APP NOTE

No. AN9758 August 1997 Harris Multimedia

Understanding Multimedia Standards

Author: Mark E. Hazen

Introduction

This application note is intended for those who need to gain a fundamental understanding of multimedia related topics in a short period of time. Today, “Multimedia” embraces a wide range of topics that include, among other things, international standards, digital audio and video processing, PC interfacing, analog video theory, and a seemingly endless stream of terminology and acronyms. For the uninitiated, this mountain of information can be overwhelming. Where does one begin? This application note is one place to start. Here, you will establish a foundation that will help you more clearly understand multimedia applications such as video conferencing, video capture and processing, and other related topics.

Major topics covered in the application note are:

1. The Need for Standardization
2. Worldwide Standards Organizations and Interaction
3. Structure and Standards of the ITU
4. Image and Sound Compression Standards
5. Component and Composite Video Standards
6. Color Space Standards
7. Pixels and Screens
8. YCbCr Digital Sampling
9. Computer Bus Standards
10. What is MMX™ Technology?

You will discover that these topics are generously supported with Internet addresses that will enable you to increase your depth and breadth of learning.

The Need for Standardization

STANDARDIZATION - A NECESSITY!

Standardization is important. It always has been. Even in ancient times, units of measure had to be established and held as the standard. The problem is, the standards established in one region of the world were not the same as those used by others. That problem still exists today. Some use gallons, others liters. Some use feet, others meters. Some use English, others use metric. Certainly these differences have added frustration but not necessarily impossibility.

But things have changed. We are now living in the age of advanced digital communications that is well beyond the simple analog communications techniques of the past. For the sake of argument, there are thousands of possible ways in which digital information can be packaged, processed, and presented. If the use of digital communications, which includes digital multimedia, were simply a local concern, any one of the possible means of packaging, processing, and presenting would probably be adequate, dictated only by local need. However, we are no longer a local or tribal people. We are now attempting to become one global community. As a result, clearly defined digital standards, applied universally, are absolutely essential for unhindered communications using digital techniques. This applies to all forms of digital communications including cellular phones, the Internet, wireless local-area and wide-area networks, and digital multimedia.

INDUSTRY STANDARDS ARE NOT ALWAYS INTERNATIONAL!

Industrialized countries of the world have national technical organizations that serve to facilitate standardization for that country or region. In the United Stated for example, their are standards-forming groups such as the Electronics Industries Association (EIA) and the Institute for Electrical and Electronics Engineers (IEEE). Large industry organizations such as these not only serve to help define national standards but, also, serve as representatives internationally. Many IEEE Standards, for example have become international standards and vice versa. On the other hand, some national industry standards never become international standards.

In addition to large organizations, such as the IEEE, that represent nearly all of the electronics industries, there are many smaller consortiums that are frequently formed between companies interested in the same thing. These consortiums may be composed of three or four companies or, perhaps, dozens of companies who have common interest in a particular technical challenge. The main reason these consortiums form is to accelerate the product development process to be first-to-market with a solution. Another reason is consortiums often receive government funding to assist with the research and development phase of the new product development. Standards established by these consortiums are by no means international in nature. In fact, many of the standards are considered to be proprietary. That is a tactful way of saying the standards are unique and offer restrictive and local solutions, especially when they apply to any form of digital communications.

PROS AND CONS OF PROPRIETARY STANDARDS

Very often, individual companies, or small groups of companies form alliances, similar to consortiums, to aid each other in creating a solution to a potential market need. As was briefly mentioned above, the resulting solution is proprietary in nature and usually has little or nothing to do with international standards. That's because, in most cases, international standards regarding the particular market opportunity do not yet exist and are slow in coming. One example of this is the video/voice conferencing software solutions offered by many different companies such as Internet Phone™ from VocolTec, WebPhone™ from NetSpeak™, and CoolTalk from InSoft™. Companies with these proprietary solutions and standards wanted to be first-to-market to reap the initial financial harvest provided by eager small-businesses and consumers. Thus, proprietary standards offer a fast initial financial gain and satisfy, at least to some extent, early customers.

But, there are many problems involved with proprietary solutions. First of all, proprietary solutions are usually never compatible with each other. For example, Internet Phone™ (VocolTec) users cannot communicate over the Internet with CoolTalk (InSoft™) users. Another problem with proprietary solutions, especially in the digital world, is that the proprietary solution is almost sure to be replaced with an international solution. Consumers that start out using a proprietary solution will almost certainly change to the more universal international solution once it is established. Thus, manufacturers who invest in their own proprietary solution realize that, in many cases, it will only yield a short-term return. On the other hand, manufacturers who wait for the international standards lose out on the initial financial gain and face a market of strong competition from other manufacturers who are entering the market with internationally-standardized products. It should also be recognized here that manufacturers with proprietary solutions often modify their products to conform to international standards as the market demands. Some, or all, of the products mentioned above are, or have been, redone to conform to international standards for teleconferencing (video/voice/data conferencing).

Harris Semiconductor has taken the long-range position based on international multimedia standards. The Harris strategy is to offer the market a combination of sophisticated, highly-integrated, multimedia building blocks along with strong applications support to meet the requirements of international standards.

THE STRENGTH OF INTERNATIONAL STANDARDS

International standards offer long-range product usability along with compatibility and interoperability. These are the product characteristics that the market really demands - not short-term proprietary solutions, but robust long-term solutions that serve as a solid international infrastructure.

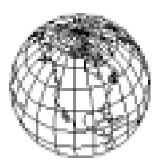
So, how does the world define and establish international standards? It's usually a very long and complicated process that requires the cooperation of strong influences representing highly industrialized nations. In regard to international standards for analog and digital communications, which includes multimedia, international influences work together through the International Electrotechnical Commission (IEC) and the International Telecommunications Union (ITU), an arm of the United Nations. In this next section, we will take time to identify and define many of the world organizations that work toward international standardization.

Worldwide Standards Organizations and Interaction

INTERNATIONAL TELECOMMUNICATIONS UNION (ITU)

The International Telecommunications Union (ITU) became an arm of the United Nations, headquartered in Geneva, Switzerland in 1947. Its origin dates back to 1865 when it was known as the International Telegraph Union in Paris. Today, the mission of the ITU is to coordinate development, regulation, and standardization of telecommunications between government and private sectors of the world. National and Industry standards organizations of the world work through the ITU to establish global standards that ensure worldwide compatibility and interoperability of products and regulation of radio-frequency transmissions. The ITU deals with issues relating to all forms of telecommunications, i.e., radio-frequency communications, telephone system-based communications, computer networks, and multimedia. Details regarding the ITU will be covered in the next main section. You can learn much more about the ITU on the Web at: http://www.itu.ch/

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO)



"The International Organization for Standardization (ISO) is a worldwide federation of national standards bodies from some 100 countries, one from each country. ISO is a non-governmental organization established in 1947. The mission of ISO is to promote the development of standardization and related activities in the world with a view to facilitating the international exchange of goods and services, and to developing cooperation in the spheres of intellectual, scientific, technological and economic activity. ISO's work results in international agreements which are published as International Standards.

Many people will have noticed a seeming lack of correspondence between the official title when used in full, International Organization for Standardization, and the short form, ISO. Shouldn't the acronym be "IOS? Yes, if it were an acronym - which it is not. In fact, "ISO" is a word, derived from the Greek isos, meaning "equal"., which is the root of the prefix "iso-" which occurs in a host of terms, such as "isometric" (of equal measure or dimensions - Shorter Oxford English Dictionary) and "isonomy" (equality of laws, or of people before the law - ibid). From "equal" to "standard", the line of thinking that led to the choice of "ISO" as the name of the organization is easy to follow.

NetSpeak™ and WebPhone™ are trademarks of NetSpeak Inc.
InSoft™ is a trademark of InSoft Inc.
Internet Phone™ is a trademark of VocolTec Inc.

The scope of ISO is not limited to any particular branch; it covers all standardization fields except electrical and electronic engineering, which is the responsibility of the IEC and ITU. The work in the field of information technology is carried out by a joint ISO/IEC technical committee (JTC 1)." - Quoted from: http://www.iso.ch/infoe/intro.html

INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC)

"The mission of the International Electrotechnical Commission (IEC) is to promote, through its members, international cooperation on all questions of standardization and related matters, such as the assessment of conformity to standards, in the fields of electricity, electronics and related technologies. It therefore, provides a forum for the preparation and implementation of consensus-based voluntary international standards, facilitating international trade in its field and helping to meet expectations for an improved quality of life." - Quoted from http://www.iec.ch/pig001-e.htm

The IEC is based in Geneva, Switzerland and works closely with the Comité Europeen de Normalisation Electrotechnique (CENELEC) and the ITU. The U.S. Member of IEC is the United States National Committee (USNC). The USNC is housed at the American National Standards Institute based in New York. Work groups and advisory groups of the IEC work closely in parallel with work groups and advisory groups of the ITU to develop international standards.

COMITÉ EUROPEEN DE NORMALISATION ELECTROTECHNIQUE (CENELEC)

The European Committee for Electrotechnical Standardization (CENELEC) was formed in 1972 and is responsible for electrical and electronic standards in the European Union. CENELEC works closely in cooperation with the IEC and the ITU. CENELEC's mission is to develop and approve European standards in the electrotechnical field. You can learn much more about CENELEC on the Web at: http://www.sis.se/engelska/2neng.htm

COMITÉ EUROPÉEN DE NORMALISATION (CEN)

The European Committee for Standardization (CEN) is responsible for European standardization in all fields except Electrotechnical (CENELEC) and Telecommunications (ETSI). CEN, CENELEC, and ETSI together are the European Union equivalent of the American National Standards Institute (ANSI). CEN, CENELEC, ETSI and ANSI are members of ISO and IEC. You can learn much more about CEN on the Web at: http://tobbi.iti.is/cen/

EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE (ETSI)

The European Telecommunications Standards Institute is the youngest of the three European standards making bodies (CEN, CENELEC, ETSI). ETSI was recognized by the European Council of Ministers by Council Directive 83/189. ETSI was established in 1988 to set standards for Europe in telecommunications in cooperation with the European Broadcasting Union (EBU), CEN, and CENELEC. Learn much more about ETSI on the Web at: http://www.etsi.fr/

ETSI, CEN, and CENELEC together are the European equivalent of the American National Standards Institute (ANSI). These agencies, among many others, work with the IEC and the ITU in establishing international telecommunications standards.

AMERICAN NATIONAL STANDARDS INSTITUTE (ANSI)

"The American National Standards Institute (ANSI) has served in its capacity as administrator and coordinator of the United States private sector voluntary standardization system for 78 years. Founded in 1918 by five engineering societies and three government agencies, the Institute remains a private, nonprofit membership organization supported by a diverse constituency of private and public sector organizations...

The Institute represents the interests of its nearly 1,400 company, organization, government agency, institutional and international members through its headquarters in New York City, and its satellite office in Washington, D.C.

ANSI does not itself develop American National Standards (ANSs); rather it facilitates development by establishing consensus among qualified groups...

ANSI is the sole U.S. Representative and dues-paying member of the two major non-treaty international standards organizations, the International Organization for Standardization (ISO), and, via the U.S. National Committee (USNC), the International Electrotechnical Commission (IEC).

ANSI was a founding member of ISO and plays an active role in its governance. ANSI is one of five permanent members to the governing ISO Council, and one of four permanent members of ISO's Technical Management Board. U.S. Participation, through the U.S. National Committee, is equally strong in the IEC. The USNC is one of 12 members on the IEC's governing Committee of Action and the current president of the IEC is from the United States." - Quoted from http://www.ansi.org/whatansi.html

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS (IEEE)

"The Institute of Electrical and Electronics Engineers (IEEE) is the world's largest technical professional society. Founded in 1884 by a handful of practitioners of the new electrical engineering discipline, today's Institute is comprised of more than 320,000 members who conduct and participate in its activities in 147 countries. The men and women of the IEEE are the technical and scientific professionals making the revolutionary engineering advances which are reshaping our world today." - Quoted from http://www.ieee.org/i3e_blb.html

The IEEE is a member of ANSI which serves as a representative to the IEC and ISO through the U. S. National Committee (USNC).

SOCIETY OF MOTION PICTURE AND TELEVISION ENGINEERS (SMPTE)

"The Society of Motion Picture and Television Engineers (SMPTE) is an international, award-winning technical society devoted to advancing the theory and application of motion-imaging technology including film, television, video, computer imaging, and telecommunications. The SMPTE currently serves 8,500 members in 72 countries. Members of the Society are engineers, executives, technical directors, cameramen, editors, consultants, and specialists in film processing, film and television production and post-production, and practitioners from almost every other discipline in the motion-imaging industry...

The Society was founded in 1916, as the Society of Motion Picture Engineers. The T was added in 1950 to embrace the emerging television industry. The SMPTE is recognized around the globe as a leader in the development of standards and authoritative, consensus-based recommended practices (RPs) and engineering guidelines (EGs). The Society serves all branches motion imaging including film, video, and multimedia." - Quoted from http://www.smpte.org/../../society.html

The SMPTE is an accredited ANSI Standards Developing Organization, and is recognized by ISO and IEC.

VIDEO ELECTRONICS STANDARDS ASSOCIATION (VESA)

The Video Electronics Standards Association is headquartered in San Jose, CA. Formed in 1989, VESA is an international nonprofit corporation led by a Board of Directors which represents a voting membership of more than 290 corporate members worldwide. It is an organization that supports and sets industry-wide interface standards for the PC, workstation, and computing environments. VESA promotes and develops open standards for the video electronics industry, ensuring interoperability and encouraging innovation and market growth. VESA developed the EISA bus and VL-Bus for PCs. Learn much more about VESA on the Web at: http://www.vesa.org/

Structure and Standards of the ITU

ITU STRUCTURE

The International Telecommunications Union (ITU) is an intergovernmental agency. That means it is comprised of governments, or nations, also referred to as states. All governing bodies associated with the ITU are made up of states and, as such, are referred to as State Members. In addition to the primary governmental membership, there are so-called Sector Members. Sector Members are non-governing who serve in advisory capacities. Sector members are companies, corporations, broadcasters, public and private operators, and regional/international organizations who have an interest in international telecommunications standards. As of February 22, 1996, the ITU had 185 State Members and 363 Sector Members.

As illustrated in Figure 1, the ITU is headed by the Plenipotentiary Conference under which a council serves to coordinate World Conferences on International Telecommunications and identify worldwide regulatory issues. The General Secretariat of the ITU is responsible for membership servicing and assisting the Council in conference coordination. The mission of the ITU is carried out through the efforts of three primary sectors:

- **Radio Communication Sector is comprised of:**
  - **World and Regional Radiocommunication Conferences**
  - **Radiocommunications Assemblies**
  - **Radio Regulations Board**
  - **Telecommunication Standardization Sector (Including World Telecommunication Standardization Conferences)**
  - **Telecommunications Development Sector (Includes World and Regional Telecommunication Development Conferences)**

The Radiocommunication Sector came out of the former International Consultative Radio Committee (CCIR). Its overall purpose is to regulate the use of the radio spectrum for all services worldwide. Areas of concern include:

- **Spectrum Utilization and Monitoring**
- **Inter-Service Sharing and Compatibility**
- **Science Services**
- **Radio Wave Propagation**
- **Fixed-Satellite Service**
- **Fixed Service**
- **Mobile Services**
- **Sound Broadcasting**
- **Television Broadcasting**

Regulatory standards established by the Radiocommunication Sector are prefaced with an ITU-R designator.

The Telecommunication Standardization Sector was formed out of the former International Consultative Telegraph and Telephone Committee (CCITT). The responsibilities of the Telecommunication Standardization Sector involve the study and standardization of the interconnection of radio systems in public telecommunications networks in order to ensure worldwide interoperability. To address a particular issue, a study group is formed of experts from public and private sectors. Members of study groups are Sector Members and include corporations and national standards committees. Areas of concern include:

- **Telecommunication Services and Network Operation**
- **Telecommunication Tariffs and Accounting Principles**
- **Maintenance**
- **Protection of Outside Plant**
- **Data Communication**
- **Terminal for Telematic Services**
- **Switching, Signaling and Man-Machine Language**
- **Transmission Performance, Systems and Equipment**
- **ISDN**

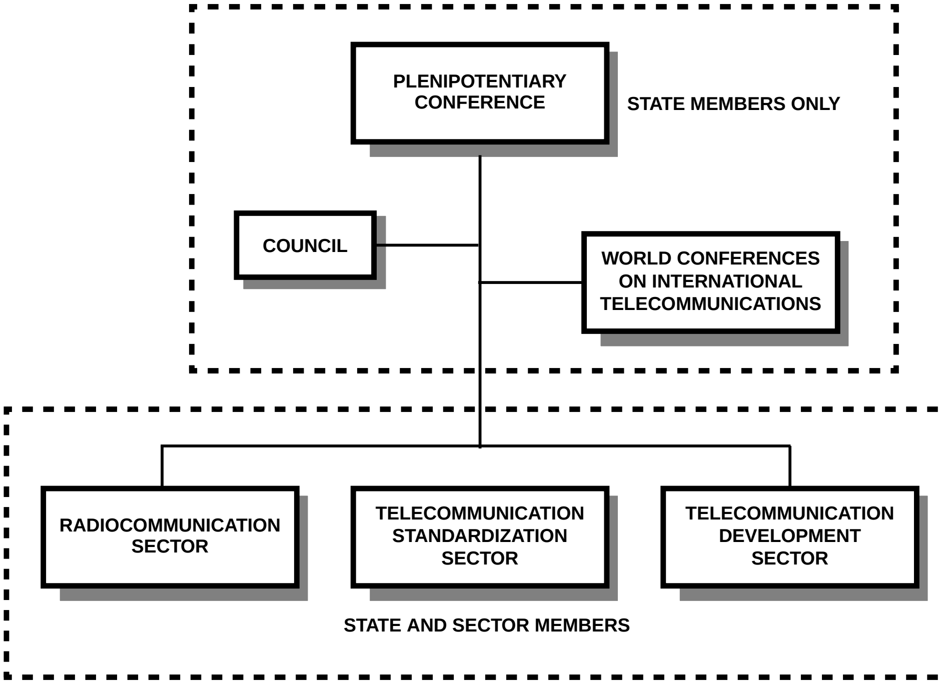


FIGURE 1. STRUCTURE OF THE ITU

Standards established by the ITU are recommended which means their adoption is voluntary and optional. However, most recognize the wisdom of adopting these international standards. Standards recommended by the Telecommunication Standardization Sector are prefaced with an ITU-T designator.

The Telecommunications Development Sector was established to address the development of communications capability throughout the world with particular attention to developing nations. This sector administers information and financial assistance on behalf of the United Nations. Recommendations of the Telecommunications Development Sector are prefaced with an ITU-D designator.

ITU-R AND ITU-T SERIES DESIGNATORS

ITU Standards are organized by "Series". A series and has an alphabetic designator. For example, ITU-T H.320 is a recommended standard of the Telecommunication Standardization Sector and "H" is the series designator. Listed here are the series for ITU-R and ITU-T. The number in parenthesis is the number of recommendations in that particular series.

ITU-R Series Recommendations

BO	Broadcasting satellite service (20)
BR	Sound and television recording (29)
BS	Broadcasting service (sound) (46)
BT	Broadcasting service (television) (63)
F	Fixed service (123)
IS	Inter-service sharing and compatibility (10)
M	Mobile, radiodetermination, amateur and related satellite services (124)
P	Radiowave Propagation (67)
RA	Radioastronomy (6)
S	Fixed satellite service (57)
SA	Space applications and meteorology (45)
SF	Frequency sharing between the fixed satellite service and the fixed service (16)
SM	Spectrum management (41)
SNG	Satellite news gathering (7)
TF	Time signals and frequency standards emissions (21)
V	Vocabulary and related subjects (12)

ITU-T Series Recommendations

- A Organization of the work of the ITU-T (8)
- B Means of expression (definitions, symbols, classification) (13)
- C General telecommunication statistics (4)
- D General tariff principles (133)
- E Overall network operation, telephone service, service operation and human factors (244)
- F Telecommunication services other than telephone (136)
- G Transmission systems and media, digital systems and networks (4)
- H Line transmission of non-telephone signals (45)
- I Integrated Services Digital Networks (ISDN) (183)
- J Transmission of sound program and television signals (65)
- K Protection against interference (33)
- L Construction, installation and protection of cable and other elements of outside plant (19)
- M Maintenance: transmission systems, telephone circuits, telegraphy, facsimile, etc. (172)
- N Maintenance: international sound program and television transmission circuits (38)
- O Specifications of measuring equipment (35)
- P Telephone transmission quality, telephone installations, local line networks (74)
- Q Switching and Signaling (4)
- R Telegraph transmission (73)
- S Telegraph services terminal equipment (33)
- T Terminal characteristics and higher layer protocols for telematic systems (155)
- U Telegraph switching (51)
- V Data communication over the telephone network (72)
- X Data networks and open system communication (7)
- Z Programming languages (35)

Standards that pertain to teleconferencing topics (data conferencing, video conferencing, etc.) that you will see associated with Harris products and related applications most often are of series BT, H, G, and T. Standards may also have an annex attached such as H.261 Annex D. An annex is simply an appendix to the standard which provides further detail or explanation. More detailed information about the ITU is readily available on the Internet at http://www.itu.int/aboutitu/.

Image and Sound Compression Standards

JOINT PHOTOGRAPHIC EXPERTS GROUP (JPEG)

JPEG, pronounced “jay-peg”, is a picture (image) compression standard whose name comes from the expert study group that created the standard, the Joint Photographic Experts Group. It was designed for the compression of full-color and gray-scale photographic-type images and is intended for the compression of still images of a photographic origin as opposed to text and line drawings.

This compression standard renders images that are somewhat lacking in sharpness due to the very high degree of compression that is possible. This is because some image information is lost in the process. For intended applications such as Internet images, the resulting image quality is acceptable. The advantage gained by giving up some picture sharpness is the compactness of the picture data file and the speed of its transfer over a network. There are so-called “lossless” compression algorithms but they do not offer the large degree of compression offered by JPEG. Also, with JPEG, the degree of lost image information can be varied by setting compression parameters trading off picture quality for file size and transmission speed.

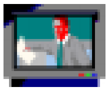
There are various versions of so-called M-JPEG which is motion JPEG or moving-picture JPEG. M-JPEG is simply a series of JPEG pictures each serving as frames that are displayed rapidly to create a motion picture. However, M-JPEG is not an international standard. Instead versions of M-JPEG have appeared as proprietary standards. The international MPEG standard is more efficient and produces better image results than the proprietary M-JPEG standards. Nevertheless, M-JPEG is used in some professional editing suites because it allows for frame-by-frame editing which MPEG does not. A JPEG tutorial is available on the Internet at: http://www.cis.ohio-state.edu/hypertext/faq/usenet/jpeg-faq/part1/faq.html

MOVING PICTURE EXPERTS GROUP (MPEG)

MPEG, pronounced “em-peg”, is an acronym for the Moving Picture Experts Group which has developed, and is developing, a number of moving picture and audio compression standards identified by the same acronym. The Moving Picture Experts Group (MPEG) was established in 1988 and is a working group that operates in the framework of the Joint ISO/IEC Technical Committee (JTC 1) on Information Technology. Their mission is the development of international standards for compression, decompression, processing, and coded representation of moving pictures, audio and their combination. ISO/IEC working groups collaborate with ITU counterpart working groups to define and recommend international standards. The working group has produced, or is working on, the following MPEG Standards:

MPEG-1 is a standard for the storage and retrieval of moving pictures and associated audio (up to two channels for stereo audio) for digital storage media at up to about 1.5 Mbps. This allows for digital storage of video and audio on compact disks with reproduction quality comparable to VHS video tape.

MPEG-2 is a standard primarily for digital television, including HDTV. It is based on MPEG-1 but provides better picture quality, higher data transfer rates (up to 40 Mbps), five-channel surround-sound audio, and copyright protection and identification.



MPEG-3 was intended specifically for HDTV but it was found that MPEG-2 was fully capable of handling HDTV. Thus, MPEG-3 is no longer mentioned.

MPEG-4 is targeted toward developing universal, efficient coding of different forms of audiovisual data, called audiovisual objects. This will involve a set of coding tools and a syntactic description language to describe the coded representation of these, audiovisual objects. MPEG-4 will not be completed until late '98 or early '99 and will be based on portions of MPEG-1 and MPEG-2, along with some ITU-T Standards such as G.723, H.261 and H.263.

MPEG-7 is a content representation standard for information search. It is also titled Multimedia Content Description Interface. It will define the manner in which audiovisual materials can be coded and classified so the materials can be easily located using search engines just as search engines are used to locate text-based information. Music, art, line drawings, photos, and videos are examples of the kinds of materials that will become searchable based on descriptive language defined by MPEG-7. This standard is expected to be completed in the year 2000.

The Moving Picture Experts Group holds three or more meetings per year. These comprise plenary meetings and subgroup meetings covering audio and video technical requirements and are attended by some 300 experts from 20 countries. To learn more about MPEG, the following Internet site is highly recommended: http://www.cselt.stet.it/mpeg/

ITU VIDEO CONFERENCING STANDARDS H.320, H.323, H.324

International standards are absolutely essential to ensure worldwide connectivity and interoperability of hardware and software used for various forms of conferencing over various mediums such as the PSTN, ISDN, and LANs which includes intranets and the Internet. The ITU has established, and is establishing, standards that govern audio conferencing, video conferencing, and data conferencing. Here, you will be introduced to some of those standards. They are the primary standards upon which many Harris Semiconductor multimedia products are based. ITU-T H.320, H.323, and H.324 are three primary, so-called, "umbrella" conferencing standards to which you will be introduced. Each of these standards are supported with specific video, audio, data, and control standards. Now, let's begin by examining the H.320 umbrella standard.

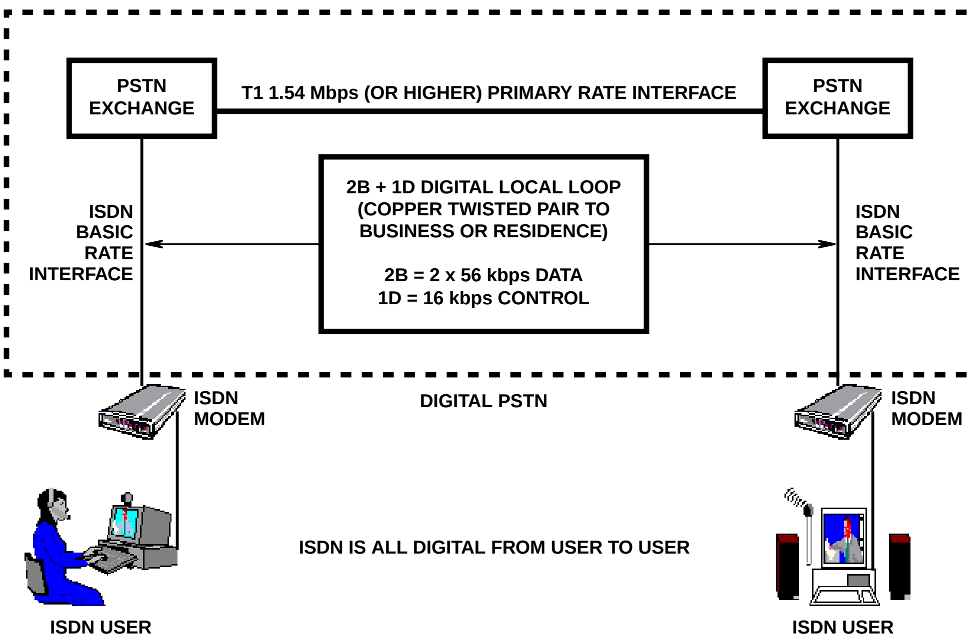


FIGURE 2. H.320 FOR ISDN VIDEO CONFERENCING

H.320 - (ISDN) NARROWBAND VISUAL TELEPHONE SYSTEMS AND TERMINAL EQUIPMENT

H.320 is actually a series, or group, of various ITU-T Standards that pertain to audio, video, and graphics communications over ISDN or Switched 56 lines. H.320 is often referred to as H.320 ISDN Conferencing Standards and is illustrated in Figure 2. The standards specify requirements for digital audio and video processing, formatting, and protocols that ensure worldwide compatibility and interoperability of software and hardware. ITU-T Standards that fall under, or are used to support, H.320 are as follows:

H.221 - Frame structure for a 64 to 1920 kbps channel in audiovisual teleservices.

H.242/H.243 - Multipoint Conferencing Standards used to coordinate capabilities of individual participants that are taking part in a multipoint conference. These standards control interaction between individual conferencing stations and a Multipoint Control Unit (MCU).

H.261 - Video Compression/Decompression Standard for ISDN Line Communications.

G.711 - Audio Compression/Decompression using an entire ISDN B channel (56 kbps) leaving a B channel for the video.

G.722 - 7kHz audio-coding within 64 kbps.

G.728 - Audio - Compression/Decompression using only 16 kbps of the entire 128 kbps ISDN channel. Thus, G.728 is preferred over G.711.

T.120 - Data Conferencing Standard that enables data conferencing over various mediums such as ISDN, PSTN, and LANs. The T.120 Standard is actually a set of T.12X Standards relating to data conferencing.

H.323 - (LAN) VISUAL TELEPHONE TERMINALS OVER NON-GUARANTEED QUALITY OF SERVICE LANS

(Audiovisual communication on LANs, and WANs, including intranets and the Internet.) Like H.320, H.323 is a series of audio, video, and data standards that are used to govern LAN-based desktop conferencing. H.323 video conferencing is illustrated in Figure 3. Standards that fall under H.323 are:

H.225 - Packetization and Synchronization of Data.

H.245 - Data Flow Control.

H.261 - Video Compression/Decompression Standard for ISDN Line Communications.

H.263 - Video Compression/Decompression Standard for low bit-rate communications such as would be required when using the PSTN to connect to the Internet.

G.711 - Audio Compression/Decompression and 56 kbps data rate.

G.722 - 7kHz audio-coding within 64 kbps.

G.723.1 - Dual rate speech coder for multimedia communications transmitting at 5.3 kbps and 6.3 kbps.

G.728 - Audio Compression/decompression and 16 kbps data rate.

G.729 - Coding of speech at 8 kbps using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP).

T.120 - Data Conferencing Standard that enables data conferencing over various mediums such as ISDN, PSTN, and LANs. The T.120 standard is actually a set of T.12X Standards relating to data conferencing.

H.324 - (PSTN) TERMINAL FOR LOW BIT-RATE MULTIMEDIA COMMUNICATION

H.324 is an umbrella standard that pertains to conferencing over the PSTN as shown in Figure 4. Sometimes this standard gets confused with H.323 because, in many cases, we use the PSTN to connect to the Internet. The H.324 Standard applies specifically to conferencing sessions that take place using the regular phone system, not the Internet. An H.324 user will typically make a normal phone call to the other person at which time they both agree to switch to the video conferencing mode. The mode is initiated by allowing the modem on each end to connect with the other. Then, voice, video, and data are exchanged via the modems under software control. H.324 can be implemented using dedicated video phones, set-top-box video phones using the TV, and PC-based video phones. Like other conferencing standards, H.324 is supported by the following function-specific standards:

H.223 - Multiplexing Protocol for low bit rate Multimedia Communication.

H.245 - Data flow control.

H.263 - Video Compression/Decompression Standard for low bit-rate communications such as would be required when using the PSTN.

G.723 - Dual rate speech coder for multimedia communications transmitting at 5.3 kbps and 6.3 kbps

T.120 - Data Conferencing Standard that enables data conferencing over various mediums such as ISDN, PSTN, and LANs. The T.120 Standard is actually a set of T.12X Standards relating to data conferencing.

V.34 - A modem operating at data signalling rates of up to 33.6 kbps for use on the general switched telephone network and on leased point-to-point 2-wire telephone-type circuits.

V.80 - In-band DCE (data communications equipment) control and synchronous data modes for asynchronous DTE (Data Terminal Equipment).

Hopefully, now you understand what ITU Standards are and why they are needed. But, there is more. There are other kinds of standards that are just as important as those about which you have already learned. As we continue from here, you will be introduced to standards that seem a little more technical. We will keep it simple so you are able to capture the important basic concepts.

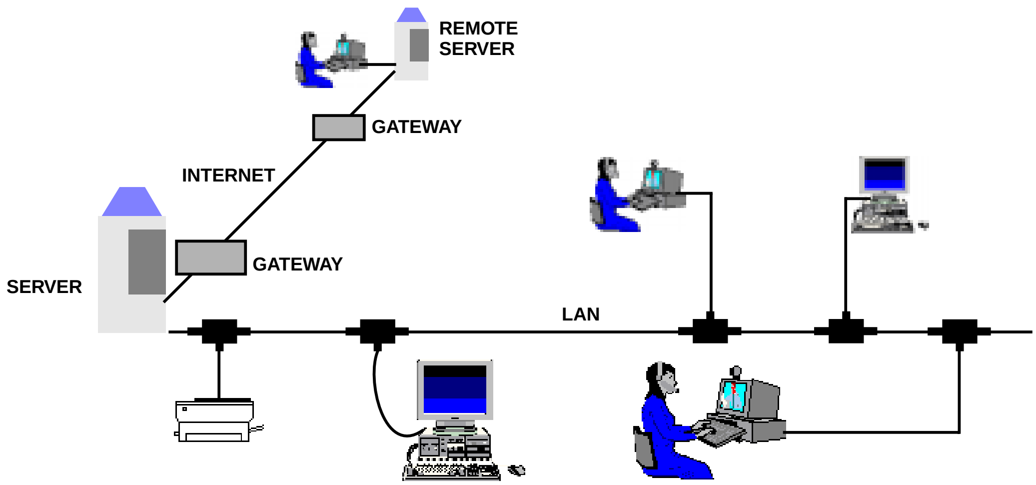


FIGURE 3. H.323 LAN VIDEO CONFERENCING

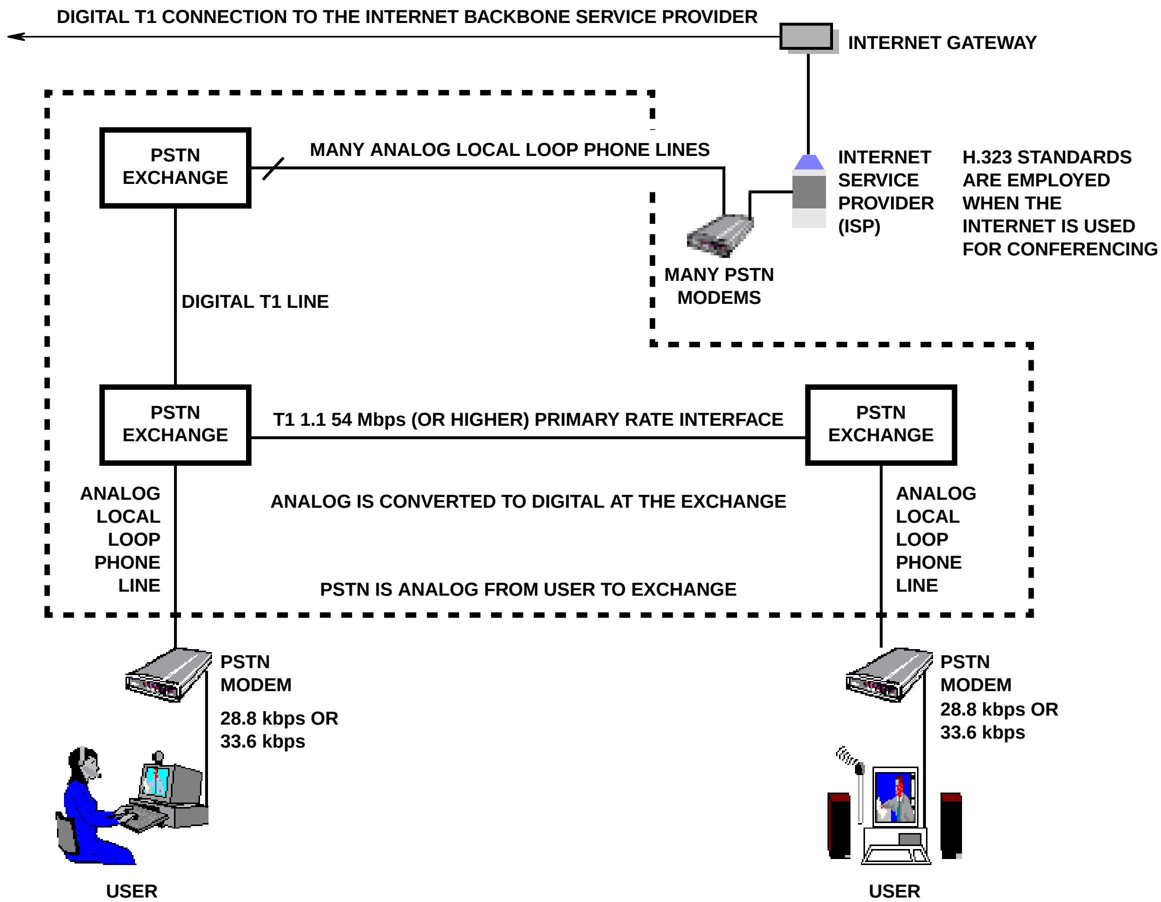


FIGURE 4. H.324 PSTN VIDEO CONFERENCING

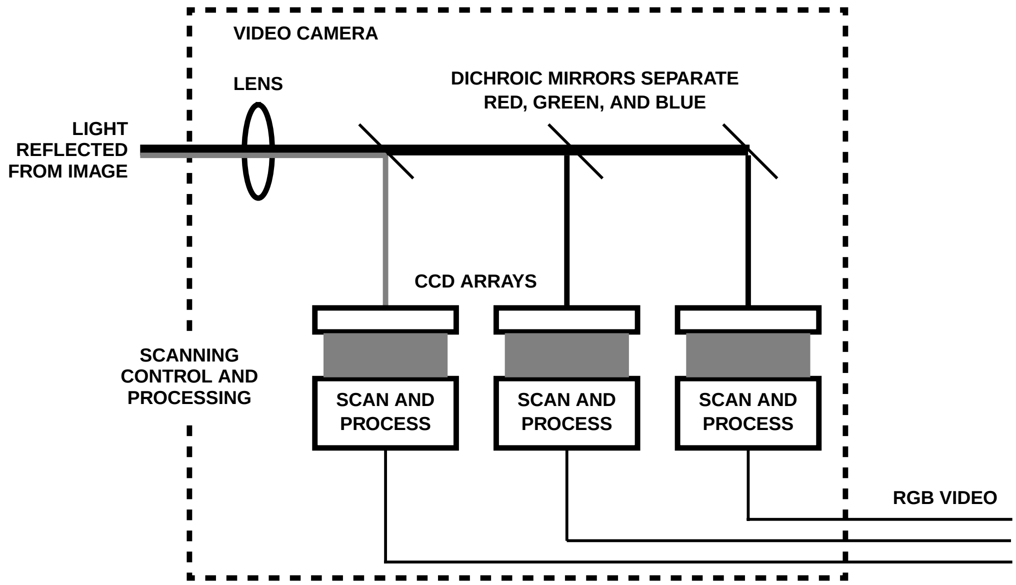


FIGURE 5. VIDEO CAMERA WITH CCD ARRAYS

Component and Composite Video Standards

COMPONENT AND COMPOSITE VIDEO

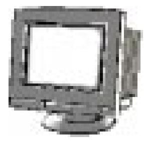
In this section, you will learn about video standards that relate to the manner in which video information is packaged and passed, or transmitted, from one piece of equipment to the next. This involves a discussion of component and composite video types. Component video exists in many different forms. In its purest form, component video is video represented as three separate color signals, a red signal, green signal, and a blue signal, RGB. RGB and other types of component video such as YUV and YCbCr will be discussed in more detail. Composite video is a video signal that contains all of the color and brightness information mixed together in a form that is easily transmitted and can later be separated in the receiver. Our investigation of component and composite video begins with a discussion of RGB.

RED, GREEN, BLUE (RGB) COMPONENT VIDEO

Red, green, and blue are three primary colors that, when mixed together as light, will produce most all visible colors. It is very interesting to note that if the three colors are of equal intensity, they will produce some shade of gray from bright white to black. That range of gray is known as the gray scale. It is also interesting how our eyes mix the three separate colors to create the perception of colors other than RGB including shades of gray. When the RGB are mixed in unequal intensities, some other color is perceived by the eye, a color other than RGB or a shade of gray.

Television screens and computer monitors (Color Cathode Ray Tubes, CRTs) and flat-panel displays, like LCD and plasma displays, all make use of the RGB principle. There are tiny color dots or color rectangles (strips), depending on the type of display, that cover the entire back side of the viewing screen. Some form of energy, whether it be an electron beam, gas plasma, or bright light source, is used to illuminate these color dots at some intensity. The dots are so small, the eye blends them together and mixes the RGB colors to create a perceived color.

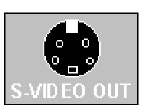
Video cameras function in a similar manner as monitors, but the process is reversed. In some cameras, color image information entering the lens of the camera is separated into primary color images using special dichroic mirrors. As illustrated in Figure 5, the separated red, green, and blue colors are then converted to electrical signals. For example, the resulting red color image is projected onto a sensitive electronic transducer such as a Charge-Coupled Device (CCD). The CCD array converts the image to electrical signals that represent all the red color in the image. The CCD is composed of a two-dimensional matrix, or array, of tiny light-sensitive elements. Each tiny element converts the light that falls on it into an electrical signal whose voltage depends on the intensity of the light. In simple terms, all of the tiny electrical elements in the CCD are electronically scanned and the individual voltage produced by each element becomes part of a signal that represents the red image. The signal may remain as an analog signal or the voltage of each tiny element may be converted to digital (a binary number). Again, in simple terms, the digital binary numbers are streamed together forming the digital signal. The same process is occurring at the same time for the green and blue images that were separated out as described above. The scanning process for the red, green, and blue CCDs is repeated at some rate. Each complete scan of the two-dimensional CCDs produces one field of video. Two fields produce a frame. The number of frames produced every second is known as the frame rate. Thus, each image field and frame is represented by three separate electrical signals that, in turn, represent the primary colors RGB of the source image. As mentioned above, the RGB can be represented as an analog signal (analog RGB) or as a digital signal (digital RGB).

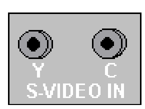


Digital RGB is the form of component video that is generated by computers and sent to the color monitor (RGB monitor). A CRT-type color monitor has three electron guns, one for each RGB color. The intensity information (luminance or luma) contained in each of the color signals is used to drive the corresponding electron gun. The resulting electron beam energizes tiny color dots (phosphorous dots) on the inside back of the screen. The strength of the beam determines the brightness of the color dot. The phosphorous color dot actually produces or emits a color of light (R or G or B). Different phosphors are used for each color. The electron beams are scanned together across the screen from left to right and top to bottom. The CRT is designed in such a way as to allow only the electron beam that carries the red information to fall on red color dots, likewise with the green and blue. Television CRTs work the same as RGB computer monitors, requiring separate RGB information to drive separate electron guns. In a color TV, the RGB signals are derived from the received composite video signal transmitted from the television station or delivered from a VCR.

S-VIDEO

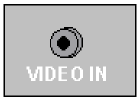
S-Video (Separated Video) is a form of component video composed of brightness (luminance, luma) information, designated with a Y, and separate color (chrominance, chroma) information, designated with a C. Often S-Video is represented as YC or Y/C. Compared to composite forms of video, such as NTSC and PAL, S-Video produces a cleaner, sharper picture, partly because the luma and chroma are already separated and there is no damage done to the color during a separation process.



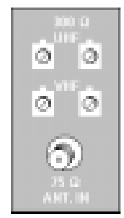


Naturally, an S-Video monitor must have an S-Video source. Some video cameras have S-Video outputs and can deliver separate luma (Y) and chroma (C) information to an S-Video monitor or a Super VHS (S-VHS) video cassette recorder/player. The S-VHS VCR also serves as an S-Video source for the monitor. Keeping the Y and C information separated all along the way yields a visibly-sharper picture than composite schemes. It should be noted here however that S-Video is not the same as S-VHS, though most S-VHS VCRs provide S-Video inputs and outputs.

NATIONAL TELEVISION STANDARDS COMMITTEE (NTSC)



In the early 1950s, the National Television Standards Committee (NTSC) was formed in the United States to tackle a somewhat difficult problem. Their task was to devise some way in which color information could be transmitted and received as normal monochrome video on black and white TV sets. Likewise, black and white TV transmissions were to be viewable on color TVs. Analyzing a black and white TV signal, we realize that the picture information, excluding blanking and synchronization pulses, is only brightness information (luminance, luma, Y). Therefore, the color signal must be formed in such a way as to supply luminance information, from all three primary colors RGB, to the black and white TV. In so doing, color information (chroma, C) is ignored by the black and white TV and only the luma information is used to form the image. When a black and white signal is received by a color TV, no chroma information is present, so the three RGB electron guns of the CRT are all driven with just luma information of equal amounts. Thus, the RGB color dots of the CRT screen are all of the same intensity and only gray scale is perceived by the eye, as was discussed earlier. The standard the committee established was named after the committee, NTSC. Humorists often remark that NTSC stands for “Never The Same Color”.



The composite NTSC signal is a combination of luminance and color information. Basically, the luminance information is contained in voltage variations of the signal and color information is contained in phase differences as compared to a color subcarrier reference signal (color burst). Figure 6 shows an NTSC signal as would be viewed using an oscilloscope. The signal waveform is shown in relation to what would be seen on the TV screen. This is a black and white rendering and, of course, color is not revealed. However, you can see the relationship between the voltage levels, color signals, and the designated color bars on the TV. Note that the black bars on the TV correspond to the black levels indicated in the composite signal.

This composite waveform contains all of the video information for one horizontal scan line from left to right as you face the screen. As the electron beam scans from left to right, its intensity varies in accordance with the up and down voltage swings in the composite waveform. The result is a horizontal scan line that varies in brightness (luminance) across the screen. In a color TV, there are three electron beams scanning left to right together, each beam falling on the proper color dots causing them to glow with a brightness related to the strength of the beam. The result is a horizontal line that not only changes in brightness but also in color. As was mentioned earlier, the composite waveform of Figure 6 is converted to RGB (component video) by the TV receiver circuits in order to drive the three electron guns of the color CRT.

At the end of the scan line, each electron beam is turned off and deflected to the left of the screen (retraced) to start the next line before it is turned back on. This is known as the horizontal blanking interval (HBI) and also includes a horizontal sync pulse which keeps the horizontal oscillator in the TV properly synchronized with the video signal. Notice also in Figure 6 that an 8-cycle to 10-cycle color burst signal at a frequency of 3.579545MHz (~3.58MHz) is riding on the blanking level in the HBI region. This burst is the color subcarrier frequency and is used by the TV to separate the phase-contained color information into component RGB. The color burst is transmitted with the composite video signal to ensure accurate color recovery in the receiver.

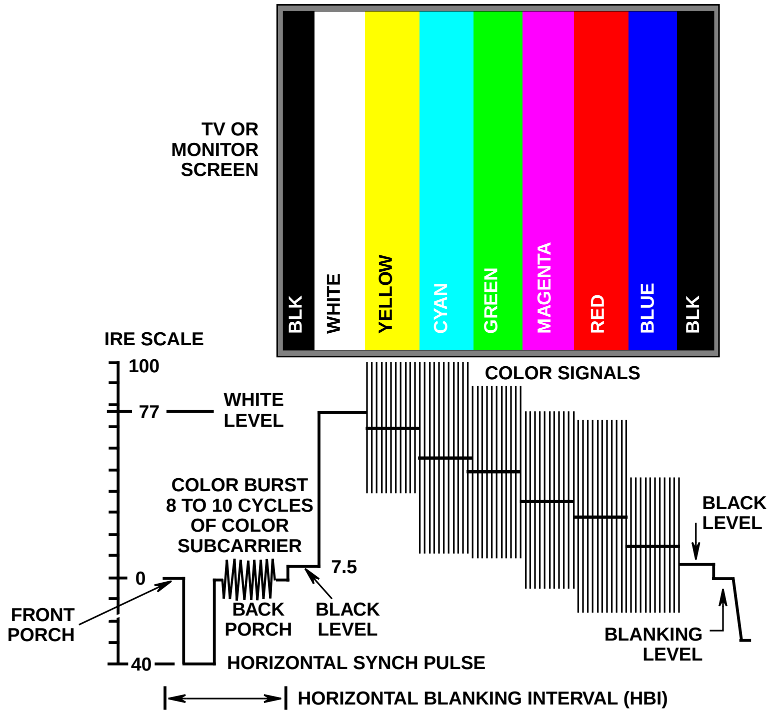
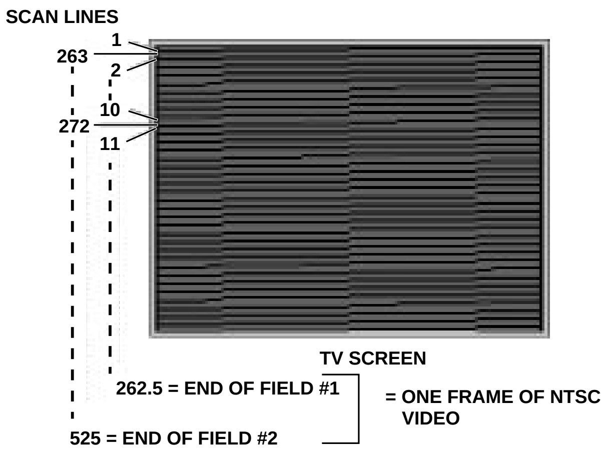


FIGURE 6. NTSC/PAL COMPOSITE VIDEO SIGNAL



NOTE: Not all of the scan lines are used. Some are sacrificed for the vertical blanking interval (VBI).

FIGURE 7. INTERLACED SCANNING - NTSC ILLUSTRATED

As illustrated in Figure 7, the NTSC Standard requires 525 scan lines per frame. The NTSC frame rate is approximately 30 frames per second (fps). Frame rate is very important because if the rate is too low, your eyes will begin to see momentary interruptions and changes to the image known as flicker. A frame rate of 30 fps is considered unacceptable, offering noticeable flicker, if interlacing is not used. This flicker is reduced and the overall video bandwidth requirements reduced by using a technique called interlaced video or interlaced scanning. NTSC makes use of interlaced scanning by scanning every other horizontal line first then going back over the screen to scan the remaining “in-between” lines. Each screen scan of 262.5 lines is called a field. Two interlaced fields make a frame. Thus, the field rate is 60 per second and the frame rate is 30 per second. The high field rate of 60 fields per second eliminates the perception of flicker.

Not all of the horizontal scan lines (262.5 per field, 525 per frame) are actually used. That's because some of the lines are sacrificed to reserve time for vertical blanking. The Vertical Blanking Interval (VBI) is the time during which each electron beam is turned off as the beam is directed back to the top of the screen to begin the next field scan. There are somewhere in the neighborhood of 480 horizontal lines per frame actually used or displayed. Today, the VBI is also used to transmit ancillary data such as closed captioning.

The NTSC Standard is used predominantly in North America. Some other parts of the world have adopted the standard as well. As such, many types of consumer video end-products are manufactured with NTSC inputs and outputs. Many video camcorders, video cameras, VCRs, and TVs have NTSC input and output jacks. The standard acronym is often prefixed with an M as in (M) NTSC. The M refers to the monochrome standard for line/field rates, video bandwidth, RF channel spacing (bandwidth), and audio subcarrier frequency. NTSC indicates the color coding technique. Countries using (M) NTSC, and other standards, are identified in ITU-R BT.470-3.

PHASE ALTERNATION LINE (PAL)

The PAL Composite Video Standard came into use first in the United Kingdom and in Germany in 1967. Since then, many European, African, South American, and Asian countries have adopted the standard in some form. There are many versions of the PAL Standard which we will introduce in a moment.

TABLE 1. NTSC AND PAL COMPOSITE VIDEO STANDARDS

TYPE	LINES	FIELD/S	HOR. FREQ. (kHz)	VERT. FREQ. (Hz)	SUBCAR. (MHz)	BLANKING (IRE)	VIDEO BW (MHz)	AUDIO CAR. (MHz)	CHANNEL BW (MHz)
(M) NTSC	525	~60	15.734	59.94	3.579545	7.5	4.2	4.5	6.0
(M) PAL	525	~60	15.734	59.94	3.575611	0	4.2	4.5	6.0y
(B) PAL	625	50	15.625	50.0	4.433619	0	5.0	5.5	7.0
(D) PAL	625	50	15.625	50.0	4.433619	0	6.0	6.5	8.0
(G) PAL	625	50	15.625	50.0	4.433619	0	5.0	5.5	8.0
(H) PAL	625	50.0	15.625	50.0	4.433619	0	5.0	5.5	8.0
(I) PAL	625	50	15.625	50.0	4.433619	0	5.5	6.0	8.0
(N) PAL	625	50	15.625	50.0	4.433619	7.5	5.0	5.5	6.0
(N) PAL Combination	625	50.0	15.625	50.0	3.582056	0	4.2	4.5	6.0

PAL is similar to NTSC is many ways. Both have horizontal synch pulses, blanking intervals, brightness information contained in amplitude variations, and color information contained in signal phase. PAL also uses interlaced scanning. However, with PAL, certain video color information is inverted in polarity, 180 degrees in phase, with every other (alternating) scan line, thus, the name Phase Alternation Line (PAL). The purpose for this phase alternation is to assist the eye in canceling slight color changes caused by phase distortions in alternating scan lines.

There are 8 different versions of the PAL Standard. These are designated with alphabetic characters B, D, G, H, I, M, N, and so-called Combination N. The differences between them are based on differences in lines per frame (most are 625), fields per second (most are 50), horizontal scan frequency (most are 15.625kHz), vertical scan frequency (most are 50Hz), frequency of the color subcarrier (most are 4.43361875 MHz), blanking level (most are 0 IRE), video bandwidth (ranges between 4.2MHz and 6MHz), audio subcarrier frequency (4.5MHz, 5.5MHz, or 6MHz), and overall channel bandwidth (6MHz, 7MHz, or 8MHz). The (M) PAL version is fully compatible with American black and white TV (monochrome TV). Table 1 summarizes the various PAL versions along with (M) NTSC.

So which is better, NTSC or PAL? PAL has a higher picture definition, and resolution, and thus, provides a sharper picture than NTSC. The number of displayed scan lines is one side of the picture-resolution coin. It defines the vertical resolution. For NTSC there are 480 displayed lines and for PAL there are 576 displayed lines. However, bandwidth is the other side of the picture-resolution coin. Notice PAL uses a bandwidth as high as 8MHz whereas NTSC is only 6MHz. This additional bandwidth gives PAL an advantage that is visible by comparison. A standard test used to demonstrate this is the display of alternating vertical black and white lines. The size of each vertical black and white line is decreased while the number of lines is increased until they are no longer distinct and discernible. An NTSC System can display roughly 330 of these vertical lines before they start to blur, thus, establishing a horizontal resolution of 330 vertical lines (165 black and 165 white). A PAL System can display 420 or so (210 black and 210 white). Thus, the picture resolution, is roughly 330 x 480 for analog NTSC and 420 x 576 for analog PAL. Therefore, the PAL System offers a sharper picture with higher horizontal and vertical (H X V) resolution. ITU-R BT.601 specifies high-definition digital resolutions of 720 x 480 for digital NTSC and 720 x 576 for digital PAL. Other HDTV proposals have resolutions of 1500 x 960 and 1280 x 720. Now that's high definition!

SEQUENTIEL COULEUR AVEC MEMOIRE (SECAM) (SEQUENTIAL COLOR WITH MEMORY)

SECAM is a composite video standard developed and first used in France in 1967. It is a more complex scheme that utilizes two color subcarriers instead of just one and alternates color information on scan lines. Preceding scan lines must be stored in a special memory so the color information of the preceding line can be mixed with color information in the current scan line to yield the primary RGB information to drive the electron guns in the CRT. SECAM has not been as widely adopted as NTSC and PAL. Thus, it is expected that SECAM will eventually be set aside. Most countries of the European Union have adopted PAL.

Color Space Standards

COLOR SPACE?

Color spaces are mathematical and graphical representations of a set of colors that apply to a certain use or application. Primary components of the color space are mixed together to produce a wide range of color and gray scale. For example, printers, both those who engage in the profession of printing and most color printing devices, use the CMYK color space which stands for cyan, magenta, yellow, and black (K). The printer knows how to mix inks in these four colors to produce whatever exact color is needed. Thus, color printing is referred to as a four-color process. RGB is another color space created by color video cameras, color scanners, and computers, and used by color TV picture tubes and computer monitors. YUV is still another color space used in the broadcast of color in the form of composite TV signals, such as NTSC and PAL, and also in the digital processing and storage of video (such as DVDs). Now, let's take a few minutes to consider each of the TV/multimedia-related color spaces.

THE RGB COLOR SPACE

As was discussed earlier when RGB was introduced, all colors including all shades of gray can be produced by mixing red, green, and blue light. If the colors are of equal intensity, at all levels of brightness, a gray scale is produced. In other words, black through all shades of gray to white is perceived by the eye as equal RGB color intensities are increased from zero.

The RGB color space is represented graphically in Figure 8. Notice this is a three-dimensional graph since any perceived color or shade of gray is created by the interaction of all three colors. Looking at the graph, you will notice that when red is 0, and green and blue are kept at equal intensity, some intensity or saturation of cyan is produced. When green is 0, and red and blue are kept at equal intensity, some intensity or saturation of magenta is produced. When blue is 0, and red and green are kept at equal intensity, some intensity or saturation of yellow is produced. When all three colors are used, all hues and saturations of colors will be produced depending on the intensity of each primary color. As we discussed earlier, this is how color is electronically captured by video cameras and electronically displayed by color monitors (including color CRTs and flat-panel displays).

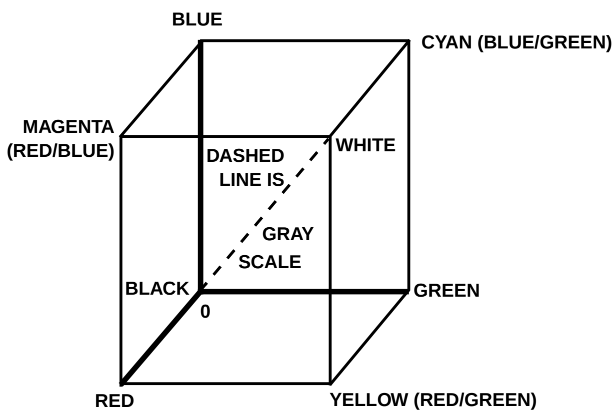


FIGURE 8. THE RGB COLOR SPACE

The intensities of RGB signals produced by video cameras are often adjusted before further conversions and processing. This is to correct for electrical-to-light conversion nonlinearities that exist in CRT displays. The different color phosphors do not produce the same intensities of light for a given level of electron-beam excitation and the amount of light they produce does not vary linearly with changes in excitation. Thus, a deliberate modification of the RGB color space is needed to correct this. This is referred to as gamma correction. Gamma is a number (usually 2.2 to 2.8) that is used as an exponent of the excitation signal voltage. For example, the brightness of the red color displayed is a function of the applied red excitation signal raised to the 2.2 power. In other words, $R_{BRIGHTNESS} = R_{VOLTAGE}^{2.2}$. If you would graph this relationship, you would see an extremely nonlinear curve. To compensate for this, the RGB signals from the camera are gamma corrected so the graph for each color bends the other direction. The corrected color information will then cancel out the nonlinear response of the CRT. Gamma-corrected RGB is represented as R'G'B'.

RGB information is usually processed in digital form using 8-bit binary numbers to express each color. Modern video cameras provide digital RGB, color scanners generate digital RGB, computers generate and process digital RGB, and other color spaces can be derived from digital RGB using DSP ICs. Harris Semiconductor makes these ICs.

THE YUV COLOR SPACE

The YUV color space, and other very similar color spaces, are used in creating composite video information such as NTSC, PAL, and SECAM. This composite information is then used to amplitude modulate a television carrier frequency for broadcasting. The Y stands for luminance, or luma, and is the brightness information which is required to provide the black and white TV picture information. The U and V are two separate color signals that contain the important color information from which all colors can be reproduced. Of course, all three video components YUV are mixed together to produce the composite video signal. However, before they are mixed they are created as three separate signals, Y, U, and V.

The Y signal is the luminance, luma, or brightness that represents, or is derived from, the luma contained in all three primary colors RGB. Recall that RGB is produced by video camera sensors like CCD arrays and by computers to drive RGB monitors. Thus, in the process of converting the RGB information to a composite signal like NTSC, Y information must be obtained that is derived from all three primary colors. This is done electronically in a mixing process. In a final mixing step, the Y information will become part of the composite video signal which is used to modulate a carrier frequency to be transmitted. The Y information contains all of the black and white picture information needed by black and white (monochrome) TVs. The conversion and processing of component video information such as RGB and YUV can be done in the analog world with analog circuits or it can be done in the digital realm where sophisticated processing can be performed.

The U and V components, of YUV, contain different color information that a color TV receiver can use, along with the luma (Y), to recover the original RGB to drive the CRT. On the transmit side, near the RGB video source (camera) U and V are derived electronically from the original RGB through a special analog mixing process or through digital processing. U and V are often described as “color difference” video signals because it can be shown mathematically that U is expressed as blue - luma (B - Y) and V is expressed as red - luma (R - Y). For our purposes here, we will not get tangled up in the actual mathematics, circuitry, and extended theory. The important concepts here are:

- **Y, U, and V Are Component Video Elements, Electronically Derived from Source RGB, Either in Analog or Digital Form**
- **Y, U, and V are Used to Create the Composite Video Signal, Such as NTSC, to be Mixed With a Television Carrier Frequency For Transmission**
- **It is Necessary to Generate YUV Components so the Y can be Used by Black and White TVs**
- **In color TV Receivers, YUV Components are Separated from the Composite Video. The Recovered YUV Components are then used to Recreate the Original RGB Signals to Drive the Electron Guns in the CRT**
- **YUV is Called a Color Space Because all Colors and Shades of Gray are Encoded in the YUV Signals**

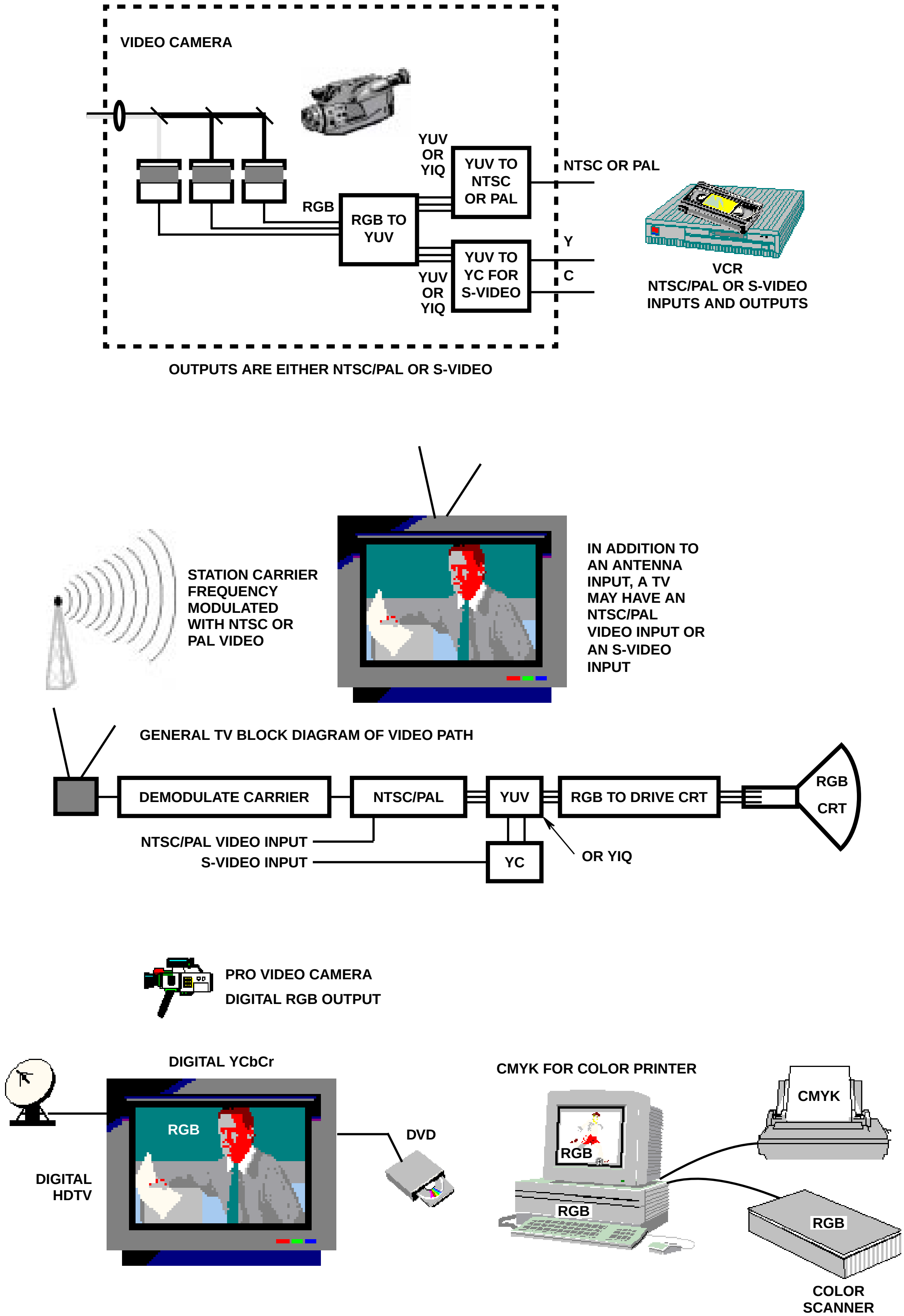


FIGURE 9. COLOR SPACES RELATED TO APPLICATIONS

OTHER COLOR SPACES SIMILAR TO YUV

Other color spaces, very similar to and derived from YUV, are also used in the conversion process from RGB to a form of composite video (NTSC, PAL, SECAM). The most common are YIQ and YCbCr. YIQ is sometimes used in place of YUV in NTSC systems. As before, the Y represents luma, but the I, in YIQ, stands for “in-phase” while the Q stands for quadrature, which means 90 degrees out of phase. This indicates that a quadrature phase modulation technique is used in the mixing process.

YCbCr is a digital color space in which each of the video components is expressed in and processed in digital binary form. It was developed as part of ITU-R BT.601 titled: Studio encoding parameters of digital television for standard 4:3 (standard TV screen dimensions) and wide-screen 16:9 aspect ratios (HDTV screen dimensions). As before, Y stands for luma, Cb stands for the blue color difference signal, and Cr stands for the red color difference signal. There are various digital formats that are used to represent YCbCr. You will be introduced to those a little later. For the moment, let's summarize the color spaces we have discussed.

SUMMARY OF COLOR SPACES

Figure 9 illustrates where each color space is used. It is designed to help you understand color spaces in context of application and also serves as a summary or review of what has been discussed. By examining Figure 9, you note the following:

- **CMYK is a Color Space Used for Color Printing - creating Color Images Using Paints, Inks, Etc.**
- **RGB is Generated by Video Cameras, Computers, and Color Scanners and is Used by Computer Monitors, Flat-panel Displays, and the CRT of Color Televisions**
- **YUV and YIQ Exist as an Intermediate Conversion Step Between RGB and NTSC/PAL on the Transmit End and Between NTSC/PAL and RGB on the Receive End. They Can Be Processed in Analog or Digital Form**
- **YCbCr is in Digital Format and is Used in Digital TV Systems as a Means of Processing, Storing, and Transmitting, Usually, From an RGB Source Through a Communication System to a Video Display**

Pixels and Screens

SCREEN PIXELS

Display screens of all types, whether CRT or flat-panel such as LCD and plasma displays, are made up of a large two-dimensional horizontal and vertical array of tiny screen elements called pixels. The name pixel is a contraction derived from the words picture element. Screen pixels are illustrated in Figure 10. Technically, as applied to color display screens, a pixel is a group of three tiny color elements (RGB) sometimes called color dots or color strips. In some CRTs, the color elements are literally tiny round dots, thus the name “color dot”. In other CRTs and flat-panel displays, the color elements are not dots but tiny rectangular strips. In LCDs the color elements are color strip filters (R, G, and B) through which light passes. In CRTs and plasma displays, the color elements are deposits of phosphorous materials that emit colored light when electrically energized. There is a different phosphor for each color, RGB.

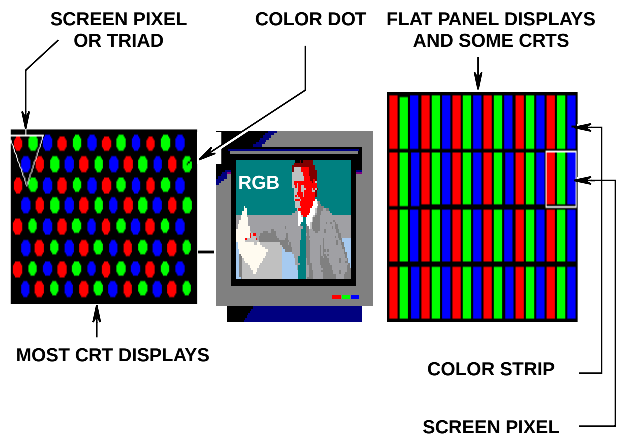


FIGURE 10. COLOR DOTS, COLOR STRIPS, AND SCREEN PIXELS

Incidentally, you will often hear and read the word pixel being used incorrectly to describe a single color dot. For example, “The red pixel...”. This is a minor infraction that usually causes little harm since the meaning of what is being explained is usually revealed in context. For black and white TV screens or monochrome monitors, the word pixel does apply to a single light-emitting element. For color displays, three color elements RGB are needed to form a screen pixel. To avoid potential confusion, pixels are sometimes called triads.

One of the most common parameters used in selecting a display is the display's dot pitch. For example, a high-resolution color monitor may have a dot pitch of 0.26mm. That means, the distance from the center of one red color dot to the center of the next red color dot is 0.26mm (millimeters) or 260 microns (likewise from blue dot to blue dot and green dot to green dot). To give you an idea of how small 260 microns is, the thickness of a sheet of copy paper (20lb wt.) is about 100 microns.

As we discussed earlier, the perceived color emitted by the entire RGB pixel will depend on the relative intensity, or brightness, of each color element. The brightness of each individual color element will depend on the signal information used to excite the element. Recall from our previous discussions on color spaces and sample formats, the signal information may be processed as analog signals or as digital sample groups of various formats. To drive the display, separated RGB information must be recovered and used to drive each color dot (element).

Using magnification of 50X or more, you can examine the screen of a display that appears to be white and discover that in fact the screen is actually made up of the individual red, green, and blue color elements just as we have discussed. Examine other areas of the screen that appear to be different colors and you will discover that each of the color dots are of different brightness, again proving what has been discussed.

SCREEN RESOLUTIONS

Screen resolution, for computer displays of all types, CRT, LCD, plasma, etc., is expressed as so many horizontal pixels by so many vertical pixels. As shown in Figure 11, a VGA display is 640 X 480. Thus, there are 640 screen pixels horizontally and 480 screen pixels vertically. The entire screen has 307,200 pixels and 921,600 individual color elements (dots or strips). Higher screen resolutions above VGA are: Super VGA (SVGA) = 800 x 600, extended VGA (XVGA) = 1024 x 768, and even higher formats such as 1152 x 900, 1280 x 1024, and 1600 x 1200. (VGA stands for Video Graphics Adapter).

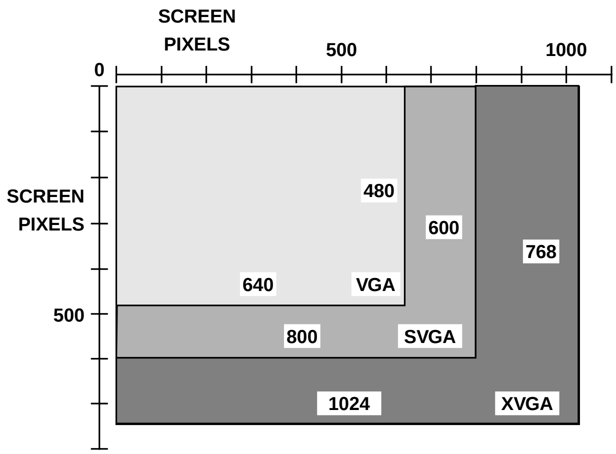


FIGURE 11. COMMON SCREEN RESOLUTIONS

ASPECT RATIOS

International standard ITU-R BT.601 specifies studio encoding parameters of digital television for Standard 4:3 and wide-screen 16:9 aspect ratios. As illustrated in Figure 12, aspect ratios are simply the ratio of the horizontal dimension to the vertical dimension. If something is square, its aspect ratio is 1:1. The screens of standard television sets have aspect ratios of 4:3 as do most computer monitors where screen resolutions of 640/480 = 800/600 = 1024/768 = 4/3 = 4:3. New wide screen TVs have aspect ratios of 16:9 which is in line with the aspect ratio of movie film viewed in theaters. Currently, movies made for theater must be converted or modified for home 4:3 TV viewing.

IMAGE PIXELS

There are pixels and there are pixels. There are the screen pixels we have just discussed and there are image pixels related to the picture itself apart from the screen on which it is displayed. The image pixel is a single image unit, of many, on any horizontal scan line. It is the smallest discernible element of the picture itself, usually based on system bandwidth or on a digital sample rate, where an image is expressed in digital form.

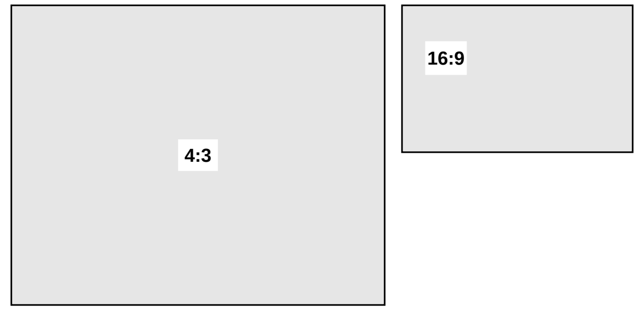


FIGURE 12. SCREEN ASPECT RATIOS

Here is something you may already know about image pixels but didn't even realize it. You may be using an XVGA (1024 x 768) monitor with your computer. However, you may have the image resolution set at just VGA (640 x 480). So, you are viewing an image composed of 640 x 480 image pixels but you are using 1024 x 768 screen pixels to form the image. In other words, one image pixel is formed using more than one screen pixel.

Another example illustrating the relationship between image and screen pixels is television images. Recall earlier how the resolution of an analog NTSC System is roughly 330 x 480. 330 discernible vertical lines define the horizontal picture resolution, and 480 horizontal scan lines define the vertical resolution. In other words, in an NTSC System there are roughly 330 x 480 image pixels. However, the screen itself has many more screen pixels that are actually used to create the image.

As far as displaying an image is concerned, the screen pixels determine the maximum possible image resolution. In other words, the maximum image resolution is limited by the number of available screen pixels. If a screen only has 320 x 240 physical screen pixels, you cannot expect to display an image at a VGA resolution. Likewise, you cannot display SVGA images on a VGA monitor.

RECTANGULAR AND SQUARE PIXELS

Now you are at a point to understand the difference between rectangular pixels and square pixels. First, it's important for you to understand that we are talking about image pixels here. For NTSC television we said there are 330 x 480 image pixels and the screen has an aspect ratio of 4:3 (width:height). That means there are 330 image pixels horizontally that corresponds to the width of the screen and 480 image pixels corresponding to the height of the screen. Thus, image pixels are stretched 330 across the 4-unit dimension and 480 down the 3-unit dimension. As you can see, the image pixel is wider than it is taller. So, we say these image pixels are rectangular. Also notice the image pixel ratio of 330:480 does not match the 4:3 screen aspect ratio. Stated another way, the image pixel is rectangular if the horizontal distance between them does not match the vertical distance between them.

If the image pixel ratio does match the screen aspect ratio, the image pixels are square. If we have 640 x 480 image pixels displayed on a screen with a 4:3 aspect ratio, the image pixels are square. This is like displaying a 680 x 480 image on a 4:3 VGA/SVGA/XVGA monitor. So, if the image pixel resolution ratio matches the screen aspect ratio, the image pixels are square. Also, if the horizontal and vertical spacings of the image pixels are equal, the image pixels are square. The number of horizontal pixels needed for the pixels to be square is calculated by multiplying the number of active scan lines by the screen's aspect ratio. For NTSC, 480 lines x 4/3 = 680 image pixels per line.

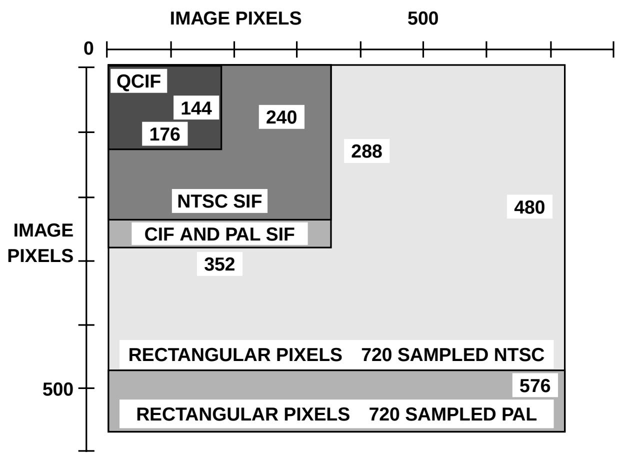


FIGURE 13. COMMON IMAGE RESOLUTIONS

IMAGE RESOLUTIONS AND FORMATS

CIF and QCIF

The ITU-T H.261 Video Compression Standard for video conferencing specifies certain resolutions of displayed video images. The highest specified resolution is called the Common Interface Format (CIF), illustrated in Figure 13. CIF has a resolution of 352 x 288 image pixels. The CIF forms a video image window that is a little more than 1/4 the area of a VGA screen, assuming no scaling is employed and one screen pixel is used to display one image pixel. This is more than adequate for video conferencing. The second resolution specified by H.261 is the Quarter Common Interface Format (QCIF) which is 176 x 144 image pixels. QCIF is not as clear or sharp as CIF but it takes much less bandwidth to transmit, an important consideration especially when using a standard modem and the PSTN for the connection. In either case, CIF or QCIF, the image can be scaled to a larger viewing size with no improvement in image resolution.

Digital NTSC and PAL Image Resolutions

ITU-R BT.601 (formerly CCIR 601) specifies the image resolution for digitally-encoded NTSC and PAL. The specification for digitized NTSC is 720 x 480. That means there are 720 image samples, represented digitally, for each of 480 displayed horizontal scan lines. For digitized PAL, the image sample resolution is specified as 720 x 576. These image resolutions are also illustrated in Figure 13. When a computer monitor is used to display a digital NTSC or PAL video image, the square-pixel image resolution is specified as 640 x 480 for NTSC and 768 x 576 for PAL. Note that these image resolutions match the 4:3 aspect ratio and the 4:3 screen pixel ratio of a VGA or SVGA display (640 x 480 and 800 x 600).

SIF and QSIF

Video Compression Standards such as MPEG-1 and MPEG-2 use one of the Standard Input Formats (SIF) or Quarter Standard Input Formats (QSIF) to define the video window (image) resolution. Where SIF is applied to applications that deliver video to NTSC monitors, the image resolution is specified as 352 x 240 image pixels with a refresh rate of approximately 30 fps. This is shown in Figure 13. Note that 352/240 does not equal the screen aspect ratio of 4:3. Thus, the image pixels are not square. Where SIF is applied to PAL televisions the resolution is specified as 352 x 288 image pixels and a refresh rate of 25 fps. Again, the image pixels are not square since 352/288 does not equal 4/3. Square-pixel SIF dimensions are employed when using a computer monitor to display the video images. For NTSC, the square-pixel resolution is 320 x 240 and for PAL it is 384 x 288. In each case, the image resolution is equal to 4:3. QSIF resolutions are half the dimensions given for SIF.

YCbCr Digital Sampling

VIDEO IMAGE SAMPLING

Video images come from video cameras and need to be edited, processed, stored, and retrieved from various media. That means the video needs to be in some robust form so as not to suffer degradation in handling (processing, etc). Thus, the video source is converted to some digital format whether the image source was an analog NTSC/PAL, analog RGB, or digital RGB. The conversion process involves the capture of image samples at some predetermined resolution, meaning, so many samples per horizontal scan line of the video image. These image samples contain YCbCr information in digital form. There are many different digital sample formats that contain YCbCr image sample information. The actual sample format used depends on the application and the standards that govern the application. The following are some of the most common sampling formats.

YCbCr SAMPLING FORMATS

YCbCr component video information is expressed in digital form. That means each component, Y, Cb, Cr, is expressed as a binary number. The value of the binary number represents the level or strength of that component. The binary numbers used are either 8-bits or 10-bits in length such as Y = 01110010. The binary numbers representing Y, Cb, and Cr are specially combined forming a data string, or word, which carries a sample of color video image information. Sampling formats are expressed numerically in relation to Y:Cb:Cr such as 4:4:4, 4:2:2, 4:1:1, and 4:2:0. Let's consider each of these to understand what the sampling formats mean.

4:4:4

The 4:4:4 sampling format indicates that every sample, of a group of 4 samples, includes an equal representation of all three components, YCbCr. In this case a 24-bit or 30-bit string is used which includes 8-bits or 10-bits for each of the components. This digital sample contains video information for a single sample (image pixel) on a horizontal scan line of the color image. In other words, the sample string is a digital representation of an image pixel. Hundreds of these samples are collected from the image scan line.

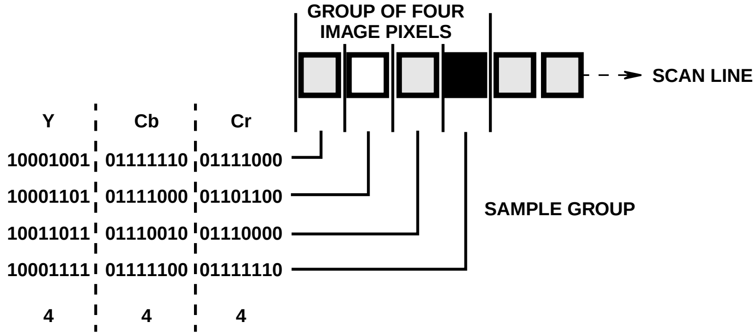


FIGURE 14. THE 4:4:4 DIGITAL SAMPLE FORMAT

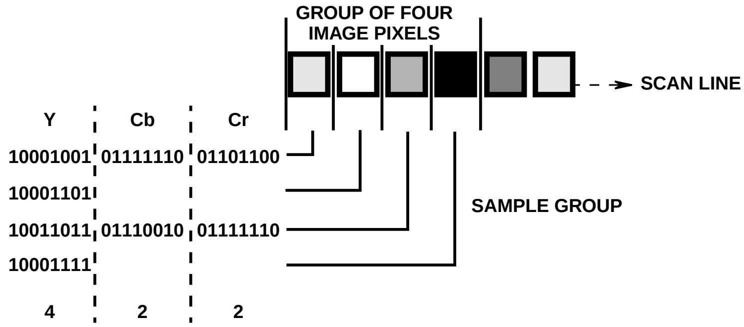


FIGURE 15. THE 4:2:2 DIGITAL SAMPLE FORMAT

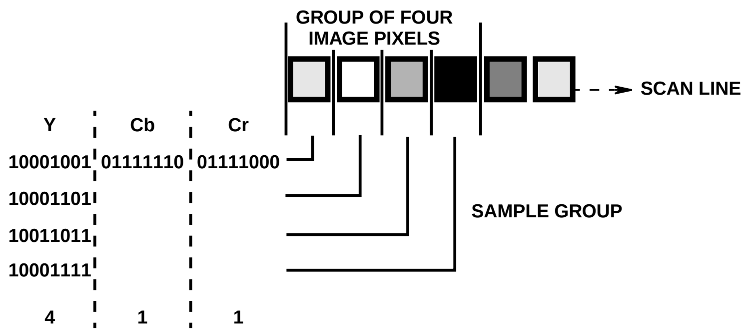


FIGURE 16. THE 4:1:1 DIGITAL SAMPLE FORMAT

So why don't they express the sample as 8:8:8 or 10:10:10? As illustrated in Figure 14, the 4 indicates that out of 4 consecutive image samples on a horizontal scan line, Y is represented in each sample as an 8-bit or 10-bit binary number. Likewise, out of 4 consecutive samples, Cb and Cr are always represented in each sample as an 8- or 10-bit binary number. You will see this more clearly as other formats are discussed. The 4:4:4 format is most often used for high-resolution computer graphics applications where the concern is not data transmission rate or bandwidth but color accuracy. Using the 4:4:4 sample format, no color information is lost.

4:2:2

The 4:2:2 sampling format indicates that the Y information is represented in every sample group of four consecutive samples just as it was in the 4:4:4 format. However, as illustrated in Figure 15, the Cb and Cr information is only sampled from every other image pixel. Thus, Cb and Cr are represented in only 2 sample strings of any 4 consecutive samples. This digital video information is then formed into 16-bit data strings for transmission or storage. Each 16-bit data string is composed of 8-bits of luma and 8-bits of the Cb or Cr information that was obtained from every other image pixel. Therefore, the transmitted or stored 16-bit data strings that originated from the 4:2:2 samples are smaller than the 4:4:4 format, 24-bit sample strings. The information has been reduced. That translates into faster data transfer of picture information, less storage room needed, and a lower system bandwidth requirement. A disadvantage to the 4:2:2 format is that picture information is less complete and full-color RGB screen pixel information has to be interpolated from the YCbCr 4:2:2 information in order to drive the display. This is accomplished by reforming the 16-bit data strings back into the original samples, shown in Figure 15, then, interpolating missing color information. This interpolation process recreates 24-bit YCbCr information which is used to create the RGB needed to drive the CRT in the monitor. The 4:2:2 sample format is used for studio production environments, professional editing equipment, distribution, servers, and some consumer products.

4:1:1

The 4:1:1 format was largely used in inexpensive consumer equipment but has, for the most part, been replaced by the 4:2:2 format. As shown in Figure 16, the 4:1:1 format indicates that there is an 8-bit Y component for every sample, there is one 8-bit Cb component for every four scan-line samples, and there is one 8-bit Cr component for every four scan-line samples. Similar to the 4:2:2 format, the 4:1:1 sample information is reformatted into data strings for storage or transmission. This time the data strings are only 12-bits instead of 16 or 24. To display this information, it must first be returned to its original format, like when it was sampled as shown in Figure 16. Then, interpolation is performed to create complete 24-bit data strings representing YCbCr. This is then converted to RGB to drive the CRT in the display. Naturally, the reproduced and displayed color is not very accurate but the digital information has been reduced allowing faster data transmission or lower system bandwidth.

4:2:0

The 4:2:0 sample format is different than 4:4:4, 4:2:2, and 4:1:1. These formats represent four consecutive samples on the same horizontal scan line while 4:2:0 represents two samples on one scan line and two samples below them on the next scan line forming a group of four. See Figure 17. Thus, a 2:1 color reduction ratio exists both horizontally and vertically. Only one Cb and Cr sample exists for every two horizontal and two vertical samples. The 4:2:0 format is used for mainstream digital television and consumer entertainment equipment using the MPEG-1 and MPEG-2 Standards.

Table 2 provides a summary of the various image formats we have discussed and the YCbCr digital sampling formats that are often used to store or transmit them. However, the YCbCr sampling formats are actually independent of the image formats.

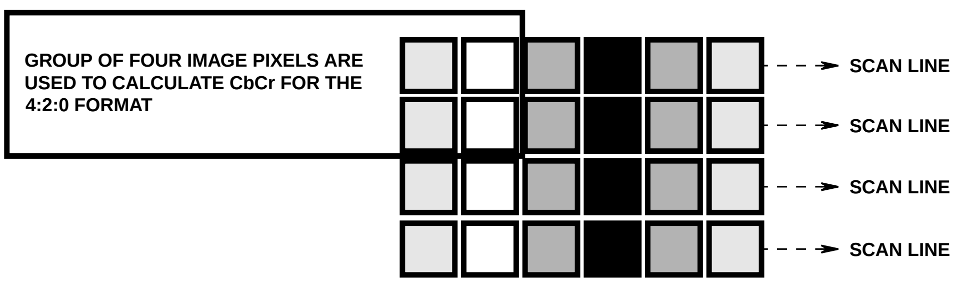


FIGURE 17. THE 4:2:0 DIGITAL SAMPLE FORMAT

TABLE 2. IMAGE AND SAMPLE FORMATS

IMAGE FORMAT	IMAGE RESOLUTION (NTSC)	IMAGE RESOLUTION (PAL)	YCbCr SAMPLE FORMAT
QCIF	176 x 144	176 x 144	4:2:0
CIF (H.261 Std)	352 x 288	352 x 288	4:2:0
QSIF	176 x 120	176 x 144	4:2:0
SIF	352 x 240	352 x 288	4:2:0
Square Pixel SIF	320 x 240	384 x 288	4:2:2
Square Pixel QSIF	160 x 120	192 x 144	4:2:2
MPEG2	720 x 480	720 x 576	4:2:0
MPEG2	720 x 480	720 x 576	4:2:2
BT.601 (Rectangular)	720 x 480	720 x 576	4:2:2
BT.601 (Square)	640 x 480	768 x 576	4:2:2

Computer Bus Standards

THE ISA BUS

ISA stands for Industry Standard Architecture. It is the architectural bus standard for the IBM™ XT (8-bit) and the IBM™ at (16-bit) and has been used in most IBM™ PC clones since 1983. In ISA systems, a system enhancement card is added by plugging the card into one of the 8-bit or 16-bit expansion slots. Obviously, the ISA bus has been around for a long time and now has little to offer modern PC systems. The reason? It's too slow for today's microprocessors and represents a serious data-flow bottleneck between the CPU (central processing unit or microprocessor) and plug-in peripheral interfaces and accelerator cards (such as audio and video capture and processing cards). While ISA crawls along at about 8 Mbyte/s, microprocessors are processing data at clock rates over 200MHz.

THE EISA BUS

EISA stands for Extended Industry Standard Architecture and was developed by the Video Electronics Standards Association (VESA). It is a computer bus standard compatible with ISA but has a 32-bit data bus and a 16 Mbyte/s data rate. This is another computer bus that is falling out of use.

THE VL-BUS

The VESA Local Bus, designated VL-Bus, was designed for high-speed data transfer between the computer mother board and adapter cards. It is intended to compete with the Intel PCI Bus, having similar specifications, but is not experiencing wide industry acceptance. VL-bus data transfer rate is in the area of 100 Mbyte/s using a 32-bit parallel data bus.

THE PCI BUS

PCI stands for Peripheral Component Interconnect. PCI is a 32-bit parallel data bus operating at data rates over 130 Mbyte/s. This bus is a purely proprietary bus standard invented by Intel. Even though PCI is a proprietary standard, it has been accepted widely throughout the computer industry for all types of desktop computers including Apple™ and IBM™ PC clones. Many customers look to see that a computer has PCI bus card slots before making a purchase. Harris Semiconductor is designing many of their multimedia ICs with PCI interface capability.

THE UNIVERSAL SERIAL BUS (USB)

The USB is a high-speed serial interface for PC peripheral devices and is expected to eventually replace the older, and much slower, RS-232 serial interface. The USB operates at data rates of 1.5 Mbps and 12 Mbps. It is now available in most PCs and peripherals such as small LCD control panels, video cameras, mice, and keyboards. This bus is a true “plug-and-play” bus making peripheral device interconnection very simple.

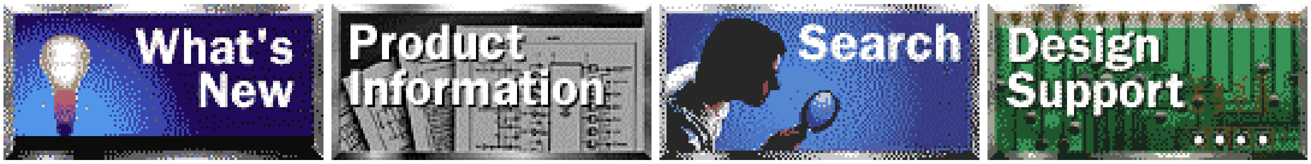
FIREWIRE IEEE 1394 SERIAL INTERFACE BUS

The IEEE 1394 serial bus was invented by Apple Computer, Inc. and was originally named FireWire. After its adoption by the IEEE, it became known as IEEE 1394. It is known by yet another number through its adoption by the IEC, IEC 1883. The rapid acceptance of FireWire is due to its very high data transfer capability of up to 400 Mbps. Data rates for FireWire are expected to climb to as high as 1600 Mbps in 1998. FireWire will be used to interconnect a wide range of devices and systems to the PC such as professional video editing equipment, home control devices, security systems, color printers, and CD ROM drives, to name a few. It is expected that FireWire will replace current serial and parallel data ports on PCs such as the parallel printer port, RS-232 serial port, and the parallel SCSI port (Small Computer System Interface). Firewire is not only faster, but it is a true “plug-and-play” interface. Devices are simply plugged into a multi-connector adapter (hub) and the system automatically sets itself up to recognize and control the device. Using 16-bit addressing, up to 64,449 devices can be connected to the same FireWire system. This is far greater than any bus of the past.

WHAT IS MMX TECHNOLOGY?

MMX was developed by Intel Corporation for a new line of Pentium microprocessors. MMX stands for Multi Media extensions. Intel extended the normal Pentium instruction set by adding 57 multimedia-related instructions. Only a Pentium MMX microprocessor, or one that is fully compatible, is able to use these new instructions. Thus, software written using MMX instructions will not run on other Pentiums or similar competing microprocessors. What MMX offers in terms of performance is faster processing of full-motion video, audio, and 3D graphics for faster, smoother animation. The Pentium MMX can perform video and audio compression and decompression under software control as opposed to using a hardware processor such as a video codec. One application for MMX technology is audio/video/data conferencing. Even though MMX offers many multimedia software solutions, such as conferencing, the consumer/customer still pays for hardware because of the premium price paid for an MMX Pentium processor. Manufacturers such as AMD, Cyrix, and Centaur Technology are currently developing, or have available, MMX compatible microprocessors. Prices on MMX based PC systems are dropping, demand is high, and sales are brisk.

http://www.semi.harris.com/



WHAT'S NEW

- **Press Releases**
- **New Services**
- **New Web Material**

PRODUCT INFORMATION

- **Organized by Device Function**
- **Product Information Page Links to:**
  - **Data Sheets**
- **>2500 Data Sheets and Application Notes**

SEARCH

- **Search Based Upon Part Number or Description**

DESIGN SUPPORT

- **Application Note Listing**
- **Tech Brief Listing**
- **Downloadable Design Software**
- **Evaluation Boards Listing**
- **Lexicon**
- **E-mail To Central Applications Group for Technical Help**

OTHER LINKS

- **Sales Office and Distributor Listing**
- **Target Application Sites**
- **Quality/Reliability**
- **Webmaster E-mail for Site Comments**

All Harris Semiconductor products are manufactured, assembled and tested under **ISO9000** quality systems certification.

Sales Office Headquarters

For general information regarding Harris Semiconductor and its products, call **1-800-4-HARRIS**

NORTH AMERICA
Harris Semiconductor
P. O. Box 883, Mail Stop 53-210
Melbourne, FL 32902
TEL: 1-800-442-7747
(407) 729-4984
FAX: (407) 729-5321

EUROPE
Harris Semiconductor
Mercure Center
100, Rue de la Fusee
1130 Brussels, Belgium
TEL: (32) 2.724.2111
FAX: (32) 2.724.22.05

ASIA
Harris Semiconductor PTE Ltd.
No. 1 Tannery Road
Cencon 1, #09-01
Singapore 1334
TEL: (65) 748-4200
FAX: (65) 748-0400

HARRIS SEMICONDUCTOR