J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*

IMAGE PROCESSING EQUIPMENT

Image processing equipment (*IPE*) encompasses a range of machines that collect or generate images, altering them in some way to improve the perceptibility of the images by human or machine. IPE can enable us to see in dark or remote locations or simply to improve the quality of a picture or synthetic scene. The equipment has historically been an effective tool for business, industry, medicine, entertainment, military, and science.

Applications of Image Processing Equipment

Image processing equipment now extends from the smallest of electronics items produced today, such as a microminiature camera based on a highly integrated chip in biomedical and hidden surveillance applications, to some of the most complex and advanced machines produced. IPE is ubiquitous, being used by virtually every industrialized nation today. While everyone may not be familiar with the direct use of the technology or the plethora of information we indirectly create with image processing equipment, most greatly benefit from its use. Many of the images we see on television and movies were created with image processing. The weather reports we watch include enhanced satellite imagery. The checks we send through the banking system are converted to digital images and then destroyed. The US Treasury uses image processing to assess automatically the quality of currency just after it's printed. The medical professions feature many imaging modalities, including X-ray and ultrasound images for diagnosis. All of these applications utilize various IPE.

Many applications of image processing have become familiar to us due to its use by the entertainment–film industry, personal computers with imaging capabilities, and the Internet. The near perfection of computer-generated imagery (*CGI*) of dinosaurs in Paramount Picture's *Jurassic Park* and scenes of the demise of the *Titanic* offer shocking realism to movie viewers. Major motion pictures, such as Paramount Picture's *Forrest Gump* (released in 1995), feature image manipulation for entertainment. By this technology actors re-live historical events, flawlessly "mixing" with actual film footage of the event. Using image processing to control lighting, color and blending of images, the actor Tom Hanks was realistically portrayed to converse with Presidents Kennedy and Johnson at White House functions of their time, long after their deaths. The actor also became a world-class table tennis player by careful frame-by-frame manipulation of a ping pong ball. In addition it has become all too common to see image processing equipment extracting recognizable images of criminals with blurry and dark photographs taken at crime scenes. Publications have seen development costs reduce with the use of computer imagery. Electronic mailing of digital images captured with the use of scanners or electronic cameras attached to personal computers (*PC*s) connected to the Internet continues to increase in popularity. Finally this technology may eventually develop to supply prosthetic artificial eyes for the blind (1).

Many Users of Image Processing Equipment

Nearly every industry uses some form of equipment that processes images. Consumers use cameras for photography and home videos. Travelers use security cameras for vacation homes and Internet "spy" cameras for seeing what awaits them. Doctors and technicians use IPE for medical instrumentation and computer-aided analysis of the images for diagnosis and treatments. West Virginia Department of Motor Vehicles staff use IPE with facial recognition software (from VisionSphere Technologies Inc.) to verify the identity of those attempting to renew a driver's license. A neurobiologist in Massachusetts uses thumbnail-size cameras on the back of horseshoe crabs ("Crabcams") to study behavior of aquarium-bound crabs when viewing the captured

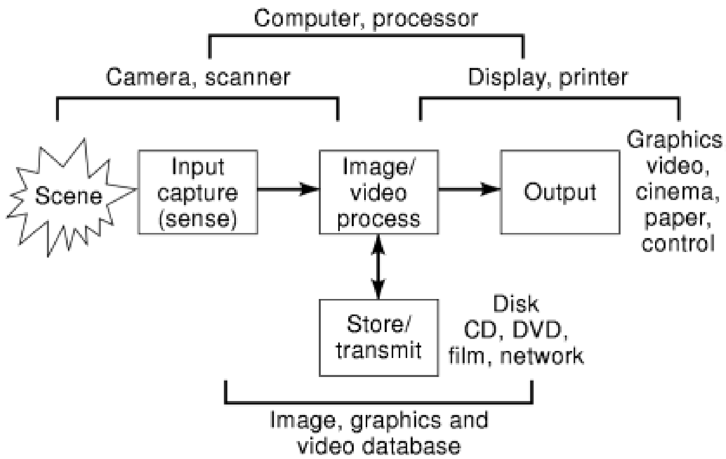


Fig. 1. Elements of image processing equipment (*IPE*) include the input block (for image/video capture/sensing), the processing block (for image/video computing), the storage/transmission block (for image/video database and compression), and the output block (for display or print of the images).

scenes. Astronomers use multispectral, high-resolution and high-sensitivity cameras to see into deep space with ground-based and space-based telescopes (Hubble Space Telescope) coupled with CCD imagers. *NASA* used imagers to navigate the *Pathfinder Rover* on Mars.

Manufacturers use vision systems for factory automation, such as inspection of units produced and ultimately product quality control. Pilots use remote sensing and vision systems for navigation. Soldiers use remote battlefield reconnaissance sensors to see the enemy prior to entering the field. Battleships use airborne, visible, IR and radar imagery for detection of possible hostile concerns. Semiconductor manufacturing technicians use X-ray imaging to align ball-grid-array (*BGA*) package components onto printed circuit boards. Digital image manipulation for explicit alteration of events has been used for "spin doctoring" of news or evidence [see (2)].

Elements of IPE

Figure 1 illustrates common elements found in most IPE. The *input* block includes the functions of capturing a scene with a camera or scanner. The time frame for the event of scene capture varies with the application. Interactive communication with images and video would dictate almost instantaneous capture and processing of the image sequence, while many applications utilize image storage for later analysis. The *image process* block includes a wide body of functions and algorithms for image processing, coding, color/texture analysis, enhancement, reconstruction and rendering, using techniques like transform domain processing, histogram analysis, and filtering. The *store/transmit* block provides image archive, distribution, communication and database mechanisms, using solutions like video disks, CDs, DVDs, films, or computer/telephony networks. Finally the *output* block includes displaying or printing the images on graphics or video monitors, cinema, paper, or for direct machine control.

Removal of Technology Limitations

As with other areas, the availability of related technologies has restricted the capabilities of IPE. Parameters such as the size of the display, the sensitivity of the camera, the number of processing cycles, the amount of memory for image storage and the number of bytes per second available for communication have historically limited its widespread use. Regardless, science, defense, manufacturing, and commerce industries have embraced the use of IPE. The equipment is widely used for automation of manufacturing processes, entertainment, detection of items not easily perceivable with human vision and minimization of travel. The technologies just recently developed will very shortly enable IPE to captivate users with immersive high-resolution displays, supersensitive cameras, powerful supercomputing technologies, extremely high-density memory cards, and high-bandwidth networks, all for cost-effective personal and portable use. Equipment to perform functions such as lifelike video conferencing and automatic face recognition, described in futuristic movies and books, is emerging.

Image Representations: *Numeric Pixels and Symbolic Objects*

Images are represented two primary ways—numerically as a multidimensional array of pixels (or picture elements) or symbolically as a list of object descriptors (also known as graphical or vector representations). While graphics processing utilizes geometric vector representations for polygon drawing and image processing utilizes pixel arrays representations, some applications use both representations.

The most common representation, the array of pixels, includes a multidimensional array of scalar quantities, proportional in some way to the scene. These include color photographs with shades and tints of visible color and thermal satellite images of earth with pseudo color shadings showing thermal signatures. This representation provides convenient application of signal processing algorithms and manipulation by specialized processor chips. Pixel arrays have been extended to higher dimensions with volume pixel elements or voxels. In this case a voxel would contain a scalar value that represents the density of a point in space (e.g., a human body) at the three coordinates associated with each value. An *IPE* could then create a 2-D image of any arbitrary plane slicing through the 3-D object by accessing the appropriate voxels.

Using the graphical or vector representation, the object descriptors contain information or instructions specific to reconstruction of the objects on a display medium. This method was developed for graphics applications and vector-oriented displays or printers (e.g., x–y pen plotters). In this case simple objects or drawings like line art and schematics can be drawn with a few strokes rather than painting an entire x–y grid image. Using an ordered list of these descriptors (which could include a 2-D image), the image-rendering system first processes and displays objects that have priority, such objects near the center of attention of the user or not even displaying those objects that have been occluded by others.

An advantage of the “graphical” representation is simpler creation of 3-D scenes as a function of the viewer’s pose or gaze at objects relative to the other objects. Either representation can be used to create 3-D scenes or views; however, depending on the scenes complexity, the graphical method can yield better compaction of data (for storage or transmission) and simpler implementations.

Organization of the Article

This article is organized by categories of IPE. The streaming of data and information through image processing equipment can follow the path represented by these categories. They are:

(1) Image sensing equipment, for acquiring the original images
(2) Image computing equipment, which can include enhancement, analysis, measurement, and compression
(3) Video and image compression equipment, for real-time communication with others
(4) Image databases equipment, for storage and on-line access of information and transmission
(5) Image display equipment, for on-line visualization by the observer
(6) Image printing equipment, for hardcopy recording or visualization by the observer

In addition to an overview of the system products or subsystems associated with each, this article includes a summary of the components (or chips) that form building blocks with the underlying technologies, enabling image processing applications today and into the future. Pioneering work in each area has created technologies that are used in nearly every industry today.

Image Sensing Equipment

Image sensing equipment (*ISE*) provide raw images for use by image processing equipment. This equipment includes electronic still cameras, video cameras, medical imaging equipment, and page scanners. *ISE* utilize imaging detectors that produce signals proportional to the intensity of the source illumination. Today CCD chips are the leading type of imaging detector, featured in virtually every class of ISE. Some cameras can collect signals well beyond the perceptibility limits of biological human vision. Examples include using X-ray or infrared spectrum sensors. Many ISE internally utilize image processing to condition the signals in preparation for distribution or to simply remove any sensor-induced artifacts such as geometric distortion.

Cameras can sense remote locations, enabling inspection of nearly any situation such as microscopic surgery, lumber processing, or viewing the surface of Mars. In many cases cameras provide the sensing function remotely from the final IPE such as in different cities, in space or under the ocean. *Remote* connection of sensing equipment, as described later in this document, requires digital compression of the images using a variety of methods such as *ISO-MPEG* (Moving Picture Experts Group) or *JPEG* (Joint Photographic Experts Group) standards for storage or transmission via radio waves, telephone, cable, or satellite transmission to the IPE. *Direct* connection to the IPE may utilize compression technology to either improve the quality or reduce the cost of the connection, or a simple full-bandwidth electrical connection.

Electronic Cameras and Detecting Methods. Electronic cameras can be as large as a shoebox and as small as a single *IC* package including an integrated lens. They capture images of the environment by sensing visible, X-ray, infrared, or other spectrums of light. Cameras can capture still images, a sequence of images used to represent "live" motion (video), or both, depending on the camera's storage and interface methods. The sensing technology, which includes the camera's optics, photo-recombination methods, sensitivity, size, power, mobility/portability, and electrical interface bandwidth, all directly limits a camera's capabilities.

Various methods have been used to detect light, including photo-emissive devices like photo-multipliers, photodiodes, phototransistors, vidicons, and silicon chips, all of which generate electrons over time proportional to photons hitting the imager. Professional television cameras were originally developed with vidicon detector tubes. These detectors featured a sensing area that collected the generated charge in a material that acted as an array of capacitors and contained circuits for electrical readout of the charged array. In effect an electron beam scanned across the sensor, regenerating the charge into those capacitors, and was bleed away by the illumination as a circuit sequentially detected the amount of charge used to replenish the sensor. For cost and performance reasons, electronic cameras have practically made a complete transition from tube technology to semiconductor sensors, including *CCD*s (charge-coupled devices) and *CMOS* (complementary-metal-oxide-silicon).

Camera Sensor Chips. The semiconductor architecture most used today, the CCD as shown in Fig. 2, produces a spatially sampled array of pixels with charge proportional to the incident illumination and uniquely transfers the charge from each storage site (or well) to its neighboring storage site in succession [see (3)]. This simultaneous transfer of charge from well to well is characterized as *bucket briggading.* A CCD generates charge, transfers that charge and converts the charge to voltage on output. Charge generation can be in an array of either MOS capacitors or photodiodes (for interline transfer). In either case, charge transfer occurs in MOS capacitors. The charge is finally converted to a voltage signal at the output of the CCD.

A CCD offers uniform response, since its charge generation and collection happens in "precisely" patterned sites or charge wells. During manufacture of CCDs, semiconductor photolithography and process steps like ion implantation create these precise pixels. With vidicon tubes, the specific location of the pixel on the detector varied as a function of input scan voltages. This created a challenge for camera designers that usually resulted in some geometric distortion of the detected images. With CCDs, the pixel location does not change, yielding a very geometrically stable pixel, with less sensitivity to scanning or clock voltages. The CCD's use of a *single* high-quality transconductance amplifier, coupled with these highly uniform pixel characteristics, leads to high performance, with excellent uniformity at the output of the device. As with other semiconductor devices, defects

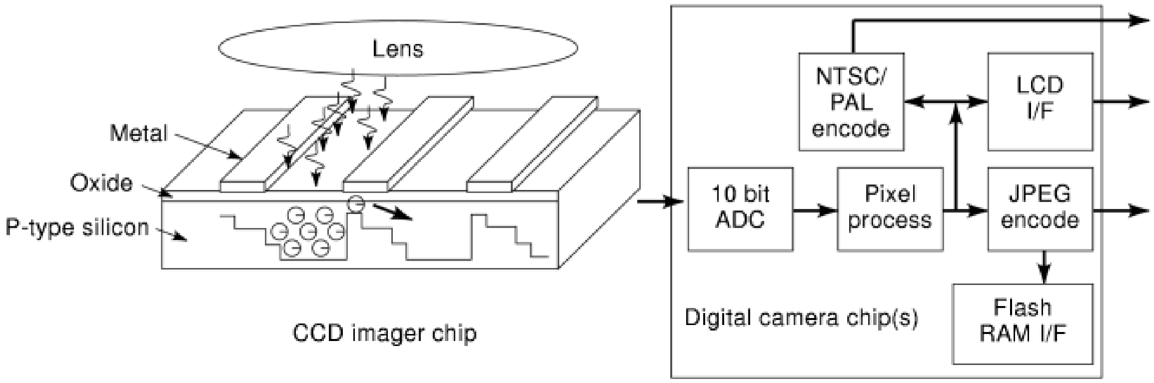


Fig. 2. The virtual phase (VP) CCD imager chip (shown on the left) and the digital camera chip or chips (shown on the right) have been used to make digital cameras. The CCD converts incident photons to electrons, collecting them in patterned potential wells or pixels. Metal lines control the transfer of the charge in each pixel in unison. These vertical metal lines and horizontal channel stops (not shown) define the pixel location. In a VP CCD another "virtual" transfer control line (implanted beside the metal line) works in concert with the metal line to orchestrate charge transfer horizontally as shown by the potential plot in the illustration. The digital camera chip(s) can be fully integrated solutions or independent chips, depending on the product requirements.

can occur. Several techniques have been developed to mask these defects, like pixel output replication [see (4)], to increase the processing yield and resultantly lower the cost to manufacture these devices.

CCDs support tremendous variations of integration time, useful for stop action or low illumination applications. Low dark current (the amount of charge generated with no light onto the detector) and high sensitivity mean that a CCD can detect scenes with extremely low illumination when using long exposure times. Cooling CCDs with thermal electric coolers (TEC) further reduce dark current. For these reasons CCD-based telescopes extend our sight to previously unobservable galaxies (e.g., with the Hubble Space Telescope). For security applications, this increased sensitivity means that intruders are seen in extremely low illumination situations. To reduce the cost of consumer products, video camera manufacturers can use inexpensive, larger *f*/number lens with light absorbing color filters and still be able to make excellent pictures in dimly lit rooms. On the other hand, a very short integration time can be used for capturing rapid motion with stop action clarity such as freezing the image of a hummingbird's wings in motion. For these "stop-action" applications, a CCD with "electronic-shuttering" again has the advantage over conventional photography's mechanical-shuttering techniques.

A CCD will use *frame-transfer* or *interline-transfer* methods (5) to quickly transfer the exposed pixels to covered sites, thereby abruptly terminating integration for those pixels. Only then does the CCD read out the frame of pixels begin, usually at conventional video display rates. Simultaneously integration of the next frame begins in the detecting array. Frame transfer techniques can offer the advantage of a single video frame acquisition, with less field mixture, for freeze-frame applications. Interline-transfer architectures can have lower efficiency than frame-transfer architectures, since the relative percentage of photon-collecting area is lower. However, interline-transfer approaches facilitate the use of photodiode detectors and anti-blooming structures that prevent charge spillage to adjacent pixels when overexposing the device.

Since CCDs contain analog pixels, with charge proportional to the light integrated on each charge well, devices have been developed to perform analog signal processing of the image directly on the imaging chip. For example, a CCD can add two pixels together by simply clocking two pixels successively into one charge well. This can reduce the system cost, size and power for several applications.

The size of the pixel affects the dynamic range of the device. A larger pixel can generate more charge. The structure of the clock lines, which facilitate charge transfer, but also reflect light away from the underlying recombination silicon, influences overall efficiency of the CCD. The physical clock structure and shape of these timing signals used to transfer charge determine the devices ability to completely move all the charge from well to well. Any residual charge will lead to poor dynamic range or picture smearing. The architecture of the sensor, including the size of each pixel, the structure of the light-blocking clock lines, and the characteristics of the photo-recombination wells, influences a CCD's performance.

CMOS Sensor Chips. *CMOS* (complementary-metal-oxide-silicon) sensors represent another architecture that directly images light onto a structure similar to conventional CMOS memories. These sensors require more complicated addressing schemes to select specific x–y cells than a CCD. In addition more charge-sensing amplifiers are used, creating more performance issues. CMOS sensors also have lower sensitivity, since their complex clocking structure again tends to block more of the underlying silicon where photo-recombination occurs, leading to pixel fill factors as low as 30%. Lower dynamic range results with higher output amplifier noise relative to a CCD. However, the CMOS approach has seen a resurgence of interest lately due to the cost-effective nature of integrating many full-digital video functions (like analog to digital conversion and compression) and sensing functions directly onto one silicon chip, as well as the use of conventional high-volume semiconductor memory fabrication techniques.

Color Imaging Chips. Professional CCD cameras incorporate three sensors and special optics for simultaneous sensing of three images corresponding to the color primaries (red, green, and blue or yellow, magenta, and cyan). Alignment of the individual sensors and the optical assembly creates challenges for manufacturers. For cost reasons, consumer-level cameras utilize only a single sensor coupled with a color filter mosaic integrated onto the top of the chip. Alignment of the filters can be performed optically before cementing onto the chip or directly patterned onto the device. Manufacturers utilize an array of color-filtered pixels which can be combined to form a color triad of pixels. Unfortunately, this color-sensing method can reduce the overall spatial resolution by as much as a third in one dimension. Taking advantage of the human eye's increased sensitivity to green, some color sequences like green-red-green-blue improve the apparent luminance resolution to only a half the original grayscale resolution (rather than a third).

Geometric Distortion and Correction with Warping. Many cameras contain optical systems that present a distorted image of the scene. The distortion can result from geometric properties of the optical system, such as the use of a fish eye lens for wide-angle imaging, or cylindrical lens for line scanning or low-quality lens. Sensor artifacts, such as with the use of vidicons, can also worsen these problems. The distortion can also directly result from applications like producing photomosaic maps of the earth, whereby the captured 2-D satellite images contain curvilinear sides of varying severity, as a function of the angle looking away from normal. Distorted and disjoint views from spacecraft such as the *Viking* and *Mariner* also require correction and mosaicing for proper viewing.

A warping process can correct for these distortions as described in Ref. 6. With warping, the algorithm matches control points located on the distorted image and the target image. For an application such as use of a fish eye lens, this process is very simple, whereas other applications need manual intervention to determine the proper mapping coordinates. Each pixel is then mapped from the source into the destination. Since this cannot be a one-to-one mapping, interpolation of the source image is required per pixel. The degree of the polynomial that estimates the missing pixels will influence the result. However, simple bilinear interpolation or first-order warping can generate effective results. Using four input pixels I_1 (pixel value at $x+1, y$), I_2 (pixel value at $x+1, y+1$), I_3 (pixel value at $x, y+1$), and I_4 (pixel value at x, y), we can compute the interpolated pixel value or intensity $I(x, y)$ for an arbitrary pixel at a location within the four pixels as

$$\begin{aligned} I(x,y) = {} & I_4 + [\Delta X^*(I_1 - I_4)] + [\Delta Y^*\{I_3 + [\Delta X^*(I_2 - I_3)] \\ & - I_4 - [\Delta X^*(I_1 - I_4)]\}] \end{aligned} \tag{1}$$

where ΔX and ΔY are fractional distances (somewhere between 0 and 1), computed as the fractional coordinate of the input pixel being mapped from the output pixel, as follows:

$$\Delta X = \mathrm{frac}[A1^{*}X + A2^{*}Y + A3] \tag{2}$$

$$\Delta Y = \mathrm{frac}[A4^{*}X + A5^{*}Y + A6] \tag{3}$$

This affine warping can produce translation, rotation, scale, and shear mapping. A set of control points can be used to determine the coefficients of $A1$ through $A6$. Once calculated, we can use incremental addressing (one addition operation) to reduce the number of computations.

For a digital signal processor (*DSP*) or media processor this warping function requires eight adds/subtracts, four multiplies, four loads, one store and several register accesses per pixel, or perhaps 15 cycles per pixel. With a 640 × 480 pixel image processed at 60 frames per second, this requires nearly 300 million operations per second (*MOPS*).

Digital Still Cameras. Many manufacturers produce digital still cameras (Olympus, Sharp, Epson) for both consumer and professional markets. Figure 2 shows the key elements of an electronic digital camera. These products enable Internet-based applications of imaging by capturing digital images for subsequent processing, printing and storage with a personal computer. Several manufacturers produce digital still cameras with 640 × 480 pixel resolution [including Nikon (http://www.nikon.com), Panasonic (http://www.panasonic.com) and Canon (http://www.ccsi.canon.com)]. With integrated *LCD* display, computer interfaces, JPEG storage of nearly 130 normal resolution or 65 fine resolution images in 4 Mbyte of memory and annotation on the images using a pen stylus or speech input, these cameras offer a portable solution for image acquisition. Another unit offers a completely enclosed camera on a small PC card for mobile use on a laptop. At the high-end, the EOS D2000 (from Canon and Kodak) offers a 2 million pixel (1728 × 1152) electronic camera for photojournalism or surveillance applications. With the cost lowering of semiconductor Flash memory cards and continuing CCD improvements, consumer digital still cameras will proliferate in the near term. Likewise development of custom chips, and recently a microprocessor-based solution for the digital camera functions shown in Fig. 2, enables cost-effective and flexible solutions. The full-integrated Mips-based DCAM-101 (from LSI Logic at www.lsi.com) offers a single-chip solution with pixel processing for color space conversion with a Bayer pattern of CCD color filter mosaics, coupled with interfaces for all other functions in the diagram.

One manufacturer (Hewlett-Packard Co.) produces a full solution home-photography kit called HP Photo-Smart. The kit includes a camera with 640 × 480 CCD with de-mosaicing, filtering, and JPEG compression, a 2400 dpi (dots per inch) scanner with feeders for color slides, negatives or prints up to 5 × 7 inches, and a printer. When used with a PC the system provides a complete picture taking, editing (using Microsoft's PictureIt software), printing and photograph storing solution.

Another manufacturer (EarthCam) produces an integrated camera solution, called the EarthCam Internet Camera, with direct connection to the Internet via a built-in 28.8 kbyte/s modem. Users simple connect the camera to a phone connection for remote control and access to pictures captured from the remotely located camera.

Video Cameras. Video cameras are frequently used a sources for IPE video capture. External frame grabber equipment can capture still images from these cameras or, with some newer video cameras, internal electronics captures still images for subsequent transfer to a computer via standard interfaces. Camcorders include tape storage of up to several hours for later playback. Camcorders have become ubiquitous today. These cameras have one limitation—image stability caused by one's natural shaking of the small handheld cameras. To correct for this jitter of the picture, some units utilize electronic stabilization. Minute repositioning of the optical system when circuits internal to the camera detect subtle motion stabilize the pictures. Alternative methods include the adaptive repositioning of the physical position of the output video on the sensor. Motion detection processing techniques drive this selective repositioning of the starting pixel of readout of the CCD sensor. Unfortunately, this technique requires a larger sensor to maintain the same picture size while repositioned.

Analog and Digital Outputs. Most video cameras have analog video outputs. Lower-end units use a single composite luma (V) and chroma (C) signal, mid-end units use separate Y and C signals, and high-end units use individual components of Y, R-Y, B-Y. These are acceptable for many applications, but they limit overall performance for critical applications. Recent digital video cameras (from Sony) have included IEEE-1394 (originally developed by Apple Computer) as a standard solution to digital interface of the camera's video to storage or other computer equipment. This approach minimizes cost and improves picture quality in comparison to conventional analog interfaces that may require digital-to-analog converters in the camera and analog-to-digital converters in the image computing/storage equipment.

Other Video Cameras. One video camera manufacture produces an MPEG camera with CCD sensor, LCD display and direct MPEG full-motion digital compressed video and audio stored on a 260 Mbyte disk (Hitachi at http://www.hitachi.com). Twenty minutes of MPEG-1 video is stored or 3000 JPEG images, in any combination. The MPEG camera's combination of JPEG still image and MPEG movie creation simplifies Internet web site development or CD-ROM creation.

Digital video (*DV*) camcorders (Sony Corp.) offer a higher-quality alternative to conventional analog video camcorders (like 8 mm or *VHS-C*). DV technology utilizes proprietary algorithm optimized for recording on a small format tape, while supporting trick modes such as fast-forward. DV also supports the 16:9 video format, which has wide acceptance in Japan.

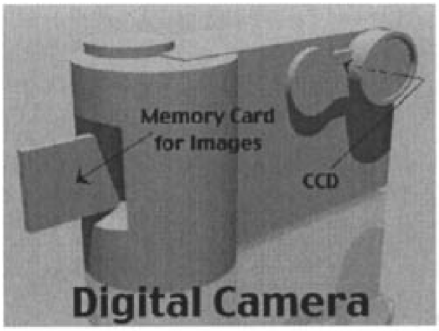
CCD-based video cameras for production studios (from Sony and Eastman Kodak Microelectronics) can produce high-resolution 1920 × 1080i video format signals. Major events such as the New Years' Rose Parade and the Olympics have featured these cameras. Video from these cameras will be used for terrestrial broadcast of digital high-definition TV (*HDTV*) in the United States.

Cameras with Pan-Tilt-Zoom. Applications such as surveillance or video conferencing require manipulation of the camera's viewing pose. To address this need, several video cameras manufacturers have integrated pan-tilt-zoom (*PTZ*) capabilities into the camera unit (e.g., the Sony D30 and Nisca). Other advanced camera systems for video have integrated microphone arrays and signal processing to create a control signal for pointing or steering the PTZ camera to the pose corresponding to the speaker's location within the conference room (PictureTel and Vtel).

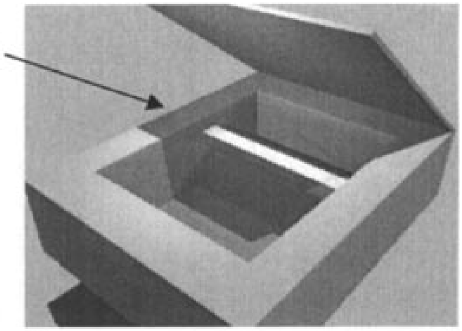
Medical Imaging. Medical IPE are used to diagnose and treat patients. Several modalities, like MRI, ultrasound, and X-ray *CT* (computed tomography), are used to detect features within patients or to measure objects with either still or moving image sequences. For example, MRI can provide images for interpretation of the presence of disease or for guidance of radiation therapy. Medical imaging applications require extreme accuracy and repeatability, with no artifacts induced by the equipment. The acquired images can be 2-D slices (cross sections) or images of the body, with pixels corresponding in some way to the density of the object. Processed images are either projected images of this acquired 2-D data or digitally reconstructed views of projections throughout the volume of data.

Magnetic resonance imaging (or *MRI*) builds a cross-sectional image of the body by first aligning the body's protons (hydrogen nuclei) spin axis with a large electromagnetic field. Radio waves then knock the protons out of alignment. As the radio waves are removed, the protons realign to their "normal" spin orientation, emitting radio waves that are finally detected by an RF scanner. An image processor then creates a cross-sectional image of the body with repeated scanning in succession to form a 3-D volume. The image processor performs filtering, false coloring by mapping intensity to color, and volume rendering tasks.

Ultrasound imaging passes harmless ultrasonic sound waves through the body and uses microphones to detect internal reflections, which are then composed into an image in real-time. Again improvements in image quality are obtained with spatial and temporal filtering. Today's research includes the use of these imaging



Digital camera



Page scanner

Fig. 3. Camera and scanning equipment with their elements, each including a CCD imager.

systems to actually manipulate or target the delivery of drugs or genes with combinations of ultrasound and MRI.

Digital X rays are produced with amorphous silicon large area sensing technology (see DpiX at http://www.dpix.com) integrated with a video camera to pick up the real-time flouroscopic image. While these can be used for medical applications, other industrial and scientific applications (like welding) are popular. The medical profession has historically utilized film radiology, coupled with scanners for digitization.

Scanners. Scanners represent a class of camera equipment which capture images by effectively sweeping a sensor (linear or 2-D) across an object, capturing sequences of data which are later reconstructed to form the actual image. The image produced by a scanner can be as simple as an image of a page, as complex as a full 3-D volumetric image of one's body, or as large as a high-resolution view of the entire surface of the earth.

Page Scanners. Xerox led the development of an entire paper scanning industry based on the flat-bed scanner and printing in one unit, the ubiquitous photocopier. Common in most businesses and many households today, page scanners are manufactured by several companies (e.g., Hewlett-Packard, Umax, and AGFA). Figure 3 shows the common elements of a flat-bed page scanner. These elements include a linear CCD imager, a scanning mechanism and a buffer that holds the scanned image. For applications like desktop publishing, manufacturers integrate paper control and advanced CCDs, creating a range of handheld to desktop scanning devices with about 600 dots-per-inch (*dpi*) up to 1200 dpi, and 30 to 36 bits of color. Several scanners also incorporate resolution enhancement functions, such as interpolation of the images, to increase the apparent resolution of the scanned image by up to a factor of 10. CCDs have typically been used for scanners due to their relatively low cost and simplicity of creating extremely high resolution in one linear dimension. CCDs are only limited by the size of a silicon wafer (8 or more inches at 1000 pixels per inch) and packaging technologies.

Airborne Scanners and Machine Vision. Large-scale, high-resolution terrain sensing is achieved by taking advantage of the extremely high-resolution detection possible with scan-line sensing. In airborne scanning, with a known ground speed, sweeping a landscape with 2-D imaging arrays generates multiple samples of the scene with only fixed time-delay between each row. These time-delay and integration (*TDI*) techniques have the benefit that repetitive samples of the same pixels are available for averaging to remove sensing noise, thereby improving the quality of the sensed scene. Similar TDI techniques are used to inspect objects moving at high speeds on a conveyor belt in machine vision applications such as lumber processing.

Infrared Cameras. Cameras that image by detecting the heat of the source or infrared have numerous applications for security, military or night surveillance situations. Photo-multiplier tubes and *CCD* arrays sense the near-infrared spectrum. Far infrared detectors have the distinct advantage of imaging through fog. Early far-infrared cameras were devised with mechanical scanning a linear array of *IR*-sensitive detectors (e.g., mercury-cadmium-telluride and lead-tin-telluride) as described in Ref. 7. For proper operation, these detectors were cooled to near 100 K with cryogenic coolers. More recent infrared cameras utilize 2-D arrays of detectors, with a resultant shape and size similar to conventional visible cameras. These thermal cameras are used in electronics manufacturing to find hot, defective circuits and for detection of heat loss in buildings.

3-D Cameras. Medical imaging, product design, and Internet 3-D applications require capture of the 3-D images. 3-D systems include those that are (1) *true 3-D,* which acquire multiple gazes for stereoscopic display, (2) *real 3-D,* which capture volumetric data for holographic display or other 3-D displays, and (3) *non-true 3-D,* which capture a 5-D plenoptic function (color as a function of position and pose), with a succession of 3-D gazes [see (8)]. Real 3-D sensing requires either passive or active scanners (like lasers or MRI) to capture the data. True 3-D and non-true 3-D data are captured with passive cameras, which in some way capture multiple views of the scene. This can be with special optics for parallax view or with successive repositioning of the camera. The "TriForm" 3-D Image Capture System uses structured light, illuminating the object with patterns of known location and dimension (like a grid), coupled with true-color CCD imager to collect the object's surface texture. A non-true 3-D camera produced by researchers at Columbia University, called the Omnicam, creates a 360° full-motion video using a curved mirror and glass dome about the size of a large grapefruit. Some manufacturers have produced similar 3-D cameras. As discussed in a later section of this article, the image is later warped to correct for geometric distortions and to allow the viewer to select a random view 360° around the camera.

Image Computing Equipment

Image computing equipment takes raw (or previously processed) images from sensors (or stored images) and performs one of a variety of image processing functions like filtering for enhancement and recognition of features. (Another section of this encyclopedia contains descriptions of these functions and algorithms.) This equipment includes a wide range of computing equipment, both hardware and software. Supercomputers, PCs, workstations, or standalone equipment, some with specialized processing chips or *DSP* (digital signal processor) chips, perform these tasks.

The properties *manipulated* with ICE include the size, aspect ratio, color, contrast, sharpness, and noise content of an image. In addition *features* or properties of the image can be determined with ICE, including the number of objects, the type of the objects (text vs. photograph of a scene), the motion of the objects, and scene changes. With ICE, images or the objects within the images can be blended or mixed with others to create new images.

The computing needs of personal computers and workstations continually increase as users demand higher quality video, graphics, and audio, using faster networks. For example, simple word processors and spreadsheets of the last decade were replaced with later versions that automatically check punctuation and grammar, on a formatted page, with voice annotation, all as one types. This was possible due to a substantial increase in performance of the underlying computing technologies within the computer. In the future, 3-D image rendering will extend from today's screen savers and 3-D games, to be integrated with word processors that feature a 3-D representation of the document moving in virtual space, all consuming more and more resources. While the computer's host processor continues to increase in computing capability by Moore's law (about 2× every 18 months) to address some of these needs, other processing chips like image processors, media processors, DSPs, and ASICs offer many advantages.

PC and Workstation Image Computing Equipment. Image computing systems commonly use a personal computer (*PC*) or workstation as a base platform. These IPE include three elements, including (1) a high-performance PC, (2) imaging hardware, and (3) image processing software. The high-performance PC includes a host processor, usually with an image computing acceleration processor such as a DSP, storage [random access memory (*RAM*) and disk], networking and high-resolution display. Imaging hardware includes an image capture board, camera, digital video tape, scanner, or other imaging hardware. Image processing software includes the code to drive I/O (input/output) equipment, such as the camera, and the code that instructs the processor to manipulate the images.

Several elements of a computer system determine its ability to be used as image computing equipment. The processor or processors must have high speed (performing billions of operations per second), with high bus bandwidth (up to hundreds of megabytes per second) and efficient I/O. The computer's operation system must support real-time scheduling, with multiple processing tasks or threads. The software tools should include application development tools, graphical user interfaces (*GUI*) and windows for concurrent presentation of code, debug information, and images. Memories needed vary as a function of the application, with several megabytes of semiconductor memory and gigabytes of disk memory a minimum today. Networks for communication of image or video should have very high bandwidth (minimum of 64 kbps), with low latency and jitter. The peripheral interfaces should support industry standards (e.g., IEEE-1394) for direct connection to cameras. The displays should support high resolution (greater than 640×480), high frame rate (60 to 85 fps) and color (24 bits).

Image/Video Capture. Paramount to *IPE* is the capturing of the image or image sequence from some source, namely camera, video tape, laser disk, or DVD. Since most of these sources utilize analog video interfaces, a video capture board is necessary to decode and convert NTSC or PAL analog signals to digital images. The video capture function includes synchronization to the input video using a sync stripper, filtering for proper extraction of chroma and luma, and possible color space conversion. Many chips are available to perform these functions (including Brooktree at www.brooktree.com and Philips at http://www.philips.com). Several commercial and consumer equipment manufacturers produce computer interface boards or boxes for video capture, including the Matrox MeteorII (http://www.matrox.com) and ATI All-in-Wonder Pro (http://www.atitech.com). Video and image capture cards are available for standard interface buses (VME, ISA, VESA, S-bus and PCI), making them available for the majority of computers. Some computers, such as versions of Apple's Macintosh, already contain this capture function in the base computer.

Analog Video Interfaces. The video industry adopted a series of international standards for interconnection of video equipment. In the United States and Japan, the NTSC and RS-170 standards identify a composite video signaling representation that encodes luminance (luma), chromanance (chroma), and timing into a single channel. The S-Video method has been adopted for higher quality separation of luma and chroma data, achieving a higher-resolution picture without the mixing of the 3.58 MHz (US) color carrier signal into the luminance space. The artifact created by mixing these signals is seen as dots crawling on colored edges of the video. PCs and workstations have adopted RGB (with sync signal on green) and RGBS methods, each with negligible artifacts.

Digital Video Interfaces. Several alternative industry standards have emerged for digital interface between cameras, video storage, video processing equipment, displays and *PCs*—Universal Serial Bus (*USB*), IEEE-1394, and VESA. IEEE-1394 supports two compressed-video types, 5:1 and MPEG. Professional level video products use standards like RS-243 or D1 serial digital video buses. In addition, several video processing chips support the CCIR-656 digital-signaling format, one that represents a CCIR-601 resolution digital image and sync information. Even with the advantages of serial digital video buses (described previously), analog video sources will exist for many years to maintain backward compatibility.

Image Processing Application Software. The processing of these captured images or sequences is accomplished with some combination of the host processor (Intel Pentium, AMD K5, DEC Alpha, SGI MIPS, etc.), attached DSPs and/or fixed logic. For host-based processing, many off-the-shelf software packages are available, supporting popular operating systems (e.g., Unix, Windows and Mac-OS). Popular image processing software includes:

(1) *Photoshop* for general-purpose still image manipulation and enhancement (from Adobe at http://www.adobe.com)

(2) *KPT Bryce 3-D* for image rendering with texture mapping (from MetaCreations at http://www.metacreations.com)
(3) *Khoros* for general image processing (from Khoral Research, Inc. at http://www.khoral.com)
(4) *Debabilizer* for image manipulation or format conversion (from Equilibrium at http://www.equil.com)
(5) *Kai's Power Goo* (*KPG*) for image morphing (from MetaCreations at http://www.metacreations.com).
(6) *KBVision* and *Aphelion* libraries and software development environment for image processing and image understanding (from Amerinix Applied Imaging at http://www.aai.com)
(7) *Gel-Pro Analyzer* for biology applications (from Media Cybernetics at http://www.mediacy.com)
(8) *Image-Pro Plus* software toolkit for image capture and analysis functions (from Media Cybernetics at http://www.mediacy.com)

These software packages offer a wide range of image conditioning and manipulation functions to the user. In addition each manufacturer of image capture boards (or frame grabbers) and processor boards provides image processing software.

Video Editing and Computing Equipment. Studio and professional applications of video editing use a plethora of video editing equipment. While computing of the video images can be performed at any speed off-line, the presentation of the processed images must appear at video rates to properly see the result. PCs and workstations (from Sun and Silicon Graphics) coupled with software packages like Media Compositor (Avid Technology) provide complex editing solutions. They utilize compression techniques, such as MPEG, to fit the large sequences on conventional PC disks and MPEG decoders to present the results. As an alternative, high-end professional digital video disk systems (e.g., Abecus or Sierra Systems digital disk recorder DDR systems), video memory systems (e.g., Sony and DVS), and digital VCRs (e.g., Sony's D1 format recorders) provide frame-by-frame manipulation of digital CCIR-601 video frames. This equipment enables compositing or processing of video frames off-line in the host computer and later playback in real-time on the DDR, VMS, or VCR. With these techniques, the algorithm developer can visualize the effects of any image or video-computing algorithm simulated on a workstation in high-level software prior to investing development resources to port the algorithms to real-time processing on DSPs or other custom hardware.

For the home or small business, several PC-based solutions have been introduced for manipulation of video clips. With these packages one can transform many hours of home videos into condensed video programs. Apple Computer Inc. has produced AV Macintosh computers with video inputs and outputs, coupled with compression technology (QuickTime), to feature digital video editing. The "Avid Cinema" software package for the Macintosh creates an effective video editing user application (Avid Technology). Also certain VCRs (Sony) include PC software when used with appropriate camcorders to index and edit home movies. Both infrared and LANC (or "Control-L") are used to control the VCRs and camcorders. PC-add-in boards (AV Media, ATI, Matrox, Videonixs, Sigma Designs, and Real Magic) provide MPEG encoders and decoders to enable this editing application.

Finally, software-only solutions for PCs have emerged (SoftPEG MPEG player from Zoran/CompCore at http://www.zoran.com and XingMPEG from Xing at http://www.xingtech.com). This approach has been limited by either lengthened editing times or reduced image quality, partly due to the limitations of MPEG-1. However, real-time soft-only versions of MPEG-1 encode and decode exist today on 300 MHz PC systems. With technology improvements including faster processor chips, this trend toward soft-only video editing and computing will continue.

Processor Chips—Image Processing, Media Processing, DSP and Coding Chips. A wide range of processors with specialized architectures and custom chips have been developed for image computing equipment [for details see (9)]. Multimedia processor (*MP*) and DSP chips, such as Texas Instruments TMS320C80 and Philips TriMedia processors, can address the computational needs of image computing. The programmable nature of the multimedia processors, coupled with architectural optimizations for the manipulation of images, and clock speeds at several hundred megahertz, creates ideal technologies for *IPE*. All applications become "soft," with the advantage of long-term adaptability to changes in requirements, algorithms, data sets, or standards. However, the degree of progammability and flexibility to these new requirements and applications varies greatly by the architecture and software development tools associated with each solution. In addition to media processors and DSPs, special enhancements to general-purpose processors, like Intel Pentium processors MultiMedia extension (*MMX*), AMD K6 processors, DEC Alpha processors and Sun SPARC processors' VIS instructions, offer speedup of imaging functions in those environments.

Table 1 illustrates several currently available image processing and coding chips. Technology limitations to overall performance are based on chip architecture, use of parallelism, and process technology. In addition software development tools also have a dramatic impact on the efficient use of the processor and the development time utilizing the processors [for details see (10)]. While a particular processor may have exceptional speed at performing a particular function, the user's ability to change the underlying algorithm depends on the capabilities of the tools. A good C compiler that automatically finds parallelism, re-orders instructions, allocates registers, and unrolls loops can greatly accelerate the development of imaging software for the processor.

Programmable DSP and Media Processors for Imaging and Video. Several DSP and MP chips have been designed to meet the computational demands of image and video processing [see (11)]. DSP processors were originally developed in 1982 to satisfy the demands of a digital finite impulse response (*FIR*) filter or convolution. Specifically, their internal architecture contains key functional units enabling very fast multi-tap filters. With parallel units performing adds, multiplies, and complex data addressing in a single cycle, they efficiently perform the following operation:

$$H(i) = \sum_{n=0}^{M-1} b_i^* G(i-n) \qquad (4)$$

While these original devices excelled at audio signal processing applications, the cycle bandwidths required for image computing were not seen until early in the 1990s. Further extensions evolved from that basic DSP architecture to perform multidimensional processing of images (more registers), video processing (higher I/O and processing bandwidths) and graphics processing (with bit-level operations)—or media processing. With programmable multidimensional filters possible, many new applications could use these chips. From Table 1, the TMS320C80, TriMedia, and MPACT media processors possess exceptional computational power, yielding between 2 and 20 billion operations per second (*BOPS*) for image and video functions. Simple functions like image math on standard size images can require a few operations per pixel on these processor, corresponding to a few million instructions per second at video rates (computed as image size times frame rate times operations per pixel). Complex filters, geometric warps, and complete codecs can require many BOPS on these processors. The same functions on general-purpose processors could take 10 to 50 times the number of cycles of a DSP or MP.

As shown in Fig. 4, these processors use architectures with large register files, instruction, and data caches and arithmetic, logic, and multiply units to achieve speedup for image processing. These functional units are partitionable during execution to support simultaneous processing of multiple bit, byte, or word sizes, depending on the algorithmic needs. The capabilities and interconnection of these units have been optimized for classes of imaging computing including filters, transforms, morphology, image algebra, and motion computation.

Three overarching advantages exist for DSP and MP approaches to these problems. First, their programmable nature maps well to the evolution of new and unique algorithms, appearing as software solutions. For example, new approaches to video coding, such as video coding based on image warping and nonrectangular DCTs [see (12)] can be implemented in a programmable chip. They are also used to implement evolving standards like MPEG-4 [see (13)], which prevents costly hardware redesign each time the standard proposal changes. Second, these DSP and MP can accelerate time-to-market over hardware approaches whereby custom logic chips like full custom or PALs or FPGAs (from Xilinx or Altera) are designed or programmed. This assumes that the task of programming is not daunting. Third, programmable solutions enable multithreaded operations

Table 1. Image Processing and Coding Chips

Coding Chip Class	Source	Architecture	Parallelism	Technology
Programmable media processor (H.263, MPEG1 codecs)	Texas Instruments (www.ti.com) MVP 320C80	MIMD 4 VLIW DSPs and 1 RISC	18 pixels, 20 execution units; 46 ops, 9 multiply issued per cycle (21 BOPS)	0.5 e, 60 MHz
Programmable media processor (H.263 codec, MPEG1 decode)	Philips (www.philips.com) TriMedia TM-1	VLIW 1 DSP	27 execution units; 18 ops, 8 multiply issued per cycle (2.5 BOPS)	NA, 100 MHz
Programmable media processor (H.263 codec, MPEG 1 and 2 decode)	Chromatic Research (www.chromatic.com) MPACT2 LG Semicon & Toshiba	VLIW 1 SIMD	5 ALU groups; 40 ops, 8 multiply issued per cycle (2 BOPS)	0.35 e, 125 MHz
Programmable DSP	Texas Instruments, 320C60	VLIW	10 ops, 2 multiply issued per cycle	0.35 e, 200 MHz
Micro-codable videoRISC processor (H.261, 3; MPEG1,2)	C-Cube CLM4740 (www.c-cube.com) (& DIVX DVexpert 6210 up to 50Mbps)	Pipeline Parallel	NA	0.5 e, 66 MHz
Micro-codable video processor MPEG1 encoding	C-Cube CLM4111	SIMD Video DSP with separate ME/VLC	4 pixels, 3 function units	0.5 e
Micro-codable (with assembler) video codec	8 3 8 (www.8x8.com) VCP	Data-path	2 Mult/ALU paths	0.5 e
Fixed-function MPEG-2 decoder	IBM (www.ibm.com) MPEGCD20	Data-path	NA	0.5 e
Fixed-function MPEG-2 encoder (multiple for HDTV)	IBM MPEGS422	Data-path	NA	0.27 mm Leff
Fixed-function MPEG-2 decoder	SGS-Thomson (www.st.com) ST135XX	Data-path	NA	0.5 e
Fixed-function DVD decoder (MPEG-2 and Dolby)	SGS-Thomson STii5500	Data-path A/V decoders with microprocessor	NA	NA
Fixed-function DVD decoder (MPEG-2 and Dolby)	Zoran (www.zoran.com) ZR36700	Data-path driven, video decoder with DSP for audio	NA	NA
Programmable general-purpose	Sun (www.sun.com) UltraSPARC I	RISC	18 ops, 16 multiply issued per cycle	0.5 e, 200 MHz
Programmable general-purpose	Intel (www.intel.com) Pentium II	CISC superscalar, translates to RISC-like instructions	10 ops, 4 multiply issued per cycle	0.35 e, 200 MHz
Programmable general-purpose	MIPS (www.sgi.com) R10000	Superscalar	4 ops, 2 multiply issued per cycle	0.35 e, 200 MHz
Programmable general-purpose	Motorola (www.mot.com) PowerPCE 604	Superscalar	4 ops, 1 multiply issued per cycle	0.5 e, 200 MHz
Programmable general-purpose	DEC (www.digital.com) Alpha 21164	Superscalar	4 ops, 2 multiply issued per cycle	0.35 e, 600 MHz
Fixed-function JPEG codec	Zoran ZR36060	Data-path driven	NA	NA
Fixed-function MPEG2 decode	Sharp (www.sharp.com) DDMP	Data-path driven	8 core processors, 1 ARM	0.25 e, 5 BOPS
Fixed-function JBIG, G3/4 decode with scale, rotate for MFS	Pixel Magic (www.pixelmagic.com) PM-2m	Data-path driven	NA	0.6 e, 33.3 MHz
Programmable core and ASIC for JBIG decode with scale and filter for MFS apps	Xionics (www.xionics.com) XipChipE	Superscalar core and Data-path ASIC	PowerPCE core, JBIG decoder, graphics processor, & DMA w/ RAMBUS I/F	36 e Leff, 80/40 MHz ASIC/core

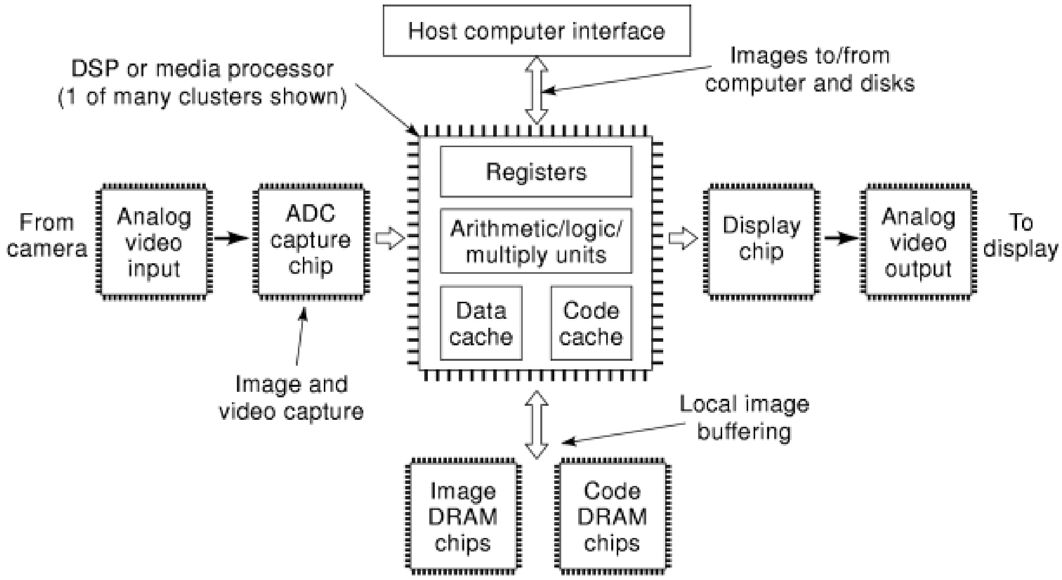


Fig. 4. This hardware diagram illustrates the elements of an image computing system. A central DSP or media processor (MP) can work in concert with a host computer to process images or video. Input, capture, display, output, and memory chips are also typically used. The DSP or MP features many internal registers, ALUs, and caches to achieve computational performance.

or the repeated switching between several independent tasks. With this feature, application developers can dynamically balance loading of the processor as a function of demand, supporting computational variation for interactive applications, enabling graceful degradation of less important functions. For example, if video quality becomes vital during a conference, the audio quality or frame rate could be lowered to allocate more cycles to video processing for that portion of time. In conclusion, the impact of using these media processors includes realization of a flexible platform for rapid development of image processing algorithms, a wider range of *practical* algorithms with several BOPS of performance and proprietary control of new algorithms in software (not dedicated to hardware).

Other Image Processing and Coding Chips. In applications where standard functions exist, such as with MPEG-2 video compression, custom or hardwired solutions find use in cases where that may be the only function to be performed. Hardwired solutions can offer efficient use of transistors and low cost with *proper* design (a time-consuming task). Many consumer products use these hardwired solutions for that reason. However, use of these hardwired devices can be so restrictive that system performance goals cannot be met. For example, the application may require higher pixel precision than that provided by the manufacturer of the hardwired chip. One can typically alter a programmable solution with software to satisfy the requirement. Other solutions, such as the CLC4000 series and VCP (C-Cube and 8×8, Inc.), include fixed datapath processors, combining either dedicated logic function blocks or micro-instruction function blocks which are hand coded with assembly. This data-path architecture limits overall efficiency for only but a select class of algorithms, since the silicon devoted to those functions must always be kept busy with those functions to justify their cost. However, for specific applications, those solutions can have a cost advantage over programmable solutions. (See web site http://www.visiblelight.com/mpeg/products/p_tools.htm for a more complete list of these MPEG processors.)

Typically these hardwired chips only perform the coding and/or decoding portion, depending on other back-end or front-end chips to pre- and post-process the video (e.g., for conditioning, color space conversion, and scaling), and align with digital audio data. Companies like S3, Brooktree, TI, Pixel Semiconductor, and Cirrus

Logic provide chips for functions like dithering and color space conversion in scalable windows. Unfortunately, the scaling and filtering differs as the application switches between standards (MPEG-1 to MPEG-2) and resolutions (source or display). With advanced media processors these functions can fit into the processor with the codec, simplifying design and minimizing costs.

MVP: An Architecture for Advanced Image and Video Applications. The multimedia video processor (*MVP*) or TMS320C80 is well suited for image and video applications, such as those required for *IPE* [(see (14,15)]. This single-chip parallel media processor performs over 2 billion operations per second. However, key architecture considerations make this processor ideal for imaging applications. The architecture was tuned for algorithms like MPEG, JPEG, 3-D graphics, DCTs, motion estimation, and multi-tap, multidimensional filters. Several manufacturers of computer add-in boards developed solutions featuring the TMS320C80's image processing abilities, including Ariel, Loughborough Sound Images (*LSI*), Matrox, and Precision Digital Images (*PDI*).

The elements within one MVP include four advanced DSPs (ADSP) parallel processors (*PP*), an embedded RISC master processor (*MP*), instruction/data caches, floating point capability, crossbar switch to connect the many on-chip memory modules, I/O transfer controller (*TC*), *SRAM*, DRAM, and VRAM control, pixel/bit processing, host interface, and JTAG/IEEE-1149.1 and emulation. The keys to the MVP's high performance include an efficient parallel processing architecture. It uses a tightly coupled array of heterogeneous processors, each of which perform complete operations in a single cycle (but with some internal pipelining), and each executing different instruction streams. This architecture forms a specialized *MIMD* (multiple instruction, multiple data) architecture to achieve high performance in image computing. The individual *ADSP*s were optimized for fast pixel processing by tuning their parameters for image, video, and graphics processing. Features like certain combinations of subtract followed by absolute value in the same cycle and bit flags to control operations are keys to its performance. In addition the chip uses intelligent control of image data flowing throughout the chip via a crossbar connection of each processor's memory, accessible in one cycle, and a complex I/O processor. This I/O transfer controller is capable of directly accessing external memory as multidimensional entities in external or internal memory (e.g., strips of data anywhere in memory), relieving the internal processors from the burden of computing every pixel's address in memory. Finally the MVP contains most functions needed in a single chip, forgoing slower chip-to-chip communications for several applications. With this collection of elements already contained in the chip, very compact and cost-effective digital image computing systems have been constructed. Reductions on the order of hundreds of components to one have been realized. Examples include a fingerprint identification system and an ultrasound imaging system [see (11)], and a group video conferencing system integrated into a CRT display (Sony).

Programming Media Processors and VLIW. Early media processing chips like the 320C80 required detailed assembly coding to port imaging algorithms onto the processor to achieve peak speeds. C compilers could create code for each of the multiple internal processors. However, to fully utilize the power of the processor, one manually partitioned algorithms to each processor without the help of the compiler. This would consume many months of design by programmers very skilled at coding in assembly language. The root cause of the complexity of programming these MP stems from the architecture itself, in which the programmer must consider optimal usage of many restricted and orthogonal resources (e.g., different combinations of register usage and functional data paths). However, with use of fixed latencies for each operation like a conventional DSP, the programmer could then systematically piece together a solution.

More recent MPs feature very-large-instruction-word (*VLIW*) architectures (TriMedia in Table 1). In this case all functional units and registers comprise a large pool of resources that are *automatically* allocated by a compiler. C compilers improve the programmer's ability to map the algorithms onto the processor by automatically performing instruction selection, scheduling, register allocation, and dealing with the unique latencies and any specific load/store dependencies. A good compiler will find parallelism, fill delay slots, and accurately map the application code to run efficiently on the processor. However, coding efficiency, or the resultant speed of the processor when using compiler generated code, *may* not be as good as detailed assembly coding. Fortunately compiler technology continues to improve.

Image-Adaptive Algorithms with Programmable Processors. The computational efficiency of media processors can open new applications to the system designer, including image pre-processing, scene change detection, object tracking, and recognition. For example, when using a programmable solution, data-adaptive routines such as Inverse DCT will have reduced computational complexity if the program ignores zero coefficients, saving adds and multiplies [see (16)]. With the ability to perform image-adaptive functions, system developers can create new applications, such as scene content processing, adaptive contrast, and real-time video object overlay with these media processor chips. Developers can use real-time object detection and image warping for insertion of region-specific information like text with local languages, news, weather, or advertisements directly into the scene. For example, telecasts of ball games can use video messages inserted directly onto the ball field, onto a players back, or on actual billboards that line the field but with different messages for different broadcast regions.

Video and Image Compression Equipment

Video and image compression equipment utilize digital compression methods to reduce storage and/or communication bandwidth requirements. This equipment includes videophones, group-room video conferencing systems, and entertainment video distribution equipment such as digital broadcast satellite (*DBS*) systems. These products can create stunning digital pictures for application from telemedicine to home theater entertainment with video decoders and encoders at each site. Video conferencing systems can offer lifelike interaction with pictures of persons in remote locations. One such application, teledining, creates a virtual dining table for simultaneous dining with others in cities around the world. In contrast, DBS systems have a few centrally located, high-quality digital encoders preparing TV programs for “wide-scale” distribution, a satellite delivery network and millions of inexpensive digital video decoders, one or more in each subscriber’s home, with minimal interactivity.

Video Conferencing Systems. Video conferencing systems include (1) consumer-oriented videophones like the ViaTV (produced by 8×8 Inc.) or the C-Phone, (2) PC desktop systems and (3) group-room systems like the Concorde (produced by PictureTel Inc.) as described in Refs. 17 and 18. Video conferencing systems completely integrate image sensing, computing, and display subsystems. They also include full *codecs* (encoders and decoders) with point-to-point network connection. To place video calls between several sites a multi-point control unit (*MCU*) serves as a centralized codec, decoding each line, continuously selecting and distributing the video associated with the most active speaker to all participants. Each audio channel is simply summed and redistributed to all callers. These video conferencing systems typically communicate via telephone lines, including packet switched networks (using ITU H.323) and circuit switched (using ITU H.320), as described in a later section of this article. Figure 5 illustrates uses of these various systems.

DBS Systems and *MPEG-2* Encoders. Other digital video and image coding systems include DBS and MPEG-2 encoding. DBS systems exist worldwide including, in the United States, the Direct Satellite System (*DSS*) by RCA or Sony (or DirecTV) and Echostar’s satellite dish systems, PerfecTV in Japan, and Canal+ or TPS in France. While DSS methods have typically been used to deliver digital quality video to homes, recent products also offer Internet access via a low-speed phone connection and high-speed multiplexed data distribution via the satellite. Higher-quality MPEG-2 encoding systems are used for compression of the video with DBS systems (Compression Labs Inc. and DiviCom Inc.) or other PC-based processor boards (produced by Minerva and Optibase).

Internet Streaming Video. As shown in Fig. 5, another application centers on streaming video media to clients via local intranets or the Internet. Programs or movies can be downloaded from a server and played to multiple viewers with multicasting techniques. Local area networks (*LAN*s) may have sufficient bandwidths for

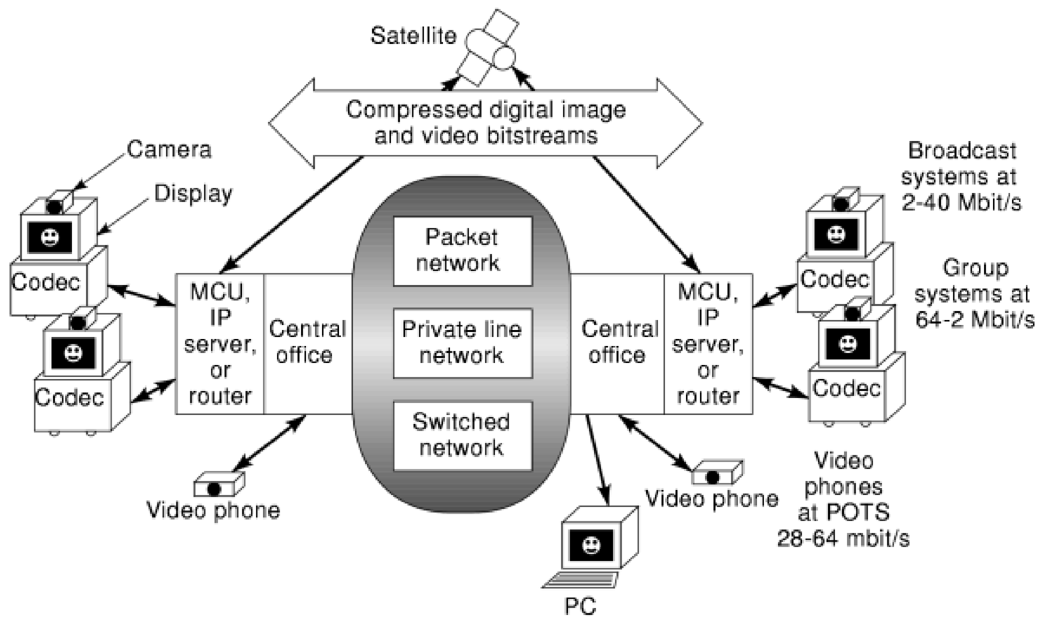


Fig. 5. Digital compressed video equipment utilizes a variety of network infrastructures to deliver video/image services to many clients. Video and image compression equipment exchange compressed digital image and video bitstreams via satellite, packet networks, private line networks, or switched circuit networks, accessed by either the central office in plain-old-telephone-service (POTS), Internal protocol (IP) servers, or routers.

this application, but web-based video streaming presents even lower bandwidths and the difficulties imposed by unpredictable slowdowns or interruptions of service. This application demands very highly compressed video, dynamic buffering, and decoding to present a viewable sequence. In a streaming application, users view the beginning of the movie, while the remainder of the clip downloads in the background. The buffering tends to smooth out variations in transmission rates. RealVideo from Progressive Networks (http://www.real.com) and VivoActive (http://www.vivo.com) are the most popular video streaming solutions today. However, video quality suffers with today's compression technology, typically achieving only 160 $\times$ 120 pixel frames at 3 to 5 fps.

Video and Image Coding Standards. The industry standards for video and image compression (ITU H.261, H.262, H.263, JPEG, JBIG, and ATSC's HDTV) have advantages for specific implementations or markets. Table 2 shows how these compression standards differ. Standards are key to many applications like video conferencing. Interoperability that results from use of standards encourages consumer acceptance of the technology. A legacy of over 50,000 H.261 group video conferencing systems are installed throughout the world, encouraging system developers to keep backward compatibility to those systems. However, while all H.261 systems can interoperate, the quality of the video varies as a function of the manufacturer's implementation of the standard.

H.262 (MPEG-2) excels when higher channel or line rates ($>$2 Mbit/s) are used, especially with interlaced video. With enhancement modes like overlapping motion vectors and PB frames, H.263 has the highest coding efficiency, which is especially useful at low bit rates (e.g., 20 to 30 kbps using H.324 POTS), however also beneficial at higher rates (1.5 Mbit/s). New enhancements, called H.263+, continue this trend in image coding efficiency improvements or increased perceptual quality, as described in Ref. 19 and the University of British Columbia (at www.ece.ubc.ca/spmg). Consequently, most of these 263+ enhancements are at the expense of additional hardware. Numerous nonstandard algorithms are in use today, like CTX+ (from CLI), SG4 (from PictureTel), wavelets, and fractals techniques; however interoperability issues limit their wide spread acceptance.

Table 2. Image and Video Coding Standards

ITU Standard Name	Image/Video Coding and Effects	Typical Application	Compression Ratio or Data Rates
H.261	Quality video at low bit rates and moderate motion	H.320 video conferencing	100 : 1 to 2000 : 1, N 3 64 kbps to 2 Mbps
H.262 or MPEG2-MP/ML	Quality with motion-intensive video at higher bit rates	TV, streaming video, video conferencing	200 : 1, 4 to 10 Mbps
H.263 and H.2631	Quality video with very low bit rates	H.323 video conferencing	100 : 1 to 2000 : 1, down to 28.8 kbps
MPEG4	Combined 3-D-graphics objects and video	Streaming video and video conferencing	100 : 1 to 2000 : 1, down to 28.8 kbps
ATSC or MPEG2-MP/HL	Quality high-resolution video at higher bit-rates	HDTV	200 : 1, 20 to 40 Mbps
JPEG	Very high quality continuous-tone still image	Image compression, printing	20 : 1
JBIG	Very high quality bi-tonal still image	Image compression, printing	
Group4 FAX	Bi-tonal FAX	FAX	

These proprietary methods can have value to closed systems like national defense or scientific applications. Unfortunately, the uniqueness of these algorithms drives special hardware requirements. On the contrary, implementation of mainstream *codecs* presents a wider range of chip solutions and lower implementation costs.

Designing High-Quality Digital Compressed Video. In applications where the quality of the compressed picture is paramount numerous system considerations are important. Proper system design increases the perceptibility of the video and reduces annoying artifacts, particularly when communicating with limited bandwidth. These design considerations include square versus rectangular pixels, pixel clock jitter, system noise, image scaling, interlace to progressive conversion, frame rate, gamma, brightness, contrast, saturation, sharpness, artifact reduction, and combining graphics with video.

Design Consideration for Square Versus Rectangular Pixels. While digital cameras can reduce system costs by eliminating the analog-to-digital converter, they can also create complications in the use of the data. If a camera uses square pixels such as with 640 × 480 resolution, horizontal scaling must be performed in a H.261/263 video conferencing system to meet the 352 pixel horizontal resolution (in addition to vertical scaling to conform to the 288 pixel common intermediate format, *CIF*, vertical resolution). Performing this scaling will increase the software or hardware costs of the codec and reduce performance somewhat due to scaling artifacts or blurring. In addition some digital cameras produce only 320 × 240 pixel images due to cost limitations and bandwidth constraints of the interface bus (e.g., Universal Serial Bus, *USB*, at near 12 Mbit/s), demanding image scaling when used with standard video conferencing.

Design Consideration for Sharpness and Compressed Artifact Reduction. At lower bit rates, the MPEG and H.26× compression process will exhibit annoying block and mosquito artifacts. These artifacts can be removed with video postprocessing. The 8 × 8 blocks are a result of the discontinuity of the image as it's broken into a mosaic. Mosquito artifacts occur due to coding errors in the high-spatial-frequency regions of the image, such as on edges. These edge artifacts change from frame to frame, creating an effect of dots dancing around edges. In addition the compression process can reduce the overall sharpness of the image. Simply sharpening the image will inappropriately enhance the mosquito and block artifacts. Edge-preserving mosquito and block filtering followed by image sharpening will improve the performance as described in Ref. 20.

Design Consideration for Pixel Clock Jitter. During quantization of analog video or NTSC/PAL/SECAM video decoding, noise can be created by jitter of the pixel clock. This can waste bits by compressing false signals. If greater than 10 ns, this jitter can create temporal differences that will be interpreted as frame-to-frame movement by the encoder, reducing overall coding efficiency.

Design Consideration for Low-Level Imaging System Noise. Noise can be injected into the imaging system from several sources, including random and 1/*f* noise, quantization noise, or finite precision effects. Temporal filtering as shown below can improve coding efficiencies by an order of magnitude, but can also create artifacts. Rapid motion will result with a somewhat blurred image.

$$\text{Pixel}\, x, y(\text{frame}\, n+1) = \frac{[\text{Pixel}\, x, y(\text{frame}\, n+1) + \text{Pixel}\, x, y(\text{frame}\, n)]}{2} \quad (5)$$

Design Consideration for Image Scaling. Converting typical camera resolutions of 720, 640, or 320 horizontal pixels, with 480 (US) or 576 (Europe) vertical rows or lines of pixels, demand the use of scaling to interoperate with other systems in the common intermediate format (*CIF*) or 352 × 480 or other display formats. Conversion of the US's 480 lines to 288 CIF resolution will require a 10 line to 6 line conversion. Simple nearest-neighbor algorithms produce distorted outputs with harsh aliasing effects for line drawings (21). Bilinear interpolation filters alias the detailed content of the image. Use of anti-aliasing filters or higher-order polynomial filters will correct many of these artifacts, however at considerable implementation cost (22).

Design Consideration for Interlace Versus Progressive Scanning. An interlaced image results when a single image is split into two sequentially transmitted sub-images or fields, each of which have only the odd or even lines. Depending on the system design, interlacing can create a temporal discontinuity in the sequence of images. It is usually manifested as every-other-line jagged edges in the parts of the scene undergoing motion. Today interlaced video (driven by the conventional television broadcast standards) leads the industry but creates complications for some international video compression standards which were based on the use of progressive images. A simple approach to the problem, namely dropping a field on input, reduces the image's resolution. Artifacts include an "aliased" appearance for images of nearly horizontal line structure. For example, the corners of a room or fluorescent lights appear to have stair steps on their edges. Converting the two fields to a progressive image prior to compression yields the best results but requires complex and expensive motion-compensated interlace-to-progressive methods. New digital television standards support both progressive and interlace methods.

Design Consideration for Frame Rate. Most cameras produce 30 frames per second video, matching the format of TV displays. Feature film is typically created at 24 fps (see http://www.smpte.org). Telecini equipment (e.g., produced by Rank) uses a technique called 3:2 pull-down to convert 24 fps film to 30 fps. In essence, three field samples are taken of one input 24 fps frame to create three output fields (two of which make one output frame) and two field samples are taken from the next input 24 fps frame to create the next output frame. One visual problem created with 3:2 pull-down methods manifests itself as a stutter of the objects in the video when panning quickly from side to side. Coding efficiency can improve if the encoder recognizes the "repeated" frames and does not code them.

Design Consideration for Display Effects: Gamma, Brightness, Contrast, and Saturation. Prior to the display of de-compressed video, significant *adjustment* may be needed to adjust for differences in the diplay used. The amount of processing depends on the characteristics of the display used and the characteristics of the source material. For example, when a video sequence is played on a computer's RGB display, the difference in display gamma from a conventional NTSC TV CRT will most likely result in a lack of visibility of detail in dark scenes. Gamma correction will remedy this problem by adjusting the intensity of each pixel to compensate for the difference. This can be accomplished by using a simple lookup table. Likewise, adapting to the video signal's setup, color demodulation errors and gain or linearity problems in the system's amplifiers can greatly enhance the picture quality.

Design Consideration for Combining Graphics and Video. Several problems occur when combining graphics and video. Video is typically overscanned on the TV by more than 4% to keep the viewer from seeing blanking signals on the screen when the set has aged to the point that the scanning electronics effectively shrinks the picture. Proper scaling of the image or cutting out blanking signals is key to a clean edge on the picture. When placing *RGB* graphics on an *NTSC* display, proper treatment must be given to the color gamut differences. For example, white text will appear pinkish on an *NTSC* monitor if not properly corrected. Blending of graphics and video can also pose a problem in terms of aliasing of lines. Finally graphics can flicker in video displays due to the interlacing effects without proper adjustment.

Video and Image Quality Validation. While design of video and image compression equipment presents many challenges, perhaps the greatest remains the determination of picture quality in compressed image applications. The most common metric used to measuring image quality has been the peak signal-to-noise ratio (*PSNR*). PSNR is calculated by comparing the processed image to the original or reference image as follows:

$$\mathrm{PSNR} = -10\log_{10}\left[\frac{255^2}{\mathrm{MSE}}\right] \tag{6}$$

where

$$\mathrm{MSE} = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}[\text{Processed pixel }(x,y) - \text{original pixel }(x,y)]^2 \tag{7}$$

MSE is the mean square error. *PSNR* has many limitations, since the value does not necessarily increase for errors that a viewer can easily see, in comparison to those that may be masked by the human vision system.

A test equipment manufacturer (Tektronix at http://www.tek.com) has produced the PQA200, which produces a single objective and numeric value corresponding to a picture-quality rating (*PQR*). The PQR utilizes technology developed to measure just noticeable differences (*JND*) of human vision (Sarnoff at http://www.sarnoff.com). The process transforms each image according to the human visual system, which includes a linear system with point nonlinearity and could include masking models for color constancy, chromatic adaptation, lateral inhibition, and Mach bands (23). A JND image or map is produced corresponding to the problems or errors in the image. A totally dark JND image corresponds to a perfect picture with respect to the original image. The PQA200 utilizes multiple TMS320C80 processors (described earlier) to process the JND image in less than a minute, rather than the alternative of using a panel of expert viewers performing analysis and calculations for many weeks to achieve the same result.

Image Database Equipment

Image database equipment (*IDE*) provides solutions to the storage and access of images and video files. Traditional database methods of using keyword labels to access text data have been replaced with methods like text labels (i.e., the database “contains a picture of Lou Whittaker on Mt. Rainier”), thumbnail images, interactive 3-D panoramic views, and content-based archival and retrieval. Refer to Ref. (24) for more details. The common elements of all *IDE* include (1) compression, (2) massive storage to hold the image files, and (3) communication networks for distribution of the data. In a typical application computer servers contain image objects that are played back by clients utilizing a communication network. These networks can be as simple as a telephone connection to the Internet or a local area network (LAN) and as complex as asynchronous transfer mode (*ATM*) networks. We describe common computer networks, some image database applications and image content retrieval.

Computer/Telephony Networks. Traditional computer networks, such as Ethernet, token-ring, and *FDDI* (Fiber Distributed Data Interface), create cost-effective LAN connections for sharing of image databases within a building or adjacent buildings. The limitations of LAN networks are bandwidth (10 up to 100 Mbit/s), number of clients and topology. For example, if dozens of clients attempt to access large images on a server, the effective bandwidth decreases and latency increases by that amount. In addition, if the topology has each computer connected in a ring, then each hop between computer on the ring can slow response. Star or centralized topologies where a central node has direct connections to each client machine minimize the hops or latency. For smaller networks the fully distributed topology is effective with each machine interconnected via a direct link. Hierarchical networks create extensions to the centralized network with a series of intermediate nodes, forming a tree structure. For interactive video conferencing with face-to-face communication, the round-trip delay (from the sending camera, through compression, transmission over the network, decompression and display on the receiving display and then back the other direction) must be less than 300 ms. For this reason centralized topologies are more useful for ITU-H.323 video conferencing over LANs.

Today's wide-area networks (*WAN*) offer connection of remote client equipment with servers (e.g., with Internet servers). These networks include circuit-switched telephony networks like analog POTS, integrated services digital network (*ISDN*), T1, and asynchronous digital subscriber loop (*ADSL*). Their advantages include predictable delay and quality of signals. Video conferencing using ITU-H.320 operates within these telephony networks. Alternatively, gateways can be used to place IP packets on these WANs within the ITU-H.323 standard (from Radvision). In a high-speed packet switched networks, like ATM (broadband-ISDN), the topology is less important, since the data are broken into packets for transmission via one of a multitude of different paths to the destination. The largest issue with these networks is that of long and unpredictable latency for those packets that travel the long routes to the server.

Modems and the Internet. Modems (Modulator-DEModulator) connected to plain-old-telephone-service (*POTS*) phone lines enable communication with digital signals. Today's modems (by 3Com and Rockwell) can achieve up to 56 kbit/s from the server and 28.8 kbit/s from the client via *ITU* standard V.90 and V.34 modems (28.8 both ways). The Internet has been facilitated by use of these client modems in the home, T1 networks in the office, all connected with multi-line modem servers operating with the *TCP-IP* (transfer control protocol-Internet protocol). Routers and hubs interconnect the variety of network connections within a particular organization before connection to the Internet. The servers are then connected using T3 connections to the backbone of the World Wide Web, providing high-speed transfer of the data packets.

Cable, Satellite, and Broadcast Networks. In addition to computer and telephony networks, video and image communication frequently occurs via cable, satellite, and terrestrial broadcast. As these networks increasingly utilize digital images and computing costs decline, an increasing percentage of IPE uses these one-to-many networks today. Broadcast network solutions offer high video bandwidths, supporting hundreds of digital video programs simultaneously. Unfortunately, their topology with a single server and many clients limits interactive communication. However, with the large quantity of channels available, image and video database applications are emerging to selectively download the client's requested data to a local disk, such as overnight (e.g., WebTV).

Network Demands of Images, from FAX to Video. Effective communication of image and video demands a variety of average bit rates and buffering to compensate for differences between the target rate and the actual rate. Communication with high-quality images may dictate up to 30 Mbit/s, whereas *ITU* Group-4 FAX could utilize merely 64 kbps or less over *POTS* phone lines. Moving a typical $640 \times 480 \times 24$ bit color image via the Internet with POTS lines could take several minutes. Sending video over networks requires (1) 64 kbit/s for *low-quality* video when using H.261, (2) 6 Mbit/s for *TV-quality* video when using *MPEG-2* MP/ML, and (3) 10 to 40 Mbit/s for *HDTV-quality* video with MPEG-2 MP/HL. Data buffering compensates for dynamic latency in the network or variations in peak network performance. Most video applications require

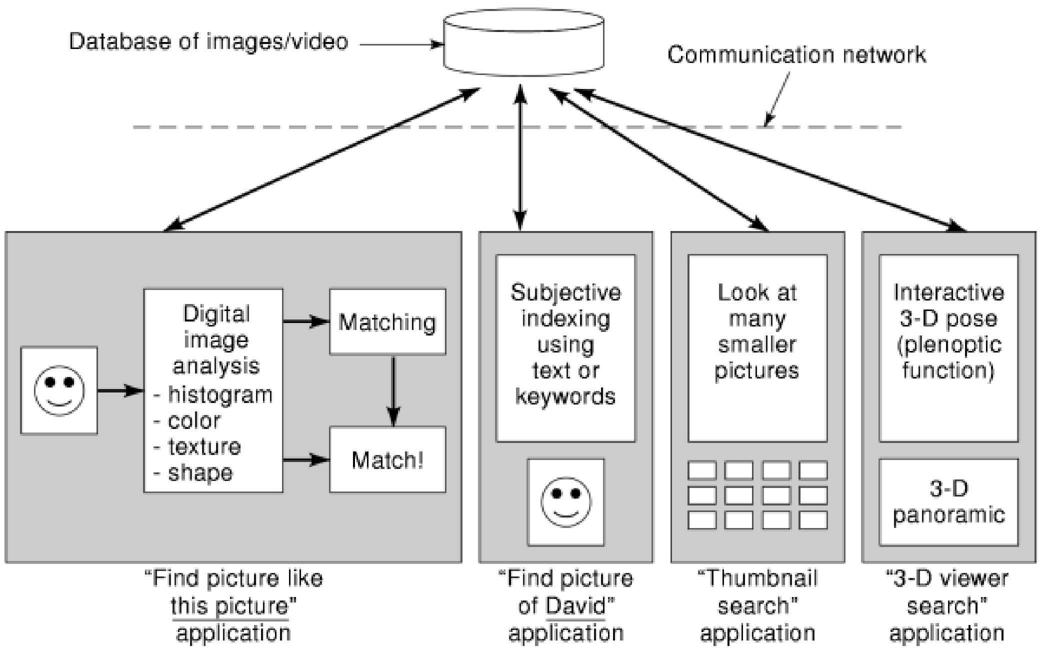


Fig. 6. Image database equipment utilizes various search applications (or engines), coupled with image/video databases and a communication network, to help users find useful images or video.

access and display of a continuous stream of images without frame drops. The size of this video buffer depends on the type of network (e.g., one-to-many or many-to-many), the required data rate and traffic on the network. Some applications can tolerate missing data but only if occurring in a predictable and infrequent fashion. In that case minimization of the frequency and duration of the frame drops increases the viewing experience [see (25)]. Traditional networks are designed to minimize errors and correct any that occur. Ignoring errors and not waiting for retransmission of bad data can improve the overall data rate realized, but input image quality will suffer.

Image and Video Content Indexing and Retrieval. With a wide variety of video and imaging material accessible either via the Internet or other electronic libraries, browsing, searching and retrieving the interesting or appropriate information can become challenging. As shown in Fig. 6, methods utilized include accessing image databases with textual headers or thumbnails. Searching by keywords, a traditional approach to databases, only has value for image databases if a textual description of each image's content (stored in the database with the images) contains pertinent information. This can include physical (size), perceptual, conceptual, objective, subjective, or statistical data.

Methods for retrieval of video increase database searching complexity due to the amount of data searched and the inter-relationship of each frame. Retrieving video clips, browsing video databases, and managing video libraries can be less than satisfying without an algebraic video model and multimedia hyperlinks [see (26)]. The video algebra keeps a working record of operations like how the video was constructed, how it was manipulated (stretched, differenced, summed, etc.), and what was composited or linked together to make the material.

Various techniques have been developed to quickly locate pertinent images or data. Current work on the ISO MPEG-7 standard (Multimedia Content Description Interface) should yield a formal standard for such image retrieval techniques. The scope of MPEG-7 includes defining a standard description of images and video for interoperable use by several search engines and feature extractors. Techniques include *content labeling* or *indexing* where headers or tags contain text which in some way relate to the content of the image [(see (27)]. Unfortunately, these labels are quite laborious to create and may not actually catalog the pertinent information. This subjective indexing has limited usefulness since it's difficult to anticipate the content that will interest all users of a particular image or sequence. However, these techniques can be useful for finding all the photos of "David Lettermen," but most likely not find those images that look similar to David, since they probably wouldn't have been stored that way.

Software tools are commercially available for determining a picture's characteristics like color, texture, or shape of objects. Solutions such as Excalibur's Visual RetrievalWare and AAI's (Amerinix Applied Imaging) KBVision software system use image understanding techniques or adaptive pattern recognition processing (*APRP*) to allow on-line identification and indexing of images and sequences. The software permits query-by-example searching. One can search for patterns, colors, or a picture that looks similar to a reference picture. Since only patterns are correlated, rather than the underlying object descriptions, a complex object can elude detection by the software. Yahoo!'s Image Surfer site utilizes Excalibur's *APRP* for Internet searching of items like charts, photos of actors, and movies (see http://ipix.yahoo.com for examples).

Other methods include use of thumbnail images, with lower-resolution snapshots of a larger image (e.g., using JPEG), and faster decoding of MPEG sequences by processing only portions of the compressed bitstream. Other techniques include scene change detection methods to aid the surfing of video sequences. Dramatic changes in motion, color and sound can indicate a scene change. Using only a single frame per scene (located by the scene change detection algorithm) facilitates rapid searching of the database.

Web-Based Imaging Database Application. Replacing home photo albums with a virtual photo album, featuring distributed storage of electronic pictures somewhere on the World Wide Web has emerged. Kodak originally produced the FlashPix architecture for Internet imaging. Users upload digital images via an Service Provider to a database managed by the service provider (Kodak in this case). Pictures can be edited, placed into a virtual photo album, sent via email or selectively accessed by others on the Internet. The use of titles and thumbnail images can minimize the bandwidth requirement when searching through your image database via the Internet. A recent consortium of companies called Digital Imaging Group (http://www.digitalimaging.org) are continuing to expand this Internet Imaging Protocol (*IIP*), enabling efficient distributed imaging database solutions using FlashPix and other tiled image formats.

3-D Imaging Database Application. Using virtual reality concepts and the Internet, another form of image database type has been created for visualization of 3-D images. These products implement a 5-D space, called the plenoptic function, projecting pixel color as a function of position and pose (or gaze) into the 3-D space. QuickTime-VR (by Apple Computer at http://www.apple.com) represents a view into a 3-D space allowing the observer to rotate 360 degrees left or right, and to shift up or down, all from the same vantage point. It's very compelling to spin through 3-D environments with this technology.

The user can selectively pan, tilt, and zoom throughout the 3-D space from the perspective of the original camera. In addition users can link to other positions in the scene that have also been captured with the same 3-D techniques. With proper design and capture, users can wander throughout a structure with a true to life feel. Live Picture provides an immersive imaging database for users to view 3-D panoramic scenes (similar to QuickTime-VR), but with the added feature that maintains resolution for selected objects as zooming progresses (see http://www.livepicture.com for examples). In effect users can virtually fly across town, into a building, through a room and look at the content of a television within that room, all while maintaining good display resolution.

New image capture techniques have been employed for this application. Rather than constructing a complex panoramic camera (e.g., those used with the popular Panavision technology in movie theaters) these images can be collected with a variety of new cameras using a special fish-eye lens. The camera captures a distorted view of the scene in a 360° circle. Subsequent image warping corrects for any distortion created by the lens. Alternatives include capturing successive images or swaths of the scene with overlapping gaze and later creating a mosaic of the entire 3-D scene. Special camera mounts maintain parallax by maintaining all rotation about the plane of the sensor. Subsequent processing of the images seams the images into a large panoramic image or a compressed file for later expansion by a QuickTime-VR application. To correct for optical distortion or misalignments of the images, image warping is performed (see the section titled "Geometric Distortion and Correction with Warping" earlier in this article for a description of this warping function).

3-D Model Database Applications. Utilizing 3-D scanners and graphical representations several applications have emerged to offer a visual walk through 3-D databases, such as a human body. These techniques have also been used in the apparel industry to scan customers for specialized fitting of clothing. Another application features a virtual globe, where users retrieve the world's historical information by interacting with a 3-D model display of the world, with successive clicks of a computer mouse to enable searches of a specific location at specific times. Multiresolution analysis and progressive meshes provide two unique techniques for representation and rendering of the 3-D image data. Each of these provide solutions for rapid display of lower resolution representations, followed by higher-resolution meshes or textures when the data become available over low-bandwidth networks. Interactive computer video games (e.g., Bungie's Myth) provide Internet-based multiplayer sessions with 3-D meshes rendered in real-time somewhat independent of the client computer's processing capabilities and the Internet's available bandwidth.

Image Display Equipment

Image display equipment provides the mechanism to view raw or processed images or video. This equipment includes computer displays, televisions, studio monitors, or projectors utilizing a variety of display technologies. (See Ref. 28 for an overview of displays and Ref. 29 for a description of emerging television display technologies.) Display technology is very application dependent. The number of viewers, the amount of power or space available, the number of colors, and the rate of image display of the application all drive unique display requirements. The quality of the display directly influences the overall performance of an image processing system. Some display equipment, like advanced consumer televisions, incorporate image processing functions directly into the equipment, conditioning the video signals for any display-dependent effects or adding features that may improve the quality of the displayed pictures. In many cases this processing can destructively interact with the image or compression processing which has occurred prior to display. Examples include special noise reduction filters, sharpness filters, or gamma correction functions, each of which may enhance artifacts in digital images, like block artifacts in MPEG/JPEG compressed images. In general, users and developers of IPE must be familiar with the features and limitations of the display equipment used.

Interlaced and Progressive Scanning Methods. Two predominant types of display scanning methods exist today—*interlaced video displays* such as NTSC, PAL, or SECAM format televisions used throughout the world and *progressive graphics displays* such as VGA format monitors typically used with today's personal computers. Both displays range in size from handheld to large wall-sized units, depending on the display technology employed. However, as displays increase in size, higher performance results with progressive techniques.

Television's interlaced video displays resulted from the limitation of transmission, processing, and display bandwidth of the video systems developed over 50 years ago. Interlacing techniques display only half the lines (odd lines) or rows each field (i.e., every other line), followed by display of the other half of the lines (even lines) the next field time. The benefit of interlacing is that it uses only half the bandwidth of progressive displays. A slight flickering artifact can result, coupled with more motion effects since the number of scan lines is halved for moving objects.

Computer graphics' progressive displays provide full scan displays with a resultant minimization of display artifacts. In contrast to the fixed frame rate of NTSC or PAL (60 or 50 fps), the frame rate of progressive graphics displays will vary depending on the system design. In addition, since the viewer of computer graphics typically sits within an arm's reach of the display, faster display refresh rates (70 to 80 Hz) are necessary to prevent off-axis flicker for most viewers. Graphics displays feature square pixels for geometric precision, whereas conventional broadcast video formats use nonsquare pixels.

Direct View versus Projection Methods. Display systems are also characterized as either *direct-view,* where the viewer looks at the display device (e.g., a face plate on a *CRT* or *LCD* panel), or *projection,* where the viewer looks at an image projected from the display device (e.g., *CRT*, *LCD*, *DMD*, or other). Projection display systems use optics to enlarge the display device's image onto a screen or directly into the viewer's retina for "heads-up" display. Projection techniques have also been used to present a virtual image that seems to float out in space, being superimposed on the real scene. Direct-view displays can have the advantage of minimal optical distortion and brightness. The trade-off also includes the size and cost. Laser scanning creates another type of display, particularly in the creation of 3-D objects, however, is quite limited due to the serial nature of scanning coupled with no phosphor retention of the scene. Some display devices created with semiconductor technology (like *DMD* with displays near one square inch) dictate the use of projection techniques for enlargement for groups viewing.

Front versus Rear Projection Methods. Projection displays either utilize front projection or rear projection techniques. Electronic front screen projectors (from Sony, Proxima, Barco, etc.) operate analogously to conventional film projectors (e.g., slide projectors, overhead projectors, or cinema), whereby the image is optically projected onto a reflective surface, such as a white wall or glass-beaded screen (from Da-Lite, Inc.). The projected image originates from the front of the reflective screen, with the projector usually somewhere in the room (or behind a wall or ceiling). Rear screen projectors (from RCA/Proscan, Sony, Mitsubishi, and Pioneer) illuminate a screen from behind and *usually* contain all the equipment within a movable enclosure that looks like a large TV enclosure. An obvious advantage of rear over front projection is the reduction of occlusion of the screen by those in the room within the line-of-sight of the projector, whereas the advantage of front screen can be portability. Rear screen projectors achieve high levels of brightness with customized screens using Ferznel lens to focus the light, usually at the expense of viewing angle—one disadvantage of rear screen projection techniques. In particular, consumer rear screen projectors have been designed for optimal viewing brightness when viewed from a seated position while dimming significantly when viewed from a standing position.

CRT (Cathode-Ray Tube) Displays. Conventional direct-view and projection displays use CRTs as a good balance between cost, performance, and lifetime. With CRTs, a single high-energy beam, scanning an x–y grid and charging a phosphor screen can create brilliant displays in fully lit rooms.

Direct-view TVs continue to exhibit high performance over other technologies like projection. Current limitations are near 40 in. diagonal in size (Mitsubishi). As one increases the size of the CRT display, the geometric limitations of pointing one beam at a glass screen dictates a deeper tube and curved surface. As well the pressure of the atmosphere dictates thicker and thicker glass to preserve a vacuum in the tube. Difficulties include the manufacturing limitations of glass and the depth growing to exceed standard consumer door widths, complicating installation of the TV.

Likewise computer monitors continue to feature CRTs for the same cost and performance reasons. However, several applications make demands on a display that CRTs cannot address. Mechanical design or architecture demands higher resolution graphics and geometric accuracy. Group activities require bigger screens. Laptop applications require flat displays. For these reasons, flat-screen display panel technologies have seen recent development, as described in Ref. 30.

LCD (Liquid Crystal Display). Liquid crystals displays (*LCD*s) use a panel of light modulators, imaged either by direct viewing of the panel (as with most laptop computers and handheld image display equipment) or by projection [for more details, see (31,32)]. Each of hundreds of thousands of light valves (each pixel) either passes or blocks light, depending on the orientation of microscopic crystals contained within each valve. The amount of light passed through each valve determines the shade or color of the pixel. The crystals are aligned by applying electromagnetic fields. Of the various methods used to produce LCDs, the most widely-accepted consists of an "active-matrix" of transistors for modulation of the crystal orientation. LCD panels today typically range from a few inches to 14 in., with some as high as 21 in. A manufacturer (Sharp) has actually seamlessly joined two 29 in. displays to make a 40 in. prototype LCD panel.

Threadlike twisted-neumatic liquid crystals are typically used in *LCD*s. In the nonenergized state the crystals (which are sandwiched between two polarizers) align with a twist. This enables light to pass between the two polarizers (oriented 90° out of phase) by twisting the light through the liquid crystals. Supertwist *LCD*s are more than 180° out of phase. When a field is applied, the molecules align perpendicular to the display surface. Light then passes directly through the crystals to the next polarizer. No light escapes since the light from the first polarizer is 90° out of phase with the second polarizer. Color can be generated by using color triads at each pixel, or with the use of multiple panels for projection displays.

LCDs have great advantage for mobile applications due to their low power consumption. In addition they perform well in graphics applications (like PCs) due to the geometric stability and 1:1 pixel mapping of the display grid to generated pixels. This 1:1 pixel mapping, like other panel displays, can yield extremely sharp text and graphics displays in comparison to analog displays like CRTs. These advantages are offset by disadvantages for many other imaging applications. LCDs typically have low pixel fill factors (the percentage of the pixel site illuminated) such that only 50% to 60% of the pixel site is illuminated, surrounded by dark patterning structures. The contrast ratio (or ratio of brightest to darkest images displayable) has also not matched CRTs. High-resolution LCDs are difficult to manufacture, increasing costs relative to other display technologies. LCDs are typically fabricated at nearly 100 pixels per inch. Xerox Palo Alto Research Center produced the first 300 dot/inch LCD with 6.3 million pixels in a 13 in. panel. Despite their limitations, LCDs can create sharp and bright displays when coupled with the appropriate illumination system.

Active and passive-matrix LCDs exist in the market place today. Less-expensive, passive LCDs use row address lines and column pulses, positive and negative about a bias, to signal pixels within a row. Unselected pixels accumulate a "cross-talk" signal as a function of the number of rows in the display. Resolution and contrast performance suffer with passive LCDs. By contrast, active-matrix LCDs use individual transistors (thin-film transistors, TFT) at each pixel to isolate on and off pixels, while maintaining a row/column addressing structure similar to DRAMs. A higher-quality image results from the isolation inherent in active-matrix LCD displays. Recent advances in reflective LCDs offer dramatic increases in contrast ratio, rapid response, brightness and can be used with ambient illumination rather than power-consuming back-lights, with great promise for future applications.

FED (Field-Emission Display). With FED technology a grid (or field) of individually addressable microscopic electron-emitters replace the CRT's single electron beam. The light then strikes a layer of phosphor, causing it to glow at the location of electron emission, forming an image. The advantages of FED include low power consumption, wide angle of view, flat screens, high quality, and spatial stability. Manufacturers, including Candescent Technologies (San Jose), PixTech (Montpellier, France), and Texas Instruments (Dallas, Texas) have produced FEDs. Currently manufacturing issues have limited their use.

Plasma Gas Display. Applications for a wall-mounted, high-brightness, and high-resolution display such as HDTV has spurred development of gas plasma displays. Manufacturers of plasma systems (Matsushita, NEC, NHK, Fujitsu, Sony, etc.) have produced displays up to 40 in.' and only a few inches thick. The advantages of plasma gas displays include large flat panels, high brightness, and spatial stability. The disadvantage with respect to LCDs has been power efficiency and contrast ratio. A gas plasma display's efficiency is near one lumens per watt, between a half and quarter as efficient as backlit LCDs. With plasma displays, an electric current activates a neon gas at each pixel by creating ultraviolet light with a gas discharge (similar to fluorescent lights). The UV light is then turned into visible red, green, and blue light by phosphors. Vertically oriented, phosphor-coated and gas-filled channels contain one electrode. Another layer in the display, an array of horizontal electrodes, crisscross the display. A pixel is then defined by the intersection of two electrodes. As the size of the display or number of pixels increase, the display does not grow in depth, keeping a thin aspect ratio.

Manufacturing issues have affected uniformity, contrast ratio and produced annoying defects. Due to the difficulty of applying pixel information to all electrode intersections simultaneously, pixel modulation methods

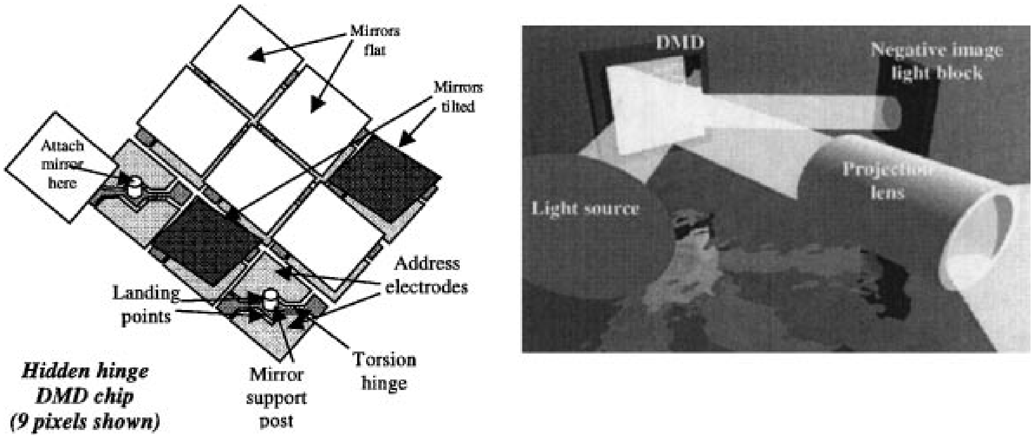


Fig. 7. *DMD* displays utilize a DMD chip and an optics system to create a projection display. The illustration shows a hidden-hinge DMD structure of nine pixels (on the left of the illustration), each of which rotate via a hinge mechanism underneath the mirrors.

are necessary to create shades of gray or color in the display. Addressing limitations had restricted plasma displays to near 6 bits of gray, or 64 shades per color primary, creating significant false contour patterns on the screen. Recent developments have extended plasma displays to 8 bits, some using dithering methods.

High-Gain Emissive Display. HGED displays use phosphor modified to operate at low voltages (100 W) with an electron beam grid, rather than a single high-energy beam like CRTs. Telegen's HGED can produce high resolution with good viewing angles, without the X-ray emissions of CRTs.

LEDs on Plastic. With a recent light-emitting diode (LED) display constructed on plastic new possibilities open for large, flexible displays. These sheets physically deformable plastic can wrap around the viewer or even replace glass-pane windows with images of any scene displayed on large sheets of plastic [see (33) for more information on LEDs-on-plastic].

***DMD* (Digital MicroMirror Display).** A display technology called Digital MicroMirror (DMD), or Digital Light Projection (*DLP*), creates an *all digital* approach to displaying images (from Texas Instruments). In essence, DMDs create a grayscale or color image by modulating a sequence of binary-images posed by an array of microscopic mirrors. Each mirror is smaller than a human hair, usually 17 μm, as shown in Fig. 7. *DMD*s can create brilliant pictures with minimal artifacts. A unique advantage of a DMD display includes its extremely low power of operation, actually creating a still binary image with no power from the light modulator. DaVis Powerbeam V SVGA data and video projector, Proxima, In-Focus, and others have based projector products on TI's DLP. For high-brightness applications, such as those with large-audiences, manufacturers have produced DLP projectors with 2000 to 5000 lm (Electrohome and Digital Projection Inc., respectively).

DMD technology emerged from work with a real-time image correlation system for use by NASA to remotely dock spacecraft. Initially micromirrors were glued onto a polymer sheet that was spatially modulated by transistors under the polymer. Later, mirrors with torsioned hinges were patterned on the silicon device to create a robust binary display. Currently conventional semiconductor processes are used to pattern the mirrors directly onto a silicon wafer. This should lead to manufacturing at a reasonable cost.

Early problems with DMDs, which have since been overcome, include artifacts relating to use of pulse-width modulation to create grayscale and color shades with the binary image display technology. DMDs present a binary image to the viewer at extremely high rates using the viewer's eye to integrate those binary images into a specific intensity per pixel. Each frame of an 8 bit system is broken into 256 time units (1/30/256 s). A saturated white pixel would require the DMD to display the binary image with that pixel "on" for the entire frame time,

Table 3. US HDTV Video Formats

Number of Pixels per Row	Number of Rows	Aspect Ratio	Frame Rate	Comment
1920	1080	16 : 9	60I, 30P, 24P	NBC adopted 1080I
1280	720	16 : 9	60I, 30P, 24P	ABC adopted 720P
704	480	16 : 9 and 4 : 3	60P, 60I, 30P, 24P	MPEG2-MP/ML
640	480	4 : 3	60P, 60I, 30P, 24P	VGA

while a half-gray pixel would only be "on" for half of the frame time. A linear range of display times result. The problem occurs as intensity number 128 (1000000_2) transitions to 127 (01111111_2). In this case, the temporal asymmetry of pulse widths creates an opportunity for the eye to incorrectly integrate the pulses if the eye moves or has its gaze interrupted (as with some object passing between the eye and the display). Solutions include breaking long pulse binary images into smaller pulses and temporally distributing them [described in (34)].

Miniature Displays or *HMD*s. Head-mounted displays (*HMD*s) use special optics and a miniature display to create images that appear to the viewer as a large display at a distance. In effect the viewer sees a 20 in. monitor at a distance of five feet when in actuality a small display is about an inch in front of the viewer. HMDs first appeared in military applications, such as displaying of instrumentation for pilots or realistic flight simulators, and now have entered consumer markets. Some problems with image quality and unsettling physical effects (e.g., dizziness or headache) have limited their widespread use.

Several display technologies have been used for HMDs. One company (Reflection Technology) uses a scanning mirror to paint an image with a linear LED array. Nintendo's Virtual Game Boy used this display. Another manufacturer (Motorola) produced a miniature 240 × 144 LED array for HMD applications. Yet another manufacturer (Kopin) produced HMD displays by direct patterning of light valves, such as AM-LCD, on silicon wafers at over 1700 lines/in. They produce devices at 320 × 240 resolution. With 15 μm pixels, 192 displays can be created from one 6 in. silicon wafer. Backlighting the LCD is accomplished with *LED*s. Color can be achieved with frame sequential illumination of red, green, and blue LEDs. Another approach, with polysilicon TFT LCD mini displays (from Seiko-Epson and Sony), uses RGB color filters. This reduces brightness and resolution by one-third. These LCDs are constructed on quartz substrates, at a higher cost that silicon wafers. Finally, products also feature direct scanning of a laser on the retina of the eye, providing small, head-mounted personal displays.

HDTV Display and *AFD*. While Japan implemented high-definition TV (*HDTV*) and enhanced-definition TV (*EDTV*) systems years ago, the recent digital HDTV standard approved by US government (see http://www.atsc.org) raises questions of what will be the display of choice. Certainly CRT technology continues to offer high quality for most video applications. However, the ideal display for high resolution and consumer costs has not yet been determined. The use of digital transmission, coupled with the advent of digital versatile disk (*DVD*), may make the digital display technologies most desirable. In that scenario, everything would be digital [as discussed in (34)]. The standard for HDTV in the United States will allow 18 different video formats as shown in Table 3.

HDTV in the United States has been a very political issue since broadcasters, cable operators, and TV-set manufacturers must adapt to the new standard—a standard that presents 18 different possibilities for noncompatibility. Others have proposed counters to the standard for alternative use of the digital spectrum. Microsoft, Sun Microsystems, and Compaq favor the introduction of a base layer (DS0) and extensions to the base layer (DS1) so that virtually any format can be supported in the future. The difficulty has been for standard developers to commit to one format as was done decades ago for NTSC. A flexible compromise to this situation has been developed by others (e.g., Hitachi and RCA Thomson), essentially constructing a display to adapt to any of the formats. Their proposed All-Format Decoder reduces the decoding, memory and display requirements of HDTV to those similar to conventional standard definition TV, while displaying all of the formats [see (35)].

TV and Computer Graphics Convergence. Today, televisions and graphics monitors are converging for many reasons. While *analog* techniques and limitations drove video standards and systems, now *digital* techniques are replacing many video functions, including the capture and storage of video, and more recently the display and broadcast of video. Size, cost, error resiliency, and longevity of video data storage started the trend. Digital transmission enables reduced artifacts, no noise problems in broadcast fringe locations, and optimization of the bandwidth for the scene and the display used. Computer graphics displays have achieved sharp, noiseless, flicker-free, and clear pictures. Since computer users sit within arms reach to the monitor, high refresh rates (70 to 85 Hz) are used to prevent the perception of off-axis flicker. In addition they can operate at reduced brightness relative to a television, since light diminishes by the square of the distance from the source. TV requires viewing from across a room, at a distance of five or more times the picture height and higher brightness. Off-axis flicker is not a problem. Now with digital broadcast and display techniques, video broadcast viewers enjoy the same benefits as computer graphics, especially if square pixels are used for video. Therefore graphics, text, and video can merge without the distorting effects of scaling one or the other. As the industry migrates to HDTV to attain the *movie and computer graphics experience* in the home, the challenge becomes finding a display technology that can support expanded screen sizes, maintain high brightness, and increase frame refresh rates to keep flicker distortion at bay. At that point, TV and computer displays will have effectively converged.

Image Printing Equipment

Image printing equipment provides hardcopy reproduction of raw or processed images. This equipment includes laser printers, ink-jet printers, offset printers, and digital copiers. Very affordable solutions exist today for producing exceptional quality pictures composed of text, graphics, and full-color for most applications. The limitations of today's printing technologies are the cost of consumables (ink, paper, toner, etc.), printing time (network, decompressing and rendering time, etc.), and durability of the printed material, with resolution continuously improving. In fact we are rapidly approaching the point of widespread replacement of conventional silver-halide film photography by electronic photography coupled with an ink or electrostatic printer.

Printers. Printers represent a class of image processing equipment that imprints images onto paper or other material for relatively permanent recording in visible form. Similar to scanners, printers usually sweep an imaging device (linear or 2-D display, pen, or inkjet, rather than image collector) across a piece of paper. Printers transport the paper while performing this function, whereas most scanners maintain a stationary page. Figure 8 shows these common elements of a printer system.

Image processing equipment use several methods to deliver images to printers. A popular method includes use of a Page Description Language (*PDL*). A PDL facilitates printing (or display) of a document independent of the source material resolution or the printer resolution. No matter what quality of printer used, in terms of chromatic or spatial resolution, the source image is communicated independent of those limitations. Methods for achieving this include coding object descriptions rather than bitmaps. For example, an object like a line or character can be coded as a sequence of pen strokes rather than a bitmap. Popular PDLs include Adobe's Postscript [see (36)], Interpress, and TrueImage.

Many other image storage formats like Tagged Image File Format (*TIFF*), Graphics Interchange Format (*GIF*), PICT (Apple MacPaint format), BMP (Windows device-independent bitmaps), or others assist with compression and transmission of images to printers [refer to (37) for a more detailed description of the image formats]. The TIFF format maintains high quality images for delivery to other computers, but without significant compression. With only 256 colors, the GIF format is suitable for synthesized images but does not work well with natural continuous-tone images sensed by cameras. ISO's *JBIG* (Joint Bitonal Image Group) and CCITT's Group 3 and 4 FAX standards are used for bi-level images and do not compress gray scale images well. However, with JBIG the images can be broken into bit planes for compression. JPEG is very popular for

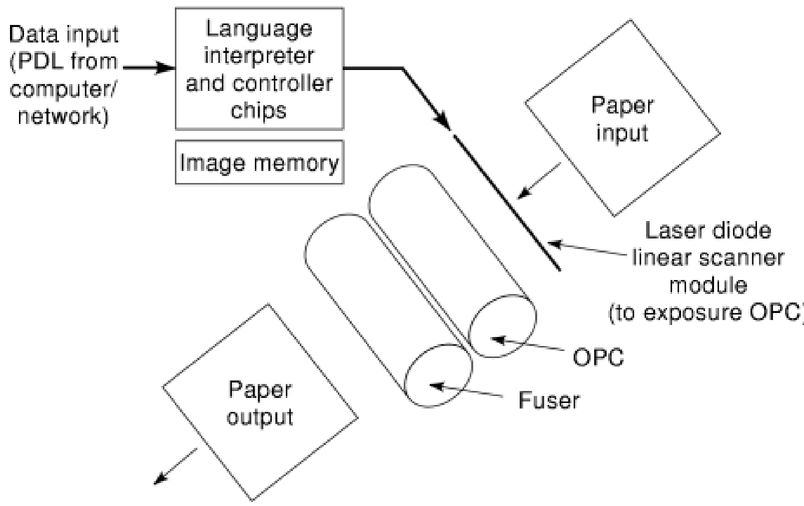


Fig. 8. Elements of a laser printer system include a processor for printer language interpretation, laser diodes, an OPC, and a fuser.

general continuous-tone images, those found in real-world scenes. The JPEG 2000 standard, with techniques such as progressive lossless reconstruction, is intended to serve applications which the original JPEG fails, such as low-bandwidth dissemination of images via the Internet, pre-press imaging, electronic photography, and laser print rendering.

RIPs. Printers use a raster image processor (*RIP*) to render printable images from the PDL input data. A *RIP* consists of the following elements: (1) compositing of images and text, (2) compression and decompression to reduce the cost of the printers by reducing memory requirements, (3) color space conversion [while the final color space is usually CMYK (cyan-magenta-yellow-black), intermediate processing is usually in an alternative color space], (4) filtering to sharpen edges, (5) affine image transforms to rotate and scale images, (6) halftoning or screening to increase apparent color resolution, (7) geometry and spline conversion to smooth fonts, and (8) memory management. Several companies produce a RIP for printer or multi-function printer (*MFP*) applications, including Electronics-For-Imaging (*EFI*) and Xionics. *RIP*s can operate in most computers, resulting in delivery of quality images to a multitude of printers. Communication speed and buffering can limit the effectiveness of these PC-based soft RIPs. A RIP implemented on fast hardware embedded within a printer offers printer manufacturers opportunity for solutions optimized for cost, speed and printed picture quality.

Ink Jet Printer. Consumer grade ink-jet printers of today range from 300 to 1200 dots per inch (e.g., those from Lexmark International, Hewlett Packard, Epson, and Canon), with at least one available at 1440 dpi (e.g., the Epson Stylus Color 800). Offering an effective balance of performance and cost, ink jet printers are limited by printing time and artifacts. The time to physically scan a pen cartridge across the page, and then up the page, fundamentally limits printing times. Artifacts include horizontal banding due to paper transport inaccuracies, and graininess of the printed image due to low resolution or the use of dithering and halftone patterns to create shades of gray. Proper color rendition, including proper tint, saturation, and contrast is controlled by (1) the ink, (2) the control of the ink jet, and (3) the type of paper. The bleeding characteristics of ink varies with the paper fibers, which in effect can change the contrast of the image. These printers make adequate printed images for desktop publishing with text, graphics, and images.

Color Laser Printer. Electrostatic color laser printers consist of a print engine and a print controller. The printer's engine determines the base resolution capabilities, the paper capacities in terms of pages per minute, and type of consumables used. These printers have been manufactured at resolutions of 300 to 1000 dpi for black and white, and color laser printers at 300 to 1200 dpi (Hewlett Packard, IBM, and Canon).

The engine contains a laser diode for imaging onto an organic photoconductor (*OPC*), a fuser, toner, and paper transport. The low-power laser diode, driven by the controller, first writes the latent image (or pattern of charge) onto the OPC. Four layers of charged powder toner (cyan, magenta, yellow, black) are rolled past the OPC drum, attracting to the oppositely charged OPC image. As the drum rotates, the toner is transferred to the paper from the OPC (in the pattern of the latent image). The fuser's application of heat and pressure with rollers bond the toner to the page. A fuser oil keeps the toner off the roller and onto the paper.

The printer's controller will directly influence the overall print quality. Considering the limitations of the print engine, the print controller will translate target pixels into a printable image. For example, the engine may only produce 300 dpi at 16 colors. Considering that limitation, the print controller may perform dithering or halftoning to increase the *apparent* resolution of the printed page, both spatially and chromatically. The controller's speed of performing image processing functions like enhancement or this dithering can limit throughput. Fast processors provide acceleration for rapid printing in the midst of many image processing functions. Some printers can also modulate the laser beam intensity to control the size and placement of the laser dots, effectively controlling resolution. In addition *apparent* resolution can be improved when incorporating image processing techniques that smooth edges or curves.

Professional and Offset Printing. High-end, professional, or production printing units utilize many techniques, including electrostatic and ink jet, to transfer an image onto paper or other material. Powerful image processing capabilities of these machines render the large high-resolution images and drive the printing devices, which may include the use of multiple drums or multiple passes, for optimal print quality. With large images and many subsequent passes of the image dictates these printers must use large memories and/or compression techniques for temporary storage of images.

Digital Copiers and Multifunction Systems (*MFS*). A multifunction system (*MFS*) combines digital scanners, image processing, printers, and digital copiers to create a new class of high-performance page copiers. This coupling of imaging subsystems with image processing offers new features. Conventional copiers simply scanned and printed line by line, whereas these newer forms of digital copiers offer additional features like “Scan-Once-Followed-by-Print-Many-Copies” modes, since they use full compression of the document, adaptive processing of images for improved print quality, and wider ranges of document resizing with affine transforms. In addition these *MFS* equipment can incorporate document image segmentation and character recognition for highly compressed storage of the document, subsequent editing of the document, and later high-quality printing of the document.

Color Management. To achieve reliable and reproducible color throughout the entire reproduction process, independent of platform or device, major providers like Adobe, AGFA, SGI, Kodak, and Microsoft developed the International Color Consortium (*ICC*) specification. Scanners use reference images (both scanned and previously stored reference images) to aid in determining the appropriate corrections. Likewise printers are measured with reference images. Using this standard, ICC profiles for the input scanner and output printer are tagged with the image file to facilitate proper correction at the time of use, mapping from any color space to any other color space. Consistent color rendition is achieved with these techniques.

BIBLIOGRAPHY

1. C. Koch Neuromorphic Vision Chips, *IEEE Spectrum*, **33** (5): 38–46, 1996.
2. W. J. Mitchell When is seeing believing?, *Sci. Am., Special Issue Comput. 21st Century*, **6** (1): 110–119, 1995.
3. C. Sequin M. Tompsett *Charge Transfer Devices*, New York: Academic Press, 1975.
4. J. Younse R. Gove Texas Instruments, Inc., *Programmable CCD Defect Compensator*, U.S. Patent 4,805,023, 2/14/89.
5. G. C. Holst *CCD Arrays, Cameras, and Displays*, Bellingham, WA: SPIE Optical Engineering Press, 1996.
6. G. Wolberg *Digital Image Warping*, Los Alamitos, CA: IEEE Computer Society Press, 1990.
7. W. Wolfe G. Zissis *The Infrared Handbook*, Washington, DC: Environmental Research Institute of Michigan, 1978.
8. B. G. Haskell A. Puri A. N. Netravali *Digital Video: An Introduction to MPEG-2*, New York: Chapman & Hall, 1997.

9. V. Bhaskaran K. Konstantinides *Image and Video Compression Standards: Algorithms and Architectures*, Norwell, MA: Kluwer Academic, 1995.
10. R. Gove *et al.* Next-generation media processors and their impact on medical imaging, *SPIE Proc. Med. Imaging*, 1998.
11. C. Basoglu *A Generalized Programmable System and Efficient Algorithms for Ultrasound Backend*, PhD dissertation, Univ. Washington, Seattle, 1997.
12. Y.-M. Chou H.-M. Hang A video coding algorithm based on image warping and nonrectangular DCT coding, *Proc. SPIE, Visual Commun. Image Proc. '97*, San Jose, 1997, pp. 176–187.
13. C. Laurent C. Boeville Parallel implementation of greyscale morphological operators on MVP, *Proc. SPIE, Visual Commun. Image Process. '97*, San Jose, 1997, pp. 502–513.
14. R. J. Gove Architectures for single-chip image computing, *SPIE Proc. Electron. Imaging Sci. Technol. Conf. Image Process. Interchange*, San Jose, CA, 1992, pp. 30–40.
15. K. Guttag R. Gove J. Van Aken A single-chip multiprocessor for multimedia: The MVP, *IEEE Comput. Graphics Appl.*, **12** (6): 53–64, 1992.
16. L. McMillan L. Westover A forward-mapping realiation of the inverse discrete cosine transform, *Proc. IEEE Data Compression Conf.*, Snowbird, UT, 1992, pp. 219–228.
17. H. Malvar M. Murphy, (eds.) *The Science of Videoconferencing 2 (Inside Picture Tel)*, Andover, MA: PictureTel Corporation, 1996.
18. J. Duran C. Sauer *Mainstream Videoconferencing: A Developer's Guide to Distance Multimedia*, Reading, MA: Addison-Wesley, 1997.
19. B. Erol *et al.* The H.263+ video coding standard: Complexity and performance, *Proc. IEEE Data Compression Conf.*, Snowbird, UT, 1998, pp. 259–268.
20. A. Jacquin H. Okada P. Crouch Content-adaptive postfiltering for very low bit rate video, *Proc. IEEE Data Compression Conf.*, Snowbird, UT, 1997, pp. 111–120.
21. J. Goel Digital video resizing, *Electron. Design News*, December 18, pp. 38–39, 1995.
22. A. Murat Tekalp *Digital Video Processing*, Upper Saddle River, NJ: Prentice-Hall, 1995.
23. W. K. Pratt *Digital Image Processing*, 2nd ed., New York: Wiley, 1991.
24. S. Khoshafian A. Brad Baker *Multimedia and Imaging Databases*, San Francisco: Morgan Kaufmann, 1996.
25. D. Kozen Y. Minsky B. Smith Efficient algorithms for optimal video transmission, *Proc. IEEE Data Compression Conf.*, Snowbird, UT, 1998, pp. 229–238.
26. R. Weiss A. Duda D. Gifford Content-based access to algebraic video, *IEEE Proc. Int. Conf. Multimedia Comput. Syst.*, 1994, pp. 140–151.
27. Y. Gong *et al.* An image database system with content capturing and fat image indexing abilities, *IEEE Proc. Int. Conf. Multimedia Comput. Syst.*, 1994, pp. 140–151.
28. A. F. Inglis A. C. Luther *Video Engineering*, 2nd ed., New York: McGraw-Hill, 1996.
29. A. Sobel Television's bright new technology, *Sci. Amer.*, **278** (5): 45–55, 1998.
30. L. E. Tannas, ed. *Flat-Panel Displays and CRTs*, Jr. New York: Van Nostrand Reinhold, 1985.
31. E. Kaneko *Liquid Crystal TV Displays: Principles and Applications of Liquid Crystal Displays*, Norwell, MA: Reidel Publishing, 1987.
32. W. E. Howard Thin-film transistor/liquid crystal display technology: An introduction, *IBM J. Res. Develop.*, **36** (1): 3–10, 1992.
33. P. Wam Plastics Get Wired, *Sci. Amer., Special Issue on Solid State Century, the Past, Present and Future of the Transistor*, 90–96, January 22, 1998.
34. R. Gove DMD display systems: The impact of an all-digital display, *Proc. Soc. Inf. Display Int. Symp.*, San Jose, CA, 1994.
35. L. Pearlstein J. Henderson J. Boyce An SDTV decoder with HDTV capability: An all format ATV decoder, *137th SMPIE Proc.*, 1995, pp. 422–434.
36. Adobe Systems Inc., *Postscript Language Reference Manual*, Reading, MA: Addison-Wesley, 1990.
37. C. Poynton *Technical Introduction to Digital Video*, New York: Wiley, 1996.
38. K. R. Castleman *Digital Image Processing*, Upper Saddle River, NJ: Prentice-Hall, 1996.

ROBERT JOHN GOVE
Equator Technologies, Inc.