levels.

Video image L_{in} transmitted with video signal voltage encoded $V_{\mathrm{s}} \propto L_{\mathrm{in}}{}^{0.455}$, to match typical receiving CRT luminance characteristic $L_{\mathrm{out}} \propto V_{\mathrm{s}}^{2.64}$, gives overall $L_{\mathrm{out}} \propto L_{\mathrm{in}}{}^{1.2}$. The display gamma ($\gamma = 2.64$) is adjusted to the viewer's preference of image gamma. The image gamma determines the distribution of grey levels in an image, where adjustment of gamma allows the viewer to optimize the appearance according to taste and viewing conditions. Video image gamma is typically set to 1.2, consistent with dimly lit viewing conditions; however, choice varies with image as well as individual.[2] It is interesting that fidelity in display engineering is usually overridden by viewer preference in choice of gamma. A similar distortion appears in choice of color temperature. Our vision evolved in response to daylight, making that the natural choice. However, given an adjustment in color temperature we will generally make color temperature higher than daylight, enhancing the blue region of the image spectrum.

1.2.6 Viewing Comfort

Adequate luminance and resolution are basic requirements for any display. The viewer's eyesight determines the upper limit on resolution. Tolerance to lower resolution depends on the information content of the display. The eye adapts to a wide range of luminance, making acceptable luminance dependent on ambient lighting and spurious display reflectance. Obviously, the display should be well engineered and in good working order, free from flicker and noticeable distortion. Direct-view displays should be set slightly below eye level, at a distance about two or three times the screen diagonal. Image quality is sensitive to viewing angle in some displays, and generally favors on-axis viewing.

Viewing comfort in head-mounted displays is much more sophisticated than in direct-view displays. Weight, size and balance have no counterpart in direct view, but are critical to HMD tolerance. Total-emersion HMDs concentrate vision on the display, shutting out all extraneous light to provide an artificial reality experience. It is very difficult to simulate normal vision effectively over a realistic field of view, including eye and head movement. Inadequate simulation gives rise to motion sickness and other discomforts known collectively as simulator sickness.

Partial-emersion HMDs allow some vision of the outside world, which preserves the viewer's orientation, and are tolerable for much longer periods, even when the outside view is restricted to peripheral vision. Eliminating the influence of head movement on the displayed image makes eye movement do all the work in scanning the image. A field of view in excess of about 35 degrees induces intolerable eye fatigue due to the scanning effort. Viewing discomfort is a difficult barrier to overcome for the HMD to gain general acceptance. The interest in wearable computers and portable internet displays provides an incentive for further development and innovation in HMDs.

1.3 Display Specifications

A display designer emphasizes resolution, luminance, etc., according to the targeted application.[5] In marketing the display, a list of specifications identifies its value in various applications.

1.3.1 Resolution and Size

Modulation transfer function (MTF) has the same functional form as Equation (1.1) for contrast, and describes the decline in contrast with spatial frequency in cycles/mm = (line pair)/mm = lp/mm. MTF is the standard method of describing the resolution of optical components such as a projection lens, and the product of component MTFs gives the overall MTF. Display resolution determined by raster scan or pixel count cannot be expressed in MTF form without loss of mathematical rigor. The number of TV lines or pixel array size characterizes display resolution. The video graphic adapter (VGA) notation, listed in Table 1.1, identifies standard pixel array formats, extended by inclusion of the high-definition television notation.

Screen diagonal characterizes the display size, a legacy of the early circular screen CRT. The cost of a direct-view display increases more rapidly than the area, since the defect probability and assembly problems increase. Plasma displays have demonstrated the largest diagonal in excess of 100 inches, followed closely by LCDs. The size and performance of flat panel displays competes with projection TV. When manufacturing costs have settled down, price will determine the target market for projection, now set at >40 inches diagonal. TV displays viewed at a distance two or three times the screen diameter barely resolve (1 arc min acuity) pixels at WXGA resolution, while computer monitors viewed at a distance comparable to screen diameter just resolve WUXGA. Higher visual acuity of lines and edges favors extended display resolution. HDTV at 1080 × 1920 resolution may pull the viewer closer to the screen, intensifying the experience, particularly for sporting events.

1.3.2 Luminance and Color Saturation

In CRT projectors and plasma flat panels, the peak luminance cannot be maintained over the entire screen due to limited power dissipation. A peak luminance and an average luminance should be specified. Consumer CRT projectors deliver an average of about 200 lumens to the screen and require high screen gain for adequate luminance of about 350 cd/m^2. Projectors in general can take advantage of screen gain to increase the screen luminance in a preferred direction at the expense of lower luminance in other directions. The focusing and scattering properties designed into the screen determine its gain.

Table 1.1 Video graphic adapter designation and array size

VGA	640 × 480	SVGA	800 × 600	XGA	1024 × 768
SXGA	1280 × 1024	UXGA	1600 × 1200	QVGA	320 × 240
WXGA	1365 × 768	WUXGA	1920 × 1200	GXGA	2560 × 2048
SDTV	729 × 480	HDTV(720P)	1280 × 720	SHDTV(1080P)	1920 × 1080

Increased screen gain has the downside of enhanced speckle as well as reduced viewing angle. Front projectors specify lumens delivered to the screen, since the luminance depends on screen area and gain. Rear-projection units have built-in screens and luminance quoted over a range of viewing directions. Acceptable luminance depends on ambient lighting. Cinema projectors illuminate a large area screen to give about 60 cd/m^2 screen luminance, and require a dark ambient environment to appreciate image quality. Microdisplay projectors aim for 500 cd/m^2 or higher to accommodate the higher ambient light favored by the business and consumer markets.

Color saturation is strongest in laser or LED driven displays, with sharply defined wavelength. Dichroic filters in arc-lamp projection displays determine the color, where stronger color saturation implies lower throughput lumens. Specification of projector output luminance is sometimes inflated by quoting the value obtained before color correction. A similar tradeoff applies to LCDs using dyed pixels for color, where optical absorption introduces severe throughput loss. Phosphor characteristics limit the color saturation of plasma and CRT displays; however, color saturation is sometimes enhanced by addition of dichroic filtering in CRT projectors.

Backlighting power and throughput efficiency determine the luminance of LCDs. Improvements in throughput efficiency should improve the already high luminance of 500 cd/m^2. Plasma displays are marketed with small-area luminance beyond 500 cd/m^2. Microdisplays reduce cost by shrinking the diagonal to 0.7 inches or less. The development of small-arc projection lamps has kept pace with microdisplay contraction, maintaining the optical collimation necessary for efficient lumen throughput. We are entering an intense stage of competition for large-screen home theatre, where several technologies vie for consumer attention.

1.3.3 Contrast Ratio and Grayscale

Contrast ratio is an important indication of image quality. It is quoted for dark ambient; room lighting always reduces the CR. The largest degradation in CR is associated with diffuse reflecting surfaces, such as the powder phosphors in CRTs and plasma displays. Projectors quote the serial all-on/all-off CR, along with the ANSI CR for a white/black checkerboard pattern; ANSI CR is always lower than the serial CR, due to internal light scattering, and provides a better indication of image quality. LCD CR is limited by the extinction ratio of polarizing optics and off-axis retardation dependence; however, optical compensation achieves $\mathrm{CR} > 500$. Projection systems achieve $\mathrm{CR} > 1000$, due to higher quality polarization optics and better compensation over the limited field angle of the projection lens. A recent development expands the effective CR by modulating the light source according to the image's average light level, to maintain excellent CR and grayscale in dark scenes.[6]

The standard 8-bit grayscale is adequate for most purposes, and applied to each primary color channel gives 24-bit color. More demanding applications such as cinema projection require 12-bit grayscale/channel. Look-up tables create appropriate luminance grayscale steps, taking into account the device characteristic. Pulse width modulation grayscale, necessary in plasma displays and other digital display devices, achieves high precision in grayscale; however, image contouring and motion artifacts associated with PWM requires an effective expansion in addressing bit depth.

1.3.4 Response Speed and Flicker

Video displays need to respond at a fast enough rate to avoid motional blurring and flicker. We are somewhat forgiving of motion blur, since our vision is less acute in observing movement. Movies shot at 24 frames/sec are acceptable, which is not very challenging for an electronic display response time. However, cinematographers structure the scenes and camera angles to minimize artifacts such as false wheel rotation; a higher frame rate is desirable. Theatres shutter the film projector to raise the flicker rate to 48 Hz or higher to suppress flicker perception. Motion picture frame time is 42 ms, but the frame changes abruptly, implying equivalent display response $\ll$42 ms to duplicate film. A video data rate of 60 frames/sec attenuates flicker, but requires response $\ll$17 ms to maintain gray level integrity and avoid trailing on fast-moving images. The response time is severely challenged in color field sequential displays, where color is formed by a rapid succession of primary color frames above the color fusion frequency of the eye. Color frame rates as high as 540 frames/sec are required to avoid color breakup caused by eye movement.

The image refresh rate introduces an intensity modulation at the refresh frequency, and half that frequency if the display mechanism is susceptible to odd/even asymmetry as in the LCD. Flicker issues are resolved in display design by minimizing the flicker modulation and raising the refresh frequency. The displayed frame rate may be doubled to reduce flicker if there is significant modulation at half frame rate.

A static image displayed for a substantial period may store some aspects of the image that persists for some time as a 'ghost image' superimposed on the newly addressed image. The effect is described as ghosting or image sticking. Liquid crystal displays are susceptible to image sticking due to ionic charging effects. Plasma displays are also prone to ghost image effects.

1.4 Displays in General

To place microdisplays and applications in context we discuss electronic displays in general and the shrinkage of direct-view displays into the micro domain. We restrict our attention to high-information-content displays, since there is little value in magnifying low information content capable of direct display.

1.4.1 Cathode Ray Tube

The CRT has been the dominant technology for many years. Reliability and low cost have outweighed its shortcomings, until recent developments in flat panels. The simplicity of electron-beam addressing remains appealing, but carries vacuum-tube baggage that occupies valuable space, and has become unfashionable. The shape of the tube has improved over the years, including attempts at a flat structure, but does not compare with a lightweight flat panel display. The tradeoff in luminance verses resolution is a technical limitation that is proving difficult to surmount. Diffuse reflection from the phosphor powder penalizes CR under modest room light. The direct-view CRT will play a prominent role in displays for many years and gradual improvement in performance will continue. However, flat panel performance is improving more rapidly and will continue to win market share from CRTs.

The luminance/resolution issue is more marked in projection CRTs, where inadequate luminance is the biggest drawback. Improved phosphors and the application of thin film phosphors should enhance performance. However, competitive pricing of high-quality microdisplay rear-projection HDTV is pushing the CRT units off the showroom floor. The development of large-area flat panel displays competes with projectors in general.

Miniature CRTs were developed for NTE displays such as military helmet displays. As microdisplays, they played an important role in the development of head-mounted displays, and are still in use. Microdisplays with superior performance and advantages in weight and volume have made the miniature CRT obsolete. Detailed discussion of microdisplays in later chapters excludes CRTs, because they do not provide much insight into current microdisplay development.

1.4.2 Matrix Addressed Displays

Electrode-addressed displays, such as liquid crystal displays, must use multiplex addressing to reduce the wiring complexity. Matrix addressing uses a rectangular network of electrodes similar to (x,y)

Cartesian coordinates, to address a pixel located at the (x,y) intersection of the electrodes. A rectangular array of $M \times N$ pixels can be addressed by $(N + M)$ electrode lines, plus the common return line. Detailed analysis of matrix addressing shows that the electro-optic response of the pixel must be highly nonlinear to address a large number of pixels in an independent manner. Moreover, line-by-line addressing implies average luminance decreases with increase in addressed lines, unless the pixel is bistable. Devices such as plasma displays have nonlinearity favoring high-resolution addressing and bistability, accounting for their early success as flat panels.

Early LCDs adopted cell designs that optimized the performance for passive matrix addressing, at the expense of response speed and CR. Active matrix addressing developed to eliminate the need for nonlinearity in the electro-optic response. A field effect transistor at each pixel activates the pixel according to the (x,y) addressing signal, and isolates the pixel between refresh cycles. The development of amorphous-silicon (α-Si) thin film transistor (TFT) arrays enabled large-area active-matrix liquid crystal displays (AMLCDs), with nematic liquid crystal (NLC) cell design optimized for display performance.

A large-area display has plenty of peripheral space to make connection to the matrix electrodes, or accommodate chip-on-glass methods to reduce the external connections. However, the small electrode pitch ($\sim$10 µm) of a microdisplay discourages external connections and peripheral space is minimal. The best solution is to incorporate the addressing electronics with the electrode matrix. α-Si transistors are inadequate for the addressing electronics, where polysilicon or single-crystal silicon (c-Si) devices are favored. High-temperature polysilicon circuitry on quartz provided the first AMLCD microdisplays used in projectors. Silicon wafer technology is cost-effective in microdisplays, but favors reflective rather than transmission optics, e.g. the liquid-crystal-on-silicon (LCOS) microdisplay.

1.4.3 Field Emission Displays

Field emission displays (FEDs) employ cathodoluminescence similar to the CRT, but the electrons are field-emitted from multiple cathodes, rather than thermo-emitted. A vacuum tube environment is still required, but multiple cathodes support a flat panel structure. Each source pixel contains many cathodes, with emission modulated by control gates, and focused to activate a similar sized phosphor pixel. The pixels are matrix addressed by driving the voltage between control gates and cathodes. A revival of interest in FEDs is making slow progress in competing with established flat panel displays, and rapid advances in OLEDs. Early development of an FED microdisplay yielded little published data, and transformed into a successful OLED program. FED remains an interesting technology, but is unlikely to compete in the microdisplay arena. In keeping with miniature CRTs, the need for a high-voltage and vacuum environment is an overwhelming handicap in NTE applications.

1.4.4 Plasma Displays

Light emitted from a gas discharge, similar to neon lighting, is the basis of plasma display panels (PDPs). Designing the gas discharges to emit UV light that excites primary-color phosphors provides a full-color display. The outstanding feature of the PDP is large screen area, with diagonal beyond 100 inches, and still growing. It provides peak luminance greater than 500 cd/m^2, and high CR $>$ 1000 in a dark ambient. In ordinary room light, the CR may drop considerably due to optical scattering from the phosphors and plasma screening filter. Pulse width modulated grayscale and susceptibility to image burn-in are handicaps. The PDP competes in the business and consumer home-theater markets. The structural complexity to address and control the discharge has prevented the plasma display from shrinking to microdisplay dimensions.

1.4.5 Liquid Crystal Displays

Displays utilize the birefringence property of liquid crystals to modulate polarized light, or create strong optical scattering. High-resolution video displays favor polarization modulation, requiring the additional complication of polarizing optics. There are a number of nematic liquid crystal (NLC) cell configurations, such as twisted nematic (TN), and vertical aligned nematic (VAN), each having some advantage in a given application. A large electro-optic effect at a low voltage is the overwhelming advantage of liquid crystals, together with video response speed.

Liquid crystals modulate light, enabling operation in sunlight using a reflective configuration. Backlighting gives the best performance in ordinary room light, and pixel dyes provide primary color pixels to form full-color images. Pixel dye losses sacrifice optical throughput efficiency, which would benefit from improved color filtration methods. The gray levels are dependent on viewing angle, but various compensation schemes have evolved to minimize the off-axis image degradation. AMLCD covers a wide range of display sizes, from hand-held devices to large scale (>100-inch diagonal). Luminance $>500\,\mathrm{cd/m^2}$ is achieved, with $\mathrm{CR} > 500$.

AMLCD adapts to microdisplay dimensions by incorporating the matrix addressing circuits on the matrix periphery. High-temperature polysilicon on quartz provides circuitry for AMLCD projectors and head-mounted displays. The application of silicon chip technology to fabricate LCOS microdisplays achieves pixel dimensions $<10\,\mu\mathrm{m}$. LCOS promises lower-cost higher-performance projectors and NTE displays.

Ferroelectric liquid crystals (FLCs) have a faster response, but require PWM grayscale. The manufacturing process of FLC displays is more difficult to control over large areas, compared with NLC, and FLC has a narrower temperature range. The advantages of FLC in microdisplays are faster switching enabling superior color sequential performance, and a thinner cell promoting higher pixel resolution.

1.4.6 Electroluminescent Displays

A thin film of electroluminescent material supports a large electric field in response to a modest applied voltage. The high field releases and accelerates electrons to several electron volts, exciting the activator atoms hosted by the film, which emit light on decaying to their ground state. Alternating-current thin film electroluminescent displays (ACTFELs) have been available for some time and have an advantage in rugged applications. Improvement in materials promotes competitiveness in large-screen consumer applications. PWM grayscale is a disadvantage.

The solid-state thin film structure of ACTFEL is readily adapted to microdisplay dimensions. A monochrome microdisplay designed for head-mounted displays employed c-Si-on-insulator addressing. The addition of a liquid crystal color shutter provided field-sequential color at the expense of lower luminance.[7] The excessive voltage requirement handicapped further development. Progress in OLED microdisplays has prompted the termination of the ACTFEL microdisplay program. The decline of interest in the ACTFEL microdisplay precludes detailed discussion.

Light emitting diodes (LEDs) exploit electron/hole recombination radiation, making them the most efficient electroluminescent elements. Inorganic LED arrays have been demonstrated, but the performance and cost is not competitive with established high-resolution displays. Recent developments in organic light emitting diodes (OLEDs) have produced full-color direct-view displays and microdisplays. The acceleration in OLED development challenges the established display technologies across the board.

1.4.7 Electromechanical Displays

Electromechanical displays have a long history, from simple clocks to motion picture film projection, and laser scanning. Development of electron-beam scanning provided a means of addressing a micromechanical structure designed to project an image by optical deflection or diffraction. Finally, microelectronic matrix addressing controls an array of micromechanical pixels. The revival of interest in micromechanics gave rise to the terms micro-electromechanical (MEM) devices and systems (MEMS), which may have non-optical functions such as pressure sensing, compared to micro-optical-electromechanical systems (MOEMS).

Scanning laser displays use mechanical scanning systems in conjunction with intensity modulation such as acousto-optic modulation. A very low power version scans the image directly onto the viewer's retina, and can take advantage of low-power LEDs. Voltage modulation of the LEDs and MEMS deflection makes a compact NTE system with full-color imaging. We have omitted further discussion of simple optical scanning systems of either high power or low power, on grounds that they do not have a physical microdisplay and magnification. However, MEM devices that started life as two-dimensional arrays, but which have found application as one-dimensional arrays plus scanning, are included.

The digital mirror device (DMD) is the first MEMS device to achieve large volume production. The DMD is designed for projection use, where the tilting mirror pixels deflect readout light out of the projection lens aperture, modulating the light on the projection screen. MEM devices are often designed in analog form, but maximum projector throughput and gray level precision are achieved by on/off binary operation. The downside is PWM gray levels. The DMD is fast enough for field-sequential color, does not require polarized light, and has high throughput and CR.

DMD-type angular modulation requires collimated light and is not suited to direct-view displays. However, other MEMS structures may have more general application. A prototype MEMS modulator based on optical interference is being developed in direct-view form.

1.5 Microdisplay Evolution

Early television investment favored CRT displays that could satisfy the huge demand for household receivers. CRT projection could also provide larger images with some sacrifice in luminance. Alternative projection methods were of interest, but did not justify sufficient investment for near term success. Work continued at a modest pace to develop alternatives to CRT technology. The emerging optical signal processing and computing community needed a two-dimensional optical modulator.[8] They used the term "spatial light modulator" (SLM) to distinguish it from shutters and acousto-optic modulators. The SLM impresses two-dimensional information onto an optical beam, where further processing with other optical components and SLMs produces the desired output signal. The resolution, response speed, and wavelength of the SLM are system issues that can extend well beyond the limits of human vision.

SLMs provide the two-dimensional parallelism essential to optical processing or computing, and experiments proceeded with whatever devices came to hand. Military funding developed SLMs for specific applications and general-purpose SLMs, where performance rather than cost was critical. If a missile is heading your way, you will pay a lot more to compute its trajectory a little faster. SLM programs had the advantage of steady funding geared to small production runs. Commercial display device programs needed to consider progression to large-scale production and competitive pricing.

Early SLMs employing magneto-optical materials were unsuited to display development. SLMs based on electron-beam scanned addressing could take advantage of the CRT manufacturing base. Electro-optic crystals were obvious candidates for the optical modulating element. Crystals such as lithium niobate require high voltage ($\sim 1000\,\text{V}$) while lower voltage crystals under development required more electrical charge. The product of charge and voltage is approximately constant for all electro-optic crystals. Fabrication, testing, and analysis of beam-addressed crystals showed that the resolution is limited by electric field fringing in the crystal, implying crystal thickness $<100\,\mu\text{m}$. The device did not win favor for large-scale manufacturing, and languished as a potential high-power, high-cost projection component.[9] Electron-beam-addressed micromechanical devices showed more promise for consumer applications, but development could not surmount the manufacturing hurdle and lifetime problems.

The arrival of liquid crystals solved a number of problems. Low voltage and low charge simplified the addressing, while layers less than 10 μm thick improved resolution. They were not compatible with vacuum tubes, and photo-addressing became the favorite method.[10] A photoactive layer, such as cadmium sulfide, controlled the voltage applied to the liquid crystal according to the local addressing light intensity. Photo-addressing has intrinsic value in optical processing systems, but requires an optical input device such as a CRT to activate the photo address in display applications. A CRT coupled to a photo-addressed SLM with arc-lamp readout forms a light amplifying system for the CRT image. Projectors based on such light amplification schemes were marketed for many years, and gave rise to the term "light valve".[11] CRT photo-addressed liquid-crystal light-valve projectors have given way to electronic-addressed devices with similar performance at lower cost.

After the flat panel display community recognized the promise of liquid crystal displays, development of materials and device configurations quickly followed. However, the need for active matrix addressing slowed the progress towards displays with high information content. c-Si MOS technology could easily handle the electronic complexity of AM addressing on small-area devices. All the pieces were now in place for liquid crystal microdisplays. The first demonstration of MOS active matrix addressing employed the dynamic scattering effect in a nematic liquid crystal,[12] as did the later demonstration with integrated drivers.[13] Devices based on polarized light modulation replaced dynamic scattering, which soon become obsolete. The early displays demonstrated live TV images in 2-inch diagonal direct-view form, but the ability to shrink the scale was obvious. Addressing technology progressed to CMOS and the microdisplay became known as liquid-crystal-on-silicon (LCOS). LCOS microdisplays preceded the development of direct-view AMLCD displays, which awaited advances in α-Si thin film transistors.

The early LCOS devices suffered severe light leakage from the gaps between pixels, making them only of value in low-light-level applications such as camera viewfinders. An aluminum light shield under the pixels solves the problem, but the necessary multilayer metallization was not available in the early days. Chemical–mechanical polishing is a further advance in silicon technology that enhances the flatness and uniformity of LCOS microdisplays. LCOS will continue to benefit from advances in silicon wafer technology.

An alternative approach to electronic addressing employed a charge coupled device (CCD) structure to create a two-dimensional charge pattern, driven to the liquid crystal interface, in the same manner as photo-addressing. A dielectric mirror on the readout side shielded the silicon from output light, similar to the photo-addressed device. It had useful optical processing and projection applications, but could not compete in price with the developing AM addressed devices.[14]

While LCOS waited for advances in wafer silicon methods, flat panel demands pushed the rapid development of α-Si TFT addressing. The same companies were also interested in projection and soon produced the high-temperature polysilicon (HTPS) on quartz TFT for transmission liquid crystal microdisplays.[15] HTPS enabled the first AMLC microdisplay projector, and continues to support the LC projector market. Transmission devices enjoy simpler optical systems, but restriction in the pixel aperture sacrifices throughput compared with reflective displays.

The arrival of ferroelectric liquid crystals (FLCs) with enhanced response speed renewed interest in color sequential systems.[16] FLC microdisplays achieve color sequential with the help of pulsed LED readout in NTE displays. Further developments in FLC technology may provide sufficient speed for projector or flat panel color sequential applications.

Micro-optical-electromechanical systems (MOEMS) developed alongside liquid crystals, also benefiting from optical signal processing support. General interest in MEMS for a variety of applications other than displays generated considerable research activity and fabrication methods. Standard CMOS production methods can accommodate the integration of MOEMS with CMOS active matrix addressing. A complete microdisplay fabricated on wafer scale requires only testing and packaging to complete the production. Recognition of the advantages of digital mechanics led to the development of the digital mirror device (DMD), making this the only MEMS technology in projection display production.[17]

1.6 Microdisplay Applications

1.6.1 Projection Displays

The development of light valves governed the development of optical projection displays. The Eidphor oil-film projector and its descendants served the high-powered segment, while CRT/liquid-crystal systems had wider appeal. The arrival of AMLCD microdisplays captured the low-powered lower-resolution market. Microdisplay projectors gradually extended their range to higher resolution and higher power, essentially eliminating the earlier technologies. Liquid crystal and MEMS microdisplays now compete for projector market share.

The projection market should grow rapidly in the near future. The business projector market is well established and still expanding with a variety of projectors. Inexpensive, lightweight projectors with >1000 lumen output at XGA resolution suitable for conference rooms are available from many sources. Performance and price extend upwards according to auditorium demands. Top-of-the-line electronic cinema projection competes with established film projection, and must match or exceed film quality. Electronic cameras are beginning to displace film, making the entire movie system electronic, despite the wealth of artistic experience invested in film. The minimum specifications for electronic cinema projection are: 7000 lumen output, 1000 line resolution, 1000:1 CR, color equal to film, and faster response speed.

Rear-projection TV is the largest market. Microdisplay rear projection, already established, should take advantage of the move to HDTV. Acceptable performance requires about 500 lumens output at WXGA resolution, with color saturation and gamut comparable to CRT values. The consumer is sensitive to price, and the replacement cost of the arc lamp powering the display. The cost of the microdisplay is roughly proportional to area, but the cost of supporting optical components has a stronger dependence on area. The area should be minimized, consistent with delivering 500 lumens to the screen. Optical collimation and lifetime dictate a high-pressure mercury arc lamp. Such lamps (100 W power) have maintained screen lumens above 50% of initial value for more than 20 000 hours. The price of components will bottom with standardization of dimensions and performance, promoting extreme volume production.

High-power LEDs supplanting the arc lamp as the projection light source is an interesting development. Miniature projectors powered by LEDs are marketed, and home rear-projection TV has been demonstrated. LED illumination may dominate consumer TV production within the next five years, bringing enhanced saturation and color gamut, with lifetime beyond the product replacement cycle.

The tradeoff in performance and price to achieve consumer best buy status has yet to be established. Field-sequential color systems employ only one microdisplay, shedding cost and color convergence issues, compared with the three microdisplays of parallel color systems. However, the cost/performance tradeoff sacrifices throughput efficiency and image quality. Some of the loss and quality are restored by color band scrolling in a single microdisplay system. High-speed liquid crystal configurations exist for the single microdisplay option, but some increase in device area is favored to enhance throughput. The DMD projector employs a single microdisplay system for consumer markets and three microdisplays for high-performance markets.

1.6.2 Near-to-Eye Displays

Near-to-eye displays create an image for a single eye or pair of eyes, greatly reducing the output light compared with direct-view or projection displays. The light output is concentrated in an output pupil set by the magnification system. A large output pupil, or eye box, makes for easy viewing by relaxing critical eye positioning. Ambient light is not critical when masking the eye to receive little more than the desired image light.

NTE applications are invariably portable, such as camera viewfinders, mobile phones, and portable computers. Total power consumption is an important consideration in extending battery life. However, low light level relegates optical efficiency to a secondary consideration in video displays, where power lost in active matrix addressing is a primary consideration. In non-video applications such as text, diagrams, and charts, image storage (e.g. bistable nematic) has an advantage.

Compact lightweight systems are essential for head-mounted displays in particular, and a general requirement for portable equipment. Single microdisplay systems have an overwhelming advantage over multiple microdisplay options. LEDs in primary colors provide simple compact field-sequential operation, moreover their rapid switching compensates the marginal response time of the LC microdisplay. Price and field-of-view handicap the DMD in NTE display applications.

OLEDs are an attractive option for NTE displays. Lambertian emission fills the aperture of the magnifying optics, and power requirements are low. RGB pixels are the natural imaging method for OLEDs, which is a drawback at high magnification. However, color sequential filtering is possible with white light emitters. Several features of microdisplays favor early application of OLEDs. They are current rather than voltage controlled, and current control is easier to implement in c-Si. OLEDs are sensitive to contamination and stringent protection favors the compact structure of a microdisplay. The emergence of high-quality, lightweight, head-mounted displays should have a strong influence on expanding market opportunities in that area.

1.6.3 Other Applications

In principle, the DMD can operate over a wide range of wavelength determined by aluminum reflectivity at short wave, and diffraction at long wave; however, it is limited to near-UV at present. The DMD finds employment as a programmable mask in photolithography and related UV imaging applications.[18]

Liquid crystal microdisplays are handicapped in the UV range by photo-induced degradation, but material selection provides reasonable life in the near-UV range at modest intensity. Infrared operation is limited by absorption bands, and the birefringence required to achieve half-wave retardation. Military applications include infrared projection, creating an infrared scene to test detector arrays.[19]

Light modulating microdisplays favor incoherent light in display applications. The same devices can act as SLMs for coherent optical systems. Phase coherence exploited in such systems requires low phase distortion in the optical components, which will need some selection among available microdisplays. SLMs find application in programmable gratings and holograms found in optical systems; diffraction efficiency is much higher for phase-modulated structures, conferring advantage to liquid crystal devices. The following examples suggest the scope of optical signal processing, but further detailed accounts are beyond the scope of the book.

In holographic storage, a laser beam splits into signal and reference beams, and an SLM imposes a pattern of amplitude modulation on the signal beam. The modulated signal beam combines with the reference beam to record a volume hologram in a photosensitive material. Readout of the hologram by a reference beam retrieves the information in the form of the original modulated signal beam. A photodetector array transforms the two-dimension information back into the electronic domain, as required.[20]

Coherent optical correlation achieves pattern recognition by matched spatial filtering. An SLM writes the input image on a laser beam. A passive lens performs a two-dimensional Fourier transform of the input image, conferring translation invariance, while phase encodes location. A sequence of Fourier plane filters written to a second SLM interrogates the transformed image. A second lens transforms the output back to image space, where a photodetector array measures the output of each filter, allowing electronic assessment of the best match, along with location identification in the input plane. The system exploits the Fourier transform ability of a lens, and the product of Fourier transforms gives the correlation function.[8]

In telecommunications, optical signals routed over fiberoptic networks require optical switching and compensation devices. Microdisplays find application in array switching, and amplitude and phase control. Two-dimensional modulator arrays form programmable holograms providing wavelength discrimination and beam deflection.[21,22]

1.7 References

1 Ishino, H. and Inoue, T. (2004) QUALIA-004 full high-definition home theater projector using silicon crystal reflective display (SXRD) technology. *IDW04*, 1687–8.

2 Barten, P. J. (1999) *Contrast Sensitivity of the Human Eye and its Effects on Image Quality*. SPIE, Bellingham.

3 Curry, D. G., Martinsen, G. and Hopper, D. G. (2003) Capability of the human visual system. *Proc. SPIE*, **5080**, 58–69.

4 Brennesholtz, M. S. (1998) Luminance contouring at the display in digital television. *Proc. SPIE*, **3296**, 62–70.

5 MacDonald, L. W. and Low, A. C. (1997) *Display Systems*. John Wiley & Sons, Chichester.

6 Toyooka, T., Yoshida, S. and Lisaka, H. (2004) Illumination control system for adaptive dynamic range control. *SID Digest* **35**, 174–7.

7 Tuenge, R. *et al*. (1997) A field-sequential color VGA AMEL display. *SID Digest* **28**, 862–5.

8 Casasent, D. (1979) Coherent light valves. In R. Kingslake (ed.) *Applied Optics and Optical Engineering*, Vol. 6, 143–201. Academic Press, New York.

9 Farrayre, A. *et al*. (1985) Geometric resolution improvement of Sodern visualization system. *SID Tech. Digest* **16**, 266–9.

10 Grinberg, J. *et al*. (1975) A new real-time non-coherent light image converter the hybrid field effect liquid crystal light valve. *Opt. Eng.* **14**, 217–25.

11 Sterling, R. D. and Bleha, W. P. (1999) Electronic cinema using ILA® projector technology, *SID Digest* **30**, 216–19.

12 Ernstoff, M. N. *et al*. (1973) Liquid crystal pictorial display. *Tech. Digest IEEE Intl Electron Devices Mtg*, 548.

13 Lipton, L. T. *et al*. (1977) A fully integrated MOS liquid crystal video-rate display panel with drivers, *SID Digest* **8**, 64–5.

14 Welkowsky, M. S. *et al*. (1987) Status of the CCD addressed liquid crystal light valve, *Proc. SPIE*, **825**, 41–8.

15 Morozumi, S. (1984) A 25-in and 1.5-in B/W and full-color LC video displays addressed by poly-Si TFTs. *SID Digest* **15**, 316–17.

16 Handschy, M. A. *et al*. (1990) Fast ferroelectric-liquid-crystal spatial light modulator with silicon-integrated-circuit active back plane. *Proc. SPIE*, **1291**, 158–64.

17 Hornbeck, L. J. (1989) Deformable mirror spatial light modulators. *SPIE Critical Review Series*, **1150**, 86–102.

18 Dudley, D., Duncan, W. and Slaughter, J. (2003) Emerging digital micromirror device (DMD) applications, *Proc. SPIE*, **4985**, 14–25.

19 Wu, S-T. *et al*. (1985) Infrared liquid crystal light valve. *Proc. SPIE* **572**, 94–101.

20 Nesbitt, R. S. *et al*. (1999) Holographic recording using a digital micromirror device. *Proc. SPIE* **3637**, 12–20.

21 Duncan, W. *et al*. (2003) DLP™ switched blaze grating; the heart of optical signal processing. *Proc. SPIE*, **4983**, 242–9.

22 Crossland, W. *et al*. (2004) Liquid crystals in telecommunications systems. *Mol. Cryst. Liq. Cryst.* **413**, 363–83.