1

DIGITAL VIDEO DELIVERY

Television has in many ways promoted understanding and cooperation among people all over the world. About 600 million people saw the first person walk on the moon and a billion people watched the 20th Summer Olympic Games. By 2012, there were over 3.6 billion viewers for the 30th Summer Olympic Games. TV watching used to be confined primarily to the living room. This has changed. The ubiquity of HD-capable smartphones and tablets equipped with powerful video decoders enables TV view time on mobile devices to surpass view time on the TV. The Internet has become a key media distribution platform that has opened up new ways for discovering, sharing, and consuming TV content anywhere, anytime, and on any device. Online Internet TV providers are trumping cable and satellite pay-TV providers with a dramatic increase in subscription and advertising revenue in recent years. Ala-carte-style Internet TV has now started to break the traditional pay-TV distribution model that is based on channel bundling. Although pay-TV providers are making TV content available online for their subscribers, they may now have to migrate to online streaming boxes and ditch the venerable set-top. Cable is now a broadband business. Among the top cable providers, broadband Internet service accounted for more subscribers than cable TV. Ultimately, pay-TV providers may have to rely on broadband to grow profits. Currently, over 10 million US households are broadband-only. In this chapter, we analyze these game-changing trends in digital video delivery.

Next-Generation Video Coding and Streaming, First Edition. Benny Bing.

1.1 BROADBAND TV LANDSCAPE

Over 80% of Internet users watch video while 30% of these users watch TV. In the United States, viewers spent an average of more than 6 h/month watching video on the Internet. Streaming live sports programs online makes truly national or global events possible, reaching millions of consumers via handheld devices. There are several challenges. Due to the mobility of subscribers and the heterogeneity of the user devices, the streaming server has to adapt the video content to the characteristics and limitations of both the underlying network and the end devices. These include variations in the available network bandwidth and user device limitations in processing power, memory, display size, battery life, or download limits.

1.1.1 Internet TV Providers

The emergence of over-the-top (OTT) online content providers such as Netflix, Hulu, and Amazon offers more choices to the consumer by providing replacement or supplementary TV services, usually TV shows and movies but no sports programming. The service is either free or much cheaper than pay-TV and this has led to a steady migration of subscribers from pay-TV to online TV, despite efforts from cable and satellite pay-TV providers in making TV content available online. Unlike OTT providers, these "TV Everywhere" Web portals may include both sports and video on demand (VOD) or time-shifted TV content.

Because broadcast pay-TV tends to surpass online TV in visual quality, it delivers better overall experience. For example, the quality of online TV service may fluctuate according to the bandwidth availability on the broadband Internet connection. However, other factors such as choice of content, flexible viewing time, and content portability are also important for the consumer. Thus, hardware set-top box (STB) vendors (e.g., TiVO) traditionally aligned to linear broadcast programming cable TV service have integrated OTT streaming content to their channel lineup.

Pay-TV operators may follow this lead and let OTT content into their STB. In doing so, subscription rates may be lower compared to traditional pay-TV subscription. For example, Walt Disney recently signed a carriage deal with Dish Network, making it the first pay-TV provider to bundle ABC, ESPN, and other channels owned by Disney in a TV service delivered entirely over the Internet. This lower-cost product will allow Dish to broaden their customer base and target new broadband-only consumers who do not currently subscribe to any form of cable or satellite TV. It is interesting to note that Dish is employing a small-scale version of the traditional multichannel subscription bundling, which provides carriage fees to the TV industry for large packages of channels.

Ala-carte-style Internet TV, where users can subscribe to individual channels, has now arrived. The new age of Web-delivered TV allows viewers to have more options to pay only for the TV networks or programs they want to watch and to decide how, when, and where to watch them. Unlike pay-TV, many of these subscription-based video on demand (SVOD) providers are currently ad-free. However, ad-based OTT service may appear in future to further reduce subscription fees. As live and on-demand Internet TV programming becomes mainstream, this development will ultimately increase competition and further drive subscription prices down. It has already forced some of the biggest pay-TV providers in the United States to merge.

1.1.2 Netflix

Netflix is currently the leader of OTT providers. It is a SVOD service where regular subscribers pay a low rate of $7.99 per month and ultra-high definition (UHD) customers pay $11.99 per month. There are over 80 million Netflix-capable devices, including TVs, smartphones, tablets, and game consoles. The number of Netflix viewers passed the number of YouTube viewers. Netflix has over 57 million subscribers in 50 countries (about 39 million US subscribers and 18 million foreign subscribers) and streams over 2 billion hours of TV shows and movies per month. Roughly half of all US households now have a Netflix subscription. Netflix accounts for nearly 30% of Web traffic in the United States at peak periods, a dominant leader among all online video websites. This percentage has increased as Netflix has added 4K UHD content to its streaming video library. Such data-heavy usage is creating a huge problem for Internet service providers (ISPs), who are demanding higher fees for the interconnection required to deliver high-quality service.

1.1.3 Hulu

Hulu provides both free TV and SVOD services. It handles over 30 million online users (over 6 million are paid subscribers) and over 1 billion video streams per month. Hulu and Amazon account for 1–2% of all Web traffic during peak hours.

1.1.4 Amazon

Amazon Prime Instant Video has a few million subscribers. Six new original TV series have been launched by Amazon in 2014, including five programs that were produced in UHD format. Amazon Studios also plans to shoot its new drama and comedy series pilots in UHD, teaming up with Samsung and major media corporations including Warner Bros and Lionsgate.

1.1.5 YouTube

YouTube has the largest library of both user-generated and premium videos. The growth of YouTube is accelerating in spite of increased competition from social networks such as Facebook. Unlike Netflix, which offers full-length movies and TV shows, YouTube's short-form videos are particularly popular. These short-duration videos are perfect for on-the-go viewing on small-screen personal devices such as smartphones. Thus, YouTube dominates other online TV websites with over 20% of all mobile downstream traffic in the Unites States. Roughly the same amount of traffic is delivered over fixed networks during peak hours. This is about half of Netflix, even though YouTube has a far greater number of views and downloads. For example, in May 2011, the number of views on YouTube hit 3 billion/day. The first video posted on YouTube was a 19-s clip called *Me at the Zoo* over 10 years ago. Today, more than 300 h of video are uploaded every minute. Google's Hangouts enable virtual participation in live events where users may record and stream videos, as well as interact in conferences, music concerts, and even football matches. The free service works on any Android and Apple device and recorded videos can be broadcast on YouTube.

1.1.6 ESPN3

Unlike OTT providers, ESPN3 offers live sports viewing online. The 2012 Super Bowl attracted over 2.1 million unique viewers when the game was streamed online by ESPN3 in the United States for the first time. In that year, all 302 events of the summer Olympics were streamed live. The service is available to Internet or pay-TV subscribers from affiliated service providers who pay fees to ESPN. Since 2008, free ESPN3 service has been made available to US college/university students and military personnel.

1.1.7 HBO

HBO has been one of pay-TV's most successful products for decades but plans to break off from the cable bundle in 2015 and distribute its shows to consumers using a standalone OTT streaming service via Apple TV. It will become a direct competitor to Netflix's SVOD service.

1.1.8 CBS

CBS launched a new subscription Internet TV streaming service on October 2014 that allows people to watch its live programming and thousands of its current and past shows on demand without paying for a traditional TV subscription. The new "CBS All Access" service costs $5.99 a month. CBS is the first traditional broadcaster that makes a near-continuous live feed of its local stations available over the Internet to non-pay-TV subscribers.

1.1.9 Sony

Sony launched the world's first "Video Unlimited" UHD movie/TV streaming service on September 4, 2013. Sony has Internet rights to carry channels from Viacom, which owns cable channels such as MTV, Nickelodeon, and Comedy Central. Sony is also developing an original TV drama series that will be available initially on its PlayStation gaming consoles.

1.1.10 Retail Giants

Retail giants such as Best Buy, Sears, and Walmart are joining the digital media ecosystem. Best Buy, for example, is providing their high-dollar customers with free online video rentals from CinemaNow.

1.2 INTERNET TV DELIVERY PLATFORMS

Content distribution platforms to store, transcode, and deliver petabytes of video on commodity hardware are readily available. The TV platforms are scalable in computing, storage, and bandwidth resources even when the data centers are interconnected through the Internet and long-distance networks. The dynamic location of these service facilities and the virtualization of hardware and software elements are increasingly complemented by a flexible software-defined network (SDN) architecture, which employs programmable application interfaces to couple the control and data planes of network hardware (e.g., routers, switches) so that data can be pulled and resources can be reconfigured from any connected device on the network.

1.2.1 Cloud TV

Cloud TV leverages on the well-known concept of cloud computing to provide video storage and streaming services for content and pay-TV providers. Cloud computing is a common metaphor for computation as well as data storage, access, and management, freeing the end user from worries about the location and management of the resources they use. Some examples of cloud computing platforms include Amazon Web Services that supports companies such as Netflix and Pinterest, and Google Cloud Platform that enables developers to build and deploy applications on Google's cloud infrastructure. Big media and entertainment companies are becoming more comfortable with cloud technologies and cloud storage solutions. By putting as much business operation as possible in the cloud, service providers can operate far more efficiently with reduced capital and operating expenses due to hands-off maintenance and customers may benefit. For example, by shifting the DVR from on-premise infrastructure to the cloud, video recordings can be scheduled and viewed from anywhere. Cloud networks are complemented by file acceleration technologies that allow large files to be transmitted to cloud processing locations quickly and securely.

Cloud TV promises ubiquitous access and security without the administrative chores of managing hardware, upgrades, and backups. Scalability issues, server bottlenecks, and server failures are alleviated to a great extent. This is achieved by providing content storage and streaming on high-end servers from strategically located data centers. The emergence of cloud TV offerings will have a significant impact on both fixed and mobile communication networks. Stronger interworking and interoperability between system and network elements are needed because data centers are becoming containers of virtual provider networks. Efficient resource management in geographically distributed data centers and cloud networks to provide bandwidth guarantees and performance isolation is particularly important with the increasing reliance on bandwidth-demanding virtual machine migrations for resource consolidation.

Cloud TV is integral to enabling the flourishing electronic sell-through business that pay-TV distributors have launched, which allow consumers to purchase permanent copies of movies through cable VOD platforms. For example, Warner Bros, NBC Universal, and other studios support the UltraViolet platform, which allows consumers who purchase DVDs in retail outlets to access the movies online. More recently, Disney Movies Anywhere launched a cloud-based movie service that allows viewers who buy movies for viewing on Apple devices to view the content online. Disney Movies Anywhere has 400 titles from Disney, Pixar, and Marvel libraries.

1.2.2 Content Delivery Network

Cloud service providers normally employ a large number of servers in one location. In contrast, content delivery or distribution networks (content delivery network, CDNs) allow the servers to be distributed in different locations and works on the principle of delivering content from a server (an edge server) that is nearest to the user's location. CDNs are widely deployed in electronic commerce, media content delivery and caching, software distribution, and multiplayer immersive games. CDNs are necessary to publish rich media, including HD video and audio, because traditional Web hosts serve media from only one location, running the risks of network congestion on the public Internet. Moreover, the performance of traditional Web hosts degrades significantly as the number of simultaneous remote connections increase. Thus, there are no predictable network guarantees because the solution cannot scale globally. The CDN replicates the stream from the source at its origin server and caches the stream in many or all of its edge servers. The user that requests the stream is redirected to the closest edge server. Thus, regardless of device and location, users will experience better performance with CDNs because the media will be automatically delivered to them from a server in their location and very often, one directly connected to their ISPs.

Super points of presence (POPs) or broadcast nodes are normally built in major cities around the world, which allows the public Internet to be bypassed, thereby improving media delivery. For example, media startup times can be minimized and higher quality media streams can be supported because the shorter network span between the edge server and client reduces the number of clients accessing the server, hence increasing the available bandwidth. Moreover, caching of static content can further improve performance whereas dynamic elements of a specific application can be tuned and delivered using an application delivery network. CDN security services prevent content theft and provide access control. Many CDNs employ HTTP servers. HTTP allows standard caching (which can scale to a large number of connections) and eliminates any issues with firewalls since it uses the same ports as Web browsing. By not using proprietary video streaming protocols or specialized streaming servers, they can avoid additional capital expenditure. CDN providers are offering management software with greater control to enterprises and operators.

1.2.3 Free CDN

Netflix's Open Connect [1] is a CDN that allows ISPs to connect directly for free. ISPs can do this either by free peering with Netflix at common Internet exchanges or placing Netflix's free storage appliances in or near their network, thereby saving additional transit costs. Free peering allows voluntary interconnection of administratively

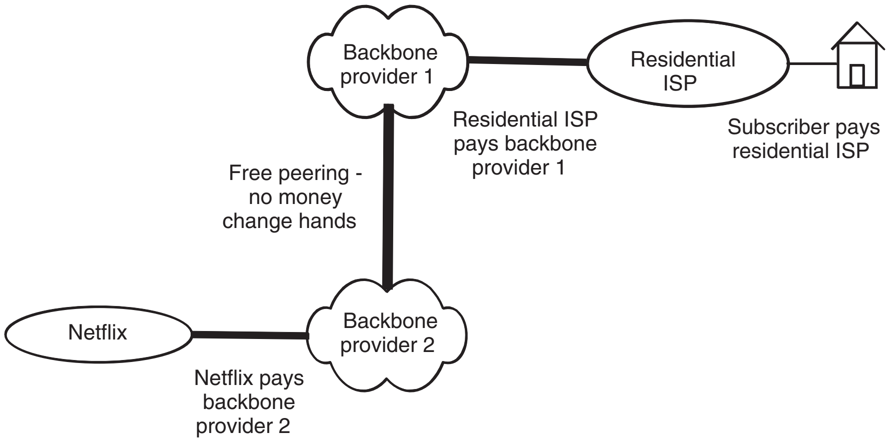

Figure 1.1 Free peering.

separate backbone Internet networks, with no money changing hands between backbone providers (Figure 1.1). Netflix also provides details of hardware design and the open source software components of the server. The Open Connect appliance hardware provides very high storage density (100 TB) and high throughput (10 Gbit/s via an optical network connection).

1.2.4 Video Transcoding

Video transcoding is equivalent to video streaming except that the output is sent to a file instead. It is common in some deployments to maintain 15 different versions of the same movie or TV program. Transcoding selects the appropriate codec and bit rate for the delivery network, a key enabler for anywhere, anytime, any device TV delivery. Software for transcoding has matured and now produces a broad range of output formats from Web and mobile standards such as MPEG-4, Windows Media, QuickTime, and Flash.

1.3 SECOND SCREEN DEVICE ADOPTION

While TVs are the traditional first screen of many residential homes, watching TV shows on second screen devices such as phones or tablets has become popular. Global smartphones sales topped one billion for the first time in 2013. Apple has sold over 1 billion iOS devices by early 2015, doubling the number 2 years ago. Currently, 10 iPhones are selling every second. According to the *New York Times* (September 29, 2013), Apple passes Coca-Cola as the most valuable brand. The rising demand for tablets is driven not only by young adults and working professionals but also kids. Many parents are buying inexpensive mini-tablets for their children. Tablets are increasingly being used in schools because the touch-based screens are a natural tool for teaching and learning. This development is far more successful than initiatives (launched several years ago) to build affordable laptops for kids.

1.3.1 Mobile Video

Video traffic constitutes a major percentage of mobile traffic, more than twice the volume of data and voice traffic. Smartphones and tablets typically allow only wireless connectivity and do not come with wired network interfaces. The proliferation of these personal devices led to pervasive mobile access. Unlike laptops, many smartphones and tablets also come with high-definition multimedia interface (HDMI) connectivity. TV programs on tablets and smartphones can also be streamed wirelessly (via Wi-Fi) to a TV using a Wi-Fi HDMI dongle that is plugged into the TV set. In essence, phones and tablets have become a DVR, portable entertainment player, TV remote, and media server conveniently integrated into one handheld device.

Verizon supports wireless OTT services and currently offers 20 TV program channels to mobile users. It expects to deliver over 100 channels and will leverage on multicast technology to enable efficient 24-h TV programming on the wireless network. Sports, news, concerts, and other live content can be viewed in real-time. An ecosystem can be created around such content that does not affect the traditional linear TV model. The mobile video service is complemented by Verizon's migration to all-IP video servers, which removes the need for STBs. This will benefit customers with faster residential service access using one piece of equipment in the home.

1.3.2 Mobile Versus Traditional TV

Although view time on mobile devices has now surpassed view time on TV, mobile TV is not a substitute for traditional TV. The most obvious difference is in the usage patterns and the length of the viewing sessions. However, tablets are like mini-TVs and are seeing much longer viewing times than smartphones. In mobile TV, audio quality and audio synchronization with the video are particularly important since this helps viewers to follow the plot in situations when image quality is degraded.

1.3.3 Over-the-Air Digital TV

Consumer desire for media content anywhere and anytime continues to stress telecom wireless networks' ability to deliver video via their one-to-one architecture. As a result, cellular operators are imposing data caps on services. These data limits, together with the size and power limitations of personal devices, place significant constraints in the delivery of high-quality videos. This has raised concerns among content providers and consumers. On the other hand, the one-to-many architecture of over-the-air digital television (DTV) has an edge in delivering video to mass audiences with few constraints. The migration of analog to DTV has led to high-quality broadcast TV transmission. The service is available on big screen HDTVs as well as mobile devices. More importantly, the broadcast service provides a good source of live sports programs and allows emergency public information to be disseminated quickly. Many consumers are turning in to free DTV service for news and sports to supplement online movies.

All HDTVs come with mandatory in-built DTV receivers. In addition, mobile DTV is gaining momentum with the ratification of the Advanced Television Standards Committee (ATSC) mobile/handheld standard in September 2009. Portable network adapters that allow DTV signals to be decoded are commercially available for phones and tablets. Due to the limited over-the-air bandwidth, DTV may not be able to support UHD transmission. Although mobile DTV remains an "after-market" phenomenon, requiring dongles and other appendages to access the service on mobile devices, consumers may prefer this free service over cellular networks.

1.3.4 Non-Real-Time TV Delivery

The ATSC has developed a standard for the delivery of non-real-time (NRT) services over legacy and mobile DTV systems. Many TV programs do not need to be delivered in real-time but can be downloaded and presented when the viewer wishes to see them. NRT service is especially attractive for mobile TV viewing, which is normally unpredictable and done on an on-demand basis. Thus, the concept of "appointment viewing" may not always be practical. NRT service enables the consumer to pick what they want to see from a menu, with the program or service preloaded on their mobile device. NRT transmissions are file-based designed to be stored in the user's device until the user chooses to access them. A recent ATSC standard allows NRT video and audio content to be encoded using H.264/AVC and high-efficiency advanced audio coding (HE-AAC) respectively.

1.3.5 NRT Use Cases

Use cases for NRT include the following:

- Clipcasting – short-form video and audio clips similar to podcasts;
- VOD – may include short-form content, sports or news programs, music videos, standard length TV programs or full-length movies;
- Micro website – offers a similar experience to browsing the Internet using a Web browser where predefined content is downloaded to user's device;
- Out-of-home content and advertising – digital signage with multiple components such as live simulcast of the mobile TV channel, news and information content, advertising;
- Mobile emergency alert system (EAS) – disaster areas, evacuation routes, harsh weather, missing persons.

1.3.6 Cable Wi-Fi Alliance

In contrast to wired networks, wireless networks operate with lower bandwidth, variable latency, and occasional disconnections that cause coverage partitions. In particular, cellular networks will need to work in tandem with wire line backhaul networks in order to overcome the bandwidth crunch associated with video transmission. A cheaper and simpler alternative is to combine Wi-Fi with a broadband wire line connection. Unlike cellular networks that employ expensive and limited licensed radio spectrum, Wi-Fi networks rely on unlicensed spectrum with broader channels that can be aggregated. The latest Wi-Fi standard (ratified as 802.11ac in December 2013) is able to support a bit rate of up to 866 Mbit/s using a single transmitting antenna. This rate is significantly higher than 4G cellular and allows Wi-Fi to support compressed UHD video streaming.

The Cable Wi-Fi Alliance [2], a consortium of US cable operators (Bright House, Cox, Optimum, Time Warner Cable, Comcast), was formed to allow broadband Internet subscribers from any operator to access over 400,000 public and residential Wi-Fi hotspots. This number is further supplemented by additional hotspots from the cable operator. For example, Comcast subscribers can access millions of Wi-Fi hotspots in public areas and new neighborhoods that are powered by Comcast gateways. Currently, 9 million of such hotspots are available in the United States. Since most user data traffic is now generated indoors (due to data caps imposed by cellular operators), these Wi-Fi hotspots provide value-added, nomadic (rather than high mobility) wireless service for fixed broadband customers.

The new Wi-Fi gateways transmit two SSIDs using two independent antennas: a private one that is used by subscribers and a public one that enables neighborhood hotspots. Custom apps from the cable operators enable their subscribers to stream video and audio from the gateways to their mobile device. These apps demonstrate how watching TV becomes more convenient when it moves to the Internet, truly enabling TV anywhere. By combining a wired broadband Internet connection with high-speed Wi-Fi, the streaming performance may easily surpass 4G. Although the Wi-Fi network coverage and reach may be more limited than 4G, users are able to enjoy consistent high-speed access even when traveling outside the home. In addition, virtually any handheld device can be connected, including cheaper devices that do not have cellular connectivity. The concept is similar to BT-FON's Wi-Fi community networks, where free Wi-Fi gateways are given to subscribers to enable shared broadband Internet service.

1.4 SCREEN AND VIDEO RESOLUTION

It is useful to distinguish between screen and video resolutions, which can be different. The screen or display resolution relates to the total number of picture elements (pixels or pels) or dots (in computer lingo) that can be packed onto the physical size of the screen or display. We shall refer to these pixels as screen pixels. The video (or image) resolution relates to the total number of pixels in a single video frame. We shall refer to these pixels as video pixels. The actual video is often scaled (or resampled) to fit the physical size of the screen. If the size of the screen is much bigger than the video resolution, the video quality may be degraded if the video is scaled to the full screen since each video pixel will require either a greater number of identical screen pixels or larger screen pixels for display. Conversely, the video quality normally remains unaffected when a high-resolution video is mapped to a small screen. However, parts of the video may be cropped. Although TV screen resolutions have improved over the years, they tend to lag behind still-frame image resolutions. For instance, the resolution of a 3840 × 2160 UHD TV is roughly two times lower than a 16 megapixel digital camera that is available in some smartphones.

1.4.1 Aspect Ratios

The original video resolution may be varied by changing the pixel aspect ratio of the screen. The pixel aspect ratio of the screen is the ratio of the length to the width of the pixel on display. A pixel aspect ratio of 1:1 corresponds to square pixels whereas a ratio of 64:45 corresponds to rectangular pixels. The screen aspect ratio is related to its pixel aspect ratio by the governing relation: screen aspect ratio = original video resolution × pixel aspect ratio of screen. For example, if the original video resolution is 720/576, a screen aspect ratio of 16/9 can be obtained by multiplying 720/576 with a pixel aspect ratio of 64/45. Thus, each video image is displayed as a grid of rectangular pixels. This is normally the case if an image is mapped to a screen whose aspect ratio is different from the image. As a result, the image may look stretched or squashed in either the horizontal or vertical directions. For example, a circle may appear like an ellipse. However, the perceived video quality may not be affected.

1.4.2 Video Resolution

A list of common video resolutions is shown in Table 1.1. The lower resolutions (quarter common intermediate format (QCIF), common intermediate format (CIF), and standard definition (SD)) have been largely superseded by high-definition (HD) resolutions. The 720p, 1080i, and 1080p HD resolutions are by far the most popular whereas the UHD-1 resolution is emerging. A UHD-1 image offers four times the number pixels and twice the horizontal and vertical resolutions of a conventional 1080p image. Similarly, the UHD-2 resolution contains four times the number

Table 1.1 Common Video Resolutions.

Format	Resolution (pixels)
UHD-2 4320p (progressive)	7680 × 4320 (16:9 aspect ratio)
UHD-1 extra wide (progressive)	5120 × 2160 (21:9 aspect ratio)
UHD-1 2160p (progressive)	3840 × 2160 (16:9 aspect ratio)
High definition (HD) 1080p (progressive)	1920 × 1080 (16:9 aspect ratio)
High definition (HD) 1080i (interlaced)	1920 × 1080 (16:9 aspect ratio)
High definition (HD) 720p (progressive)	1280 × 720 (16:9 aspect ratio)
Standard definition (SD)	720 × 480, 720 × 576
Common intermediate format (CIF)	352 × 240, 352 × 288
Quarter common intermediate format (QCIF)	176 × 120, 176 × 144

of pixels of UHD-1. The majority of TV vendors and some service providers have primarily focused on the UHD-1 resolution. The difference in UHD-1 and UHD-2 video quality may not be significant. However, the difference in UHD-1 and 1080p video quality can be significant, especially at short viewing distances. Nevertheless, video publishers continue to target resolutions under 1080p for connected TVs, tablets, and phones to maximize compatibility and playability.

HD video may contain 720 or 1080 vertical progressive scan lines (720p or 1080p), 1080 interlaced scan lines (1080i), and is capable of displaying a 16:9 image (with square pixels) and output one or more channels of digital audio. Only progressive scan lines are allowed for UHD and new video coding standards such as H.265/HEVC. Although the aspect ratio is currently specified at 16:9 by ITU-R Recommendation BT.2020, many UHD TVs support an aspect ratio of 21:9. With progressive scan, all active lines are displayed in each video frame whereas in interlaced scan, odd and even lines are displayed in successive frames at half the frame rate. Clearly, the disadvantage of interlacing is that the horizontal resolution is reduced by half, and the video is often filtered to avoid flicker and motion rendering artifacts. Interlacing relies on the phosphor persistence of a cathode ray tube (CRT) screen to blend successive frames together. Since light emitting diode (LED) TVs and liquid crystal displays (LCDs) are rapidly replacing CRT screens, interlaced content must be deinterlaced at playback time. However, a screen with poor deinterlacing can result in a jittery image. Thus, deinterlacing is sometimes performed before video encoding. The interlaced video format remains popular with broadcast and pay-TV services (480i for SD, 1080i for HD) whereas the progressive format has been widely adopted by online video portals. The frame rate is normally specified together with the scan type. For example, a 1080p video with a frame rate of 60 Hz or 60 frames per second (fps) may be specified as 1080p60. The frame rates for UHD include 24, 50, 60, and 120 Hz. The new iPhone 6 is able to capture videos at 240 Hz.

Unlike common video formats, which are usually defined in terms of the vertical resolution (e.g., 720, 1080), digital cinema formats are usually defined in terms of horizontal resolution. These resolutions are often written as multiples of a base value of 1024 pixels using a “K” or “Kilo” suffix. Thus, a 2K image will have a length of 2048 pixels. A 4K format comprises 4096×2160 pixels whereas an 8K format comprises 8192×4320 pixels. UHD follows a similar format as digital cinema with a slight modification – a minimum horizontal resolution of 3840 pixels is specified. Since 3840 is exactly two times of 1920 and 2160 is exactly two times of 1080, a UHD-1 resolution of 3840×2160 allows convenient upscaling for 1080p HD formats. Thus, although the terms 4K and UHD are often used together, the first term originates from digital cinema whereas the second term is created for HDTVs. Both Samsung and LG Electronics have demonstrated 105-in. curved 5K UHD TVs with a resolution of 5120×2160 and an aspect ratio of 21:9. The new models are extra wide and almost six times sharper than 1080p HDTVs. The increase in spatial resolution may require a corresponding increase in the frame rate due to camera panning. These bendable TVs employ moving gears behind the TV panel to shift it from a flat screen position to one that curves the sides toward the viewers.

A UHD-2 image of 7680 × 4320 pixels contains 33 million pixels or 16 times as many pixels as a 1080p image. However, the resolution is still much lower than the processing power of 126 million pixels of a single human eye. An experimental UHD-2 digital video format was developed by NHK Science and Technology Research Laboratories [3]. The frame rate is 60 Hz progressive and the bandwidth is 600 MHz, giving a bit rate that ranges from 500 Mbit/s to 6.6 Gbit/s. Several years ago, NHK demonstrated a live satellite relay over IP for display over a 450-in. (11.4 m) screen. Video was compressed from 24 Gbit/s to a rate of about 100 Mbit/s whereas 22.2 channels of surround sound audio was compressed from 28 Mbit/s to roughly 7 Mbit/s.

1.4.3 Visual Quality

In addition to the screen and video resolution, other key factors that impact video quality include the number of pixels per inch (ppi) and the quantization level used in video compression. The ppi relates to the physical spatial resolution or pixel density of the screen. The ppi metric can be applied to any screen size and tends to be lower for devices with larger screens. For example, the maximum ppi for the iPhone and the iPad are 401 and 326, respectively. In general, a smartphone or tablet may be more suited for displaying 720p than 1080p or QCIF videos. This is because the finer details of a 1080p image may not be discernible on a small screen. On the other hand, a low resolution QCIF video may appear grainy. Thus, although many tablets and smartphones can support 1080p or higher, the majority of video publishers prefer to limit the resolution to 720p or lower for optimized playback across all devices rather than the best screen resolutions possible.

The number of bits representing each pixel (i.e., the color depth) and the quantization level for compressing the video also impacts video quality. For instance, 256 color levels are available for an 8-bit pixel. A typical video codec breaks the video image into discrete blocks (e.g., 8 × 8 pixels). These blocks are digitally transformed to the frequency domain, both horizontally and vertically. The values of the resulting block are then rounded off according to the quantization level. A coarse quantization level will lead to less accurate rounding and additional loss of color information. Thus, the video quality of a decompressed HD video may suffer if the quantization level is chosen for the compression is too coarse. On the other hand, a CIF video coded with a fine quantization level can achieve decent video quality.

1.4.4 Matching Video Content to Screen Size

Just as Web pages need to be redesigned for handheld devices, matching the nature of the video content to the screen size is important for the user experience. For example, watching a 3D video in the movie theater is a much better experience than watching the same movie on a smartphone. However, watching a 30-min cartoon or comedy may work well for the smartphone. Moreover, mobile users seldom watch full-length movies on 4-in. displays but will watch trailers, music videos, news conferences, and sports highlights. Thus, short videos are perfect for informing and entertaining mobile users. Clearly, visual effects, storyline, and video content all play a role in enhancing the user experience for multiscreens.

1.5 STEREOSCOPIC 3D TV

Several stereoscopic 3D (S3D) networks have emerged following the success of James Cameron's blockbuster science fiction 3D movie *Avatar* in 2009, surpassing another movie epic *Titanic*. However, S3D programming has not gained significant traction with viewers because of the higher price of 3D TVs and the requirement that viewers wear active glasses. These glasses act like shutters that display one of two 3D images for each eye to give the perception of depth. S3D sports was supposed to be a success but due to limited viewer adoption of S3D services to the home, ESPN shut down ESPN 3D, the first 3D network, in June 2013. The cost of S3D filming is incrementally higher because it requires additional cameras and crew. More importantly, a 3D TV has a more narrow viewing range than conventional TVs and requires the viewer to sit or stand upright. Wearing glasses can be uncomfortable or inconvenient for TV viewing. Family and friends may not be able to watch a game together – everyone will have to wear glasses and only a few people will be at an optimal angle to the screen to get a good viewing experience. These glasses are fairly expensive if they need to be replaced or if additional pairs are needed. Newer displays place a device called a parallax barrier in front of an image source, such as a LCD, to allow it to show a stereoscopic image without the need for the viewer to wear 3D glasses. A drawback is that the viewer must be positioned in a well-defined spot to experience the 3D effect.

1.5.1 Autostereoscopic 3D

Autostereoscopic 3D allows S3D images to be displayed without the use of glasses. This is normally achieved by employing two or more pairs of S3D images. Glass-free autostereoscopic 3D screens are used in a limited number of TVs, game consoles, and smartphones. For instance, the HTC Evo 3D smartphone comes with an autostereoscopic 3D touchscreen. The phone's 3D camcorder captures 720p videos (in addition to 2D 1080p videos), which are stored in full-resolution temporal format for easy preview and playback. Game consoles such as Nintendo 3DS also employ autostereoscopic 3D screens. Glass-free UHD TVs are also emerging and have been demonstrated.

1.5.2 Anaglyph 3D

Anaglyph 3D requires passive viewing glasses and generally offers wide-angle 3D viewing with less visual discomfort than S3D. Unlike S3D, which requires stereoscopic displays that are synchronized with active viewing glasses, a key advantage of anaglyph 3D is the compatibility with legacy TVs and regular displays. This means one can watch anaglyph 3D videos on a smartphone using cheap passive glasses.

These glasses allow vertical or horizontal polarized light to enter the left or right eye. Alternatively, clockwise or counter clockwise circularly polarized light may enter the left or right eye. Google's Street View employs anaglyph 3D videos, allowing users to view the streets in 3D.

1.6 VIDEO CODING STANDARDS

The Joint Video Team (JVT) is a partnership between the ITU-T (SG16) Video Coding Experts Group (VCEG) and the ISO/IEC (JTC 1/SC 29/WG 11) Moving Picture Experts Group (MPEG). JVT developed the H.264 or MPEG-4 Part 10 advanced video coding (AVC) video coding standard, which is widely deployed in many online video streaming systems. It succeeds the equally popular H.262/MPEG-2 standard. The emergence of UHD video services and the ubiquity of online video streaming have created a demand for coding efficiencies superior to H.264/AVC. To this end, MPEG and VCEG worked together to develop a new H.265 or High-Efficiency Video Coding (HEVC) standard. H.265/HEVC targets to improve the coding efficiency of H.264/AVC by at least two times under the same level of video quality. A comparison of the typical video compression efficiencies of MPEG coding standards is illustrated in Figure 1.2. H.265/HEVC supports technologies that enable parallel processing of large blocks of video data. Given its hierarchical block structure and significantly increased number of coding parameters, H.265/HEVC presents many new challenges, including coding optimization, mode decision, rate-control, and hardware design.

1.6.1 Exploiting Video Content Redundancies

The key objective behind video coding is to exploit redundancies that may be present within a video frame as well as in a sequence of video frames. If such redundancies

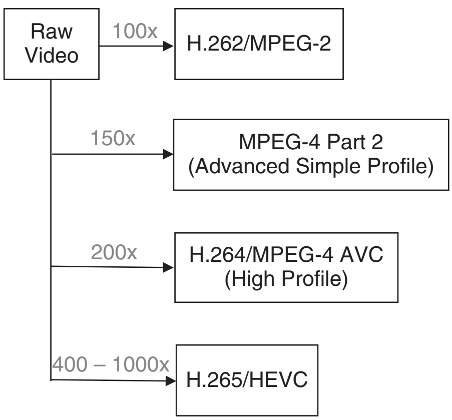

Figure 1.2 Efficiencies of video coding standards.

are present, only incremental changes are coded. For example, a pixel and its neighbors in a frame may be spatially correlated and such redundancy can be intracoded or intrapredicted. Similarly, pixels a video sequence may be temporally correlated in successive frames and such redundancy can be intercoded or interpredicted. These redundancies are encoded as residual prediction errors and are employed by the decoder to predict current frames from previously decoded frames. Motion compensation provides additional coding gains for interprediction by using motion vectors to efficiently represent any motion of objects between successive frames. The differences between the original and predicted pixels are transformed and quantized. The final step in the coding process is entropy encoding, which exploits the redundancy in data representation to further compress the binary data using the fewest bits possible. The decoder follows the same steps in reverse but is computationally less complex than the encoder (typically 5–10 times simpler). In general, a HD video may require more time to encode (compress) and decode (decompress) than a SD video because more pixels need to be processed. Similarly, a UHD video will require more compression resources than HD videos.

1.6.2 High-Quality Versus High-Resolution Videos

HD interfaces such as HDMI send decompressed or decoded videos to the display. When a video is decompressed, the overall size of the video will be the same as the original raw video. This is because each pixel in each raw and decompressed video frame is represented by a fixed value, typically ranging from 0 to 255 using 8 bits. Hence, the frame sizes will remain the same. However, the video quality of the raw and decoded videos may differ since the decoded video may have undergone a lossy compression process.

It is important to distinguish between high-quality and HD videos. A high-quality video is encoded with a fine quantization level and it need not be a HD video. This is because the video quality of a HD video can be poorer than a SD video if the HD video is quantized at a very coarse level, which leads to significant information loss. In general, we refer to videos encoded with a fine quantization level as high-quality videos. These videos typically require a higher network bandwidth for transmission and a longer time to download because more color details from the original video are retained during the compression process, especially if the video resolution is high. Thus, transporting high-quality videos over bandwidth-constrained networks (e.g., cellular networks) can be a challenge.

1.6.3 Factors Affecting Coded Video Bit Rates

Coding efficiency, video content, frame rate, and quantization all impact the video bit rate. If coarse quantization is used, the quantization parameter (QP) value becomes larger. A larger QP value leads to a higher degree of lossy encoding, which reduces video quality and bandwidth requirements. Figure 1.3 shows the bit rates for a compressed video trailer. Video trailers are ideal for streaming because the initial sections of the video are usually simple scenes, allowing the receiving device to quickly buffer

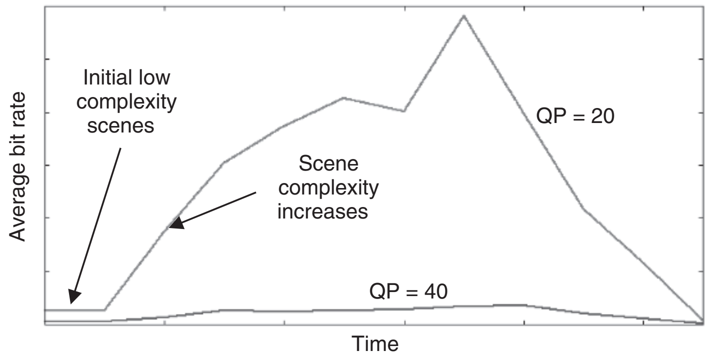

Figure 1.3 Bit rates for a compressed video trailer.

the frames for playback before more complex scenes arrive. The video bit rate variation is minimized for a QP of 40 and becomes content independent because more information is removed from each compressed video frame. As a result, the video quality may suffer, especially for more complex scenes. However, the low bit rates improve streaming performance and reduce the need for caching. Refining the quantization level using a lower QP of 20 improves the video quality but this is done at the expense of greater bit rate variation as well as significantly higher peak rates, which must be supported by the network and managed with sufficient caching at the receiver. These rates can be smoothed (or averaged) to a lower rate by buffering some video frames at the source before transmitting them, at the expense of greater delay jitter.

Many sports programs typically require higher frame rates to capture fast movements and this further increases the video bit rate. For example, a 60 Hz video requires twice the bit rate of the same video with a frame rate of 30 Hz. It is more challenging to transport UHD videos over a network. However, because consumer electronics have progressed rapidly with UHD TVs becoming available, both pay-TV and OTT providers are making UHD content available to their subscribers. To summarize, the variability of the compressed video bit rate is highly dependent on the quantization level and this should be appropriately chosen based the content type, frame rate, and the available network bandwidth.

1.6.4 Factors Affecting Coded Frame Sizes

The variation of the coded frame sizes is tightly correlated with the video resolution and content type. A higher resolution video tends to produce a broader range of coded frame sizes (Figure 1.4). For a fixed resolution, high-motion video content tends to produce a broader range of coded frame sizes (Figure 1.5). This implies that the amount of caching for the coded frames can be difficult to predict for high-resolution and high-motion videos. There is however, no correlation between the duration of the video and the frame size distribution. Thus, a longer video may not imply a higher frame size variability compared to a shorter video with the same resolution.

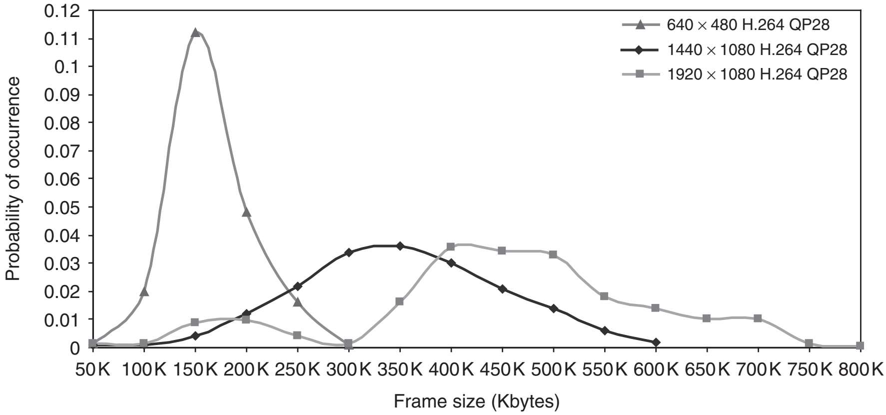

Figure 1.4 Coded frame sizes for videos with different resolutions.

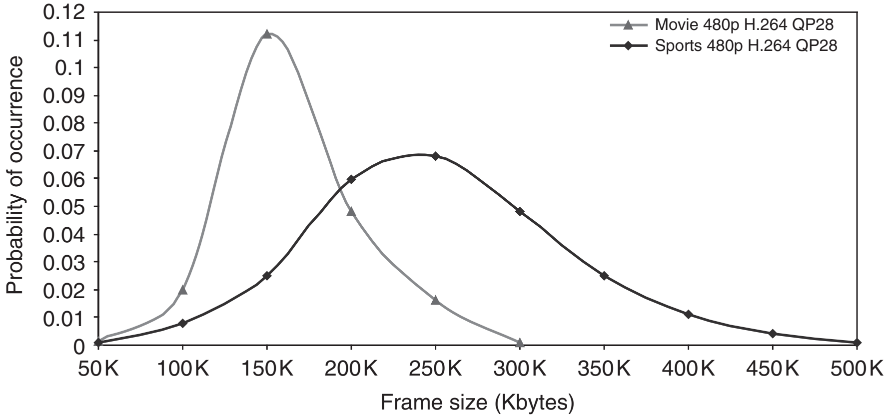

Figure 1.5 Coded frame sizes for videos with different content.

1.7 VIDEO STREAMING PROTOCOLS

Streaming video entertainment is the most efficient and convenient way compared to stored media using DVDs and old video home system (VHS) tapes. Streaming protocols determine how well compressed video can be transported across the network and the memory requirements at the user device. They simplify digital rights management (DRM) or content protection since small pieces of the video content are continuously delivered to the user device. In addition, the content is cached at the user device for a short period of time, thereby reducing the risk of content piracy for the entire video or movie. This makes it more acceptable to stream a video directly to a personal device (such as a PC, phone, tablet) and not via a proprietary STB, which employs hardware controls and encryption to prevent someone from reusing the video content. While streaming is the most popular mode of watching video, implementing trick modes (i.e., DVR functions such as fast forward and replay) can be challenging due to the network propagation delay. There are situations where video downloads may be more appropriate (e.g., downloading movies to watch on a plane). DVDs provide the best video quality, which is why they are still in demand in spite of the popularity of streaming.

1.7.1 Video Streaming over HTTP

Real-time Transport Protocol (RTP) over User Datagram Protocol (UDP) is a unidirectional IP-based network protocol designed for real-time multimedia traffic. Because UDP is a connectionless protocol that may result in packet losses and has difficulty passing through firewalls, it is normally employed by private managed networks (e.g., pay-TV cable networks) where packet losses are low. On the other hand, online video streaming is typically achieved using HTTP, a two-way connected-oriented network protocol that resends any data packet that is corrupted or dropped during transit. Thus, HTTP guarantees that all data will eventually be delivered correctly to the destination. Highly compressed video is sensitive to information loss, especially random packet losses caused by network congestion in the public Internet. Thus, loss-free protocols are more desirable and practical for video streaming over the Internet, although the end-to-end delay has to be calibrated properly for live video service. HTTP is a Web delivery standard that has become ubiquitous because it can work with standard Web servers and existing infrastructure, including CDNs, caches, firewalls, and NATs. The protocol can scale to a large number of media streams and is driven by clients that send GET requests to the servers and URLs that define the locations of the media data. Enhanced security for streaming can be achieved via secure protocols such as HTTPS.

1.7.2 Adaptive Bit Rate Streaming

HTTP's congestion control and reliability may overly burden real-time video streaming whereas UDP's nonreliable connections can affect user experience. Adaptive streaming is able to minimize the shortcomings of these protocols by dynamically adapting the video quality to the network bandwidth, connection quality, and device capabilities, thereby maintaining smooth playback. For example, when the client device detects network congestion or low bandwidth at any instant in time, it can request the server to deliver the video content at a reduced bit rate by selecting a stream with lower video quality. This prevents choppy playback when video frames do not arrive on time at the receiver (due to buffering at the sender) as well as serious video artifacts such as image breakup (due to packet losses). Thus, playback interruption is minimized when a client's network speed cannot support the quality of the video.

Clearly, the adaptive streaming process may increase the latency because the server and the player will need to switch to a lower quality video and make the necessary adjustments. Depending on the severity and frequency of the network congestion, this latency can sometimes exceed several seconds. However, because the coded video content is segmented into smaller chunks, it may be faster to switch chunks than change video coding parameters. Unlike peer-to-peer video streaming (e.g., Skype), which may consume large amounts of bandwidth in both inbound and outbound channels regardless of network conditions, adaptive streaming is more network friendly and generally consumes video bandwidth on the inbound. Adaptive streaming employs HTTP and is originally pioneered by Move Networks for online streaming of sports events. Apple, Microsoft, and Adobe have also developed their own approaches with new operating systems that support adaptive streaming and a compatible media player.

1.7.3 Benefits and Drawbacks of Adaptive Streaming

Adaptive streaming obviates users having to make choices between high- or low-quality video streaming and to adjust the video quality when switching from small to full screen mode. It also simplifies and reduces the cost of multiscreen (e.g., phone, tablet, PC, TV) deployment for service providers. The main drawback of this approach is a noticeable change in video quality as well as reinforced artifacts whenever there is a sudden degradation or improvement in network bandwidth or connection quality. If such instances occur frequently, it may compromise end-user quality of experience. On a private managed network, bandwidth availability and connection quality are more reliable. However, the use of adaptive streaming can still help maintain good video quality while smoothing the peaks that are associated with variable bit rate videos. For example, peak rates due to complex or high-motion scenes can be lowered by switching to a lower video quality for a short period of time. Thus, precise bandwidth dimensioning is not needed when adaptive streaming is applied to managed networks.

1.7.4 HTTP Progressive Download

Adaptive streaming relies on HTTP's rate control mechanism to determine the quickest sending rate and the appropriate level of video quality. To achieve fine-grained control of the video transmission, a progressive download mechanism based on HTTP can be employed. In this case, the user device may begin playback of the media before the download is complete. Low bit rate or short-form videos tend to be smaller in size and rather than streaming the videos continuously or changing the video quality levels, they can be downloaded instead when a broadband connection is available. Progressive download is widely used for the download and playback of short-form videos (e.g., YouTube) with a higher burst at the beginning to reduce the startup delay for playback. However, this method of video transmission may not allow the video to be fast forwarded quickly.

1.7.5 HTML5

HTML5 adds many new features such as audio, video, and canvas elements as well as support for scalable vector graphics (SVG). These features are designed to make it easy to add and handle multimedia and graphical Web content without relying on proprietary plugins and APIs. The canvas element enables dynamic, scriptable rendering of 2D shapes and bitmap images. Unlike canvas, which is raster-based, SVG is vector-based where each drawn shape is represented as an object and then rendered to a bitmap. This means that if attributes of an SVG object are changed, the browser can automatically re-render the scene. A canvas object would need to be redrawn.

1.8 TV INTERFACES AND NAVIGATION

Accessing digital movies should be as simple as flipping a channel. The STB has been the bedrock of cable and IPTV operators for over two decades. The box scans for available TV channels and allows the user to select the desired channels via a TV remote. A tablet may work better as a user and social TV interface than a remote. Several years ago, Comcast developed an app to allow tablets to act as a TV remote and communicate with the STBs. The app connects the tablet to the STB using the enhanced TV binary exchange format (EBIF), the cable industry's specification for delivering one-to-one interactive applications to STBs. This pioneering personalized remote allows users to view channels, browse the full channel lineup, and invite friends to watch movies. With the emergence of Internet TV, STBs have been gradually replaced by video streaming boxes and HDMI adapters that enable connections to HDTVs. Using a broadband connection, media entertainment can be streamed via Wi-Fi using these devices. There are a variety of streaming boxes (e.g., Apple TV, Roku, Amazon Fire TV) and adapters (e.g., Chromecast, Roku). These devices also allow smartphones and tablets to stream content to the TVs.

1.8.1 Streaming Adapters

With a thumb-sized Chromecast HDMI dongle plugged into a TV, online entertainment (e.g., movies, TV shows, and music from Netflix, YouTube, HBO GO, Hulu, Google Play Movies and Music, Pandora, Vudu, Crackle, Rdio) and any content from the Google Chrome Web browser can be streamed wirelessly from any Wi-Fi device to the TV. Users select the media to play using Chromecast-enabled mobile apps and Web apps, or via a beta feature called "tab casting" that can mirror content from the Chrome browser tab. The apps also allow users to control playback and volume. Android and Apple mobile apps are supported using an extension. Chromecast employs a system-on-a-chip, which includes video codecs for hardware decoding of VP8 and H.264/AVC. Like Chromecast, which conserves space, the Roku HDMI adapter provides access to Netflix, Hulu, Amazon Instant Video, ESPN, HBO Go, Showtime, YouTube, Redbox Instant, among more than 1000 music, movie, TV, and sports channels.

1.8.2 Streaming Boxes

The Apple TV streaming box provides Netflix, Hulu, ESPN, YouTube, Vevo, Vimeo, MLB.TV, NBA League Pass, NHL GameCenter, along with any music or videos purchased through Apple's iTunes. AirPlay allows users to broadcast anything on their iOS device to the HDTV through the box. Users can also play games with dual screens, leveraging the TV and either an iPhone, iPad or iPod Touch. The Roku streaming box provides access to more than 1000 music, movie, TV and sports channels, including Netflix, Hulu, Amazon Instant Video, ESPN, HBO Go, Showtime, YouTube, and Redbox Instant. The remote includes a headphone jack for private listening, and features games including Angry Birds and classics such as Pac-Man or Galaga. Amazon has launched a powerful Fire TV streaming box with a voice-activated remote, providing instant access to over 200,000 TV episodes and movies, plus all subscriptions and streaming services, including Amazon Instant Video, Netflix, Hulu, ESPN, Vevo, Bloomberg, and Showtime.

1.8.3 Media-Activated TV Navigation

The last digital island, the TV, has joined the PC, tablet, and smartphone as an Internet-connected device. The broadband TV brings all of the Internet to the TV, including the full range of video websites. This development has a profound impact on the distribution and consumption of digital media. Media-activation can help improve the navigation, selection, and playback of videos, photos, and music. As an example, Google Talk personalizes the Web search engine with speech input from the user. Google also provides a voice-recognition application that converts speech commands to actions via devices with speech input such as a smartphone. Media-activated communications can also be nonverbal. For example, popular immersive games employ video sensors on game consoles to recognize human body movement while hand gestures can also be analyzed to navigate TV or video thumbnails.

1.8.4 Smartphone and Tablet TV Navigation

Smartphones and tablets can serve as TV remotes (more sophisticated but better), allowing the user to browse for video content, control playback, and adjust volume. Such a capability is more flexible than the Internet-connected TV, where users can only watch online content from preprogrammed video portals. In addition, navigation using a TV remote is less intuitive than using a phone or tablet with a familiar keyboard. Thus, the widespread deployment of Internet TV may eventually remove the traditional remote as well as Internet/cable connections for the TV, which may potentially evolve to become a display-only device.

1.8.5 Digital Living Network Alliance

The Digital Living Network Alliance (DLNA) [4] has developed open technical specifications to ensure interoperability among TVs, STBs, Blu-ray players, and other mobile devices (e.g., tablets, phones) so that these devices can seamlessly stream multimedia content to each other over point-to-point wireless (e.g., 802.11) or wired (e.g., Ethernet, MoCA) connections. DLNA employs the Universal Plug and Play

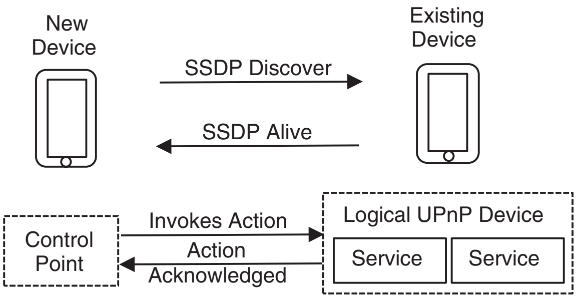

Figure 1.6 SSDP discover message flow.

(UPnP) networking protocols that enables networked devices to seamlessly discover each other on the network and establish functional network services for media sharing and communications. UPnP is the network analogy of plug-and-play.

The Simple Service Discovery Protocol (SSDP) forms the basis of the discovery protocol of UPnP. SSDP is a network protocol that allows HTTP clients and services to discover each other without depending on server-based configuration mechanisms, such as the Dynamic Host Configuration Protocol (DHCP) or the Domain Name System (DNS), or relying on static configuration of a network device. SSDP defines the communication protocols between control points and devices that provide services. A UPnP device may contain both functions. A typical message flow for discovering services between UPnP devices is shown in Figure 1.6.

1.8.6 Discovery and Launch

Discovery and Launch (DIAL) is a mechanism for discovering and launching applications on a local area network such as a home network. It enables second screen devices (e.g., phones, tablets) to send content to first screen devices (e.g., TVs, Blu-ray players, streaming boxes and adapters). DIAL is a protocol co-developed by Netflix and YouTube with help from Sony and Samsung. It is used by Chromecast and relies on UPnP.

1.8.7 UltraViolet

Unlike DVDs, which play on any DVD device, consumers are not comfortable with limited usage of online video content. For example, videos that are compatible with Flash may not be compatible with Windows Media or Apple QuickTime. Leading entertainment and consumer-electronics companies have therefore formed a consortium, the UltraViolet to develop technical specifications that content distributors and manufacturers can follow to ensure that consumers are not locked to a specific platform. The goal is to let consumers know that content and devices carrying a special logo will play seamlessly with one another.

REFERENCES

1. Netflix Open Connect, https://openconnect.netflix.com.
2. Cable Wi-Fi Alliance, http://www.cablewifi.com.
3. NHK, http://www.nhk.or.jp/strl/english.
4. Digital Living Network Alliance, http://www.dlna.org.
5. UltraViolet, https://www.myuv.com.

HOMEWORK PROBLEMS

1.1. Explain whether user interactivity, ease of use, content sharing, and video quality are important differentiators between online and pay-TV. Will the age of the consumer play a role in your assessment?

1.2. Explain whether Wi-Fi or DLNA enabled HDTVs are better than the use of HDMI streaming adapters such as Google's Chromecast. State-of-art STBs are becoming thumb-sized. Will they compete directly with streaming adapters?

1.3. Will autostereoscopic 3D require the viewer to sit or stand upright? Will it provide wider viewing angles than stereoscopic 3D?

1.4. Consider the following streaming method that sends video frames at a constant frame rate. Instead of sending a video at a bit rate that is based on the file size, we now send the video at the same frame rate that the video is meant to be played at. Typical frame rates are 30 and 60 Hz. A few frames are buffered by the receiver before they are played back. Will this method show good performance for constant frame rate videos? Consider another situation when videos are encoded with variable frame rates where fast moving scenes (such as action scenes) require more frames to be generated by the encoder over the same time interval. How will such variable frame rate videos impact the performance of the constant frame rate streaming method as outlined earlier?

1.5. Streaming a video at a constant frame rate may achieve good performance for constant bit rate videos. For such videos, the average bit rate for each frame remains more or less fixed and the natural frame rate of the video is held constant. This is achieved by adjusting the quantization level for every encoded frame to produce a constant output bit rate. With variable bit rate videos, both the frame rate and the quantization level (hence video quality) are held sconstant for all frames and so fast changing or complex scenes will require more bits per frame. Explain whether video streaming at a constant frame rate will achieve good performance for variable bit rate videos. Note that constant bit rate videos may not imply a constant sending rate at the source since there may be instances of network congestion or poor quality links that mandate the source to buffer or retransmit some of the video frames.

1.6. An alternative to adaptive streaming is to maintain an acceptable video quality suitable for the subscribed network bandwidth by encoding the video as efficiently as possible to ensure smooth playback. How will network congestion (that lead to packet dropping) impact adaptive and fixed video quality streaming? Can constant bit rate video encoding replace adaptive streaming or can they work together?

1.7. Will the new iPad be able to display UHD-1 videos? What is the best resolution for video display on a smartphone?

1.8. Suppose the same movie is recorded in analog VHS and digital DVD formats. Which format will generally produce better video quality? Does quantization play a role in the video quality? How do they compare to the video quality of pay-TV and online TV distribution? Note that in analog television, there are roughly 480 active lines with each line holding about 440 pixels. Thus, each video frame has slightly more than 200,000 color pixels. This should be compared with the resolution of the SD video, which has 720×480 or 345,600 pixels/frame.

1.9. In some US cities, high-speed fiber-optic Internet and TV access to the home is readily available. Will this development remove the need for video compression?

1.10. If adaptive streaming is implemented over managed pay-TV networks, will the user still be able to record all or part of the video on a DVR?

1.11. Will adaptive streaming lead to lower buffering and faster start times for the client device when compared to traditional constant video quality streaming?

1.12. When adaptive streaming starts, the client may request chunks from the lowest bit rate stream. If the client discovers that the download speed is greater than the bit rate of the downloaded chunk, it will request higher bit rate chunks. If the download speed deteriorates later, it will request a lower bit rate chunk. Suppose the highest bit rate stream is selected at the start instead. Discuss the tradeoffs for doing this.

1.13. What are the benefits of using fixed size versus variable size chunks in adaptive streaming? Note that ultimately, these chunks are transported by network technologies such as Ethernet and IP that employ variable size packets.

1.14. Discuss the roles of the following methods to improve video transport over a network: (a) buffering at the source, (b) buffering at the receiver, (c) segmentation of compressed video into chunks, and (d) encoding the video at constant bit rate.

1.15. Which of the following achieves the best video quality when displaying a 1080p video: (a) smartphone, (b) tablet, (c) HDTV, (d) UHD TV, and (e) giant TV screen at a football stadium? Note that the viewing distance is a key factor that affects perceived video quality.

1.16. Why do consumers still prefer Netflix even though the service does not offer live sports, unlike Web portals from service providers?

1.17. Will the traditional remote enable more effective discovery of video content than a personal device such as a phone or tablet?

1.18. Analyze why Internet protocols (e.g., Real-time Transport Streaming Protocol (RTSP), Real-time Transport Protocol, (RTP), Real-time Transport Control Protocol, (RTCP)) that are designed with real-time guarantees for transporting multimedia traffic may not perform better than HTTP. RTP facilitates packet resequencing and loss detection using sequence numbering of packets, synchronization between multiple flows using the RTP timestamp and RTCP Sender Reports, and identification of the payload type. Evaluate the advantages of using RTP for video streaming over the public Internet.

1.19. Suppose a user employs a tablet to video chat with a friend who is using a smartphone. Is adaptive streaming useful for such an application? Suppose the same user now connects the tablet to an OTT provider to watch a movie. Explain whether the streaming requirements for this application will be different to the video chat application. Note that the first video application is an example of human–human communications (a live application, possibly interactive) whereas the second video application is an example of machine–human communications (an on-demand application). Are there video applications that involve machine–machine communications?

1.20. Periodic traffic is broadly defined as a traffic type where the average bit rate can be determined. This is in contrast to bursty traffic where the average rate is unpredictable and may therefore cause packet losses due to network congestion. Such congestion is often solved by adding buffers, at the expense of increased end-to-end latency. Delay jitter is a key metric that determines the useful playback lifetime of periodic traffic whereas fairness in network bandwidth consumption is a key metric for bursty traffic. Explain whether adaptive video traffic should be classified as periodic or bursty traffic.

1.21. Many consumers will watch the best available content on the best available screen at the best available video quality and price. Rank these four criteria in the order of importance and justify your answer.

1.22. Cloud gaming requires graphics steaming, where game objects are represented by 3D models and textures, and are streamed to the players' devices, which perform rendering of the scenes. Alternatively, video streaming can be employed, where the cloud not only executes the game logic but also the game rendering, and streams the resulting game scene to the players' devices as video. Evaluate the strengths and weaknesses of each of these methods by balancing bandwidth and delay limitations with wider accessibility and possibility to run the game on thin clients.

1.23. One reason why music piracy is more prevalent than video piracy is because compressed audio requires much smaller storage space and bandwidth requirements. Will next-generation video codecs such as H.265/HEVC with vastly improved coding efficiency create DRM problems in video distribution?

1.24. Suppose a compressed video quantized with QP 28 is decoded to raw video and then compressed again using QP 28. Will the original and resulting file sizes and video quality match up?

1.25. Compute the pixel aspect ratio to convert an old 4:3 video to 16:9. Compute the screen dimensions of a 401 ppi smartphone to display a 1080p video using square pixels.