Multimedia Contents for Mobile Entertainment

Hong Yan
City University of Hong Kong, Hong Kong
University of Sydney, Australia

Lara Wang
Tongji University, China

Yang Ye
Tongji University, China

INTRODUCTION

Electronic mobile devices are becoming more and more powerful in terms of memory size, computational speed, and color display quality. These devices can now perform many multimedia functions, including rendering text, sound, images, color graphics, video, and animation. They can provide users with great entertainment values, which were only possible with more expensive and bulky equipment before.

Technological advances in computer and telecommunications networks have also made mobile devices more useful for information exchange among users. As a result, a number of new mobile products and services, such as multimedia messages services (MMSs) and online games, can be offered to users by the industry.

It is commonly believed that "contents are the king" for multimedia products and services. As the mobile handsets and networks become more and more advanced, there is a stronger and stronger demand for high-quality multimedia contents to be used in new hardware systems. Content creation involves both computer technology and artistic creativity. Due to a large number of users, mobile entertainment has become an important part of the so-called creative industry that is booming in many countries.

In this article, we provide an overview of multimedia contents for mobile entertainment applications. The objective is for the readers to become familiar with basic multimedia technology, and the concepts and commonly used terms. Our focus will be on multimedia signal representation, processing, and standards.

SOUND

The original sound, such as speech from humans, can be represented as a continuous function $s(t)$. The sound can be recorded using electronic devices and stored on magnetic tapes. It can also be transmitted through telecommunications systems. The traditional telephone is used to send and receive waveform information of sound signals. The function $s(t)$ here is called an analog audio signal. This signal can be sampled every T seconds, or $f_s = 1/T$ times per second. The output signal is called a discrete-time signal. Each data sample in a discrete-time signal can have an arbitrary value. The sample values can be quantized so that we can represent them using a limited number of bits in the computer. These two processes, sampling and quantization, are called digitization, which can be achieved using an analog-to-digital converter (ADC) (Chapman & Chapman, 2004; Gonzalez & Woods, 2002; Mandal, 2003; Ohm, 2004). A signal $s(n)$ that is discrete both in time and in amplitude is called a digital signal. To render a digital sound signal, we must use a digital-to-analog converter (DAC) and send the analog signal to an electronic speaker.

The parameter $f_s = 1/T$ is called the sampling frequency. For telephone applications, usually $f_s = 8000$Hz, and for audio CD usually $f_s = 44100$Hz. To achieve stereo effects, the audio CD has two channels of data. There is only one channel in telephone applications. The sampling frequency f_s must be greater or equal to twice the signal bandwidth in order to reconstruct or recover the analog signal correctly. This is called the Nyquist sampling criterion. In telephone systems, a sound signal may have to be filtered to remove high-frequency components so that the sampling criterion is satisfied. This is why audio CDs have a higher sound quality than telephones.

The sound information can be stored as raw digital data. Some sound files, such as WAVE files (with extension ".wav") used on PCs, store raw sound data. This kind of format requires large storage space but has the advantage, since there is no information loss and the data can be accessed easily and quickly. To reduce the amount of data for storage and transmission, sound data are often compressed. Commonly used sound data compression methods include the m-law transformation, adaptive differential pulse code modulation (ADPCM), and Moving Picture Experts Group (MPEG) audio compression (Chapman & Chapman, 2004; Mandal, 2003; MPEG, n.d.; Ohm 2004).

Currently, MPEG Audio Layer 3 (MP3) is a very popular sound data compression technique for mobile devices. MP3 is also a well-known sound file format. During data compression in MP3, a sound signal is decomposed into 32 frequency bands, and psychoacoustic models are used to determine the masking level and bit allocation pattern for each band. Modified discrete cosine transform (MDCT) is used to compress the data. The discrete cosine transform (DCT) has the so-called energy compact property—that is, it is able to pack most of the energy of a signal in a small number of DCT coefficients. For example, if $s(0) = 2$, $s(1) = 5$, $s(2) = 7$, and $s(3) = 6$, then the DCT coefficients are $S(0) = 10$, $S(1) = 3.15$, $S(2) = 2$, and $S(3) = 0.22$ (Mandal, 2003). In this case, from $S(0)$ to $S(3)$, the coefficients become smaller and smaller. We can simply retain the low-frequency components $S(0)$ and $S(1)$ and discard high-frequency ones $S(2)$ and $S(3)$ to obtain an approximation of the original signal. In practice, we can allocate more bits to code $S(0)$ and less and less bits for $S(1)$ to $S(3)$ to achieve a high compression ratio and at the same time maintain good signal quality.

Short sound files can be completely downloaded to mobile devices before being played. To reduce waiting time for downloading long sound files, audio streaming technology can be used (Austerberry 2005). In a streaming system, audio data is transmitted through a network and played by a mobile device as the data become available. That is, a sound does not have to be completely stored in the mobile device before being played. Steaming is useful if a large amount of data need to be received by a mobile device or live broadcasting is required.

We have focused on how to process sound waveform information above. In fact, sound can also be generated according to its parameters or a set of instructions. The musical instrument digital interface (MIDI) standard is used for such purpose (Chapman & Chapman, 2004; Mandal, 2003). A MIDI file contains information on what kind of instruments, such as different types of pianos, should be used and how they should be played. Several instruments can be arranged in different channels and played at the same time. MIDI files are much smaller than waveform-based sound files for music and is widely used for ring tones on mobile phones.

IMAGES

The imaging ability of mobile devices has been improved rapidly in recent years. Now most new mobile phones are equipped with digital cameras and can take pictures with millions of pixels. Software programs are available to edit an image, such as to enhance its contrast and change its color appearance. Mobile devices can also be used to send or receive images and browse the Web.

An important parameter for images is the resolution, which is closely related but should not be confused with the image size. Image resolution is usually measured by dots per inch (dpi). Higher resolution for the same physical area of an object would generate a larger image than lower resolution, but a large image does not necessarily mean a high resolution as it depends on the area that the image covers. The concept of resolution is often used in image printing and scanning. Typically, laser printers have a resolution of 300dpi or 600dpi and fax documents have resolutions from 100dpi to 300dpi.

A digital image can be considered as a two-dimensional (2D) discrete function $i(x, y)$. Like sound data, images need to be compressed to save storage space and transmission time. There are a number of methods that can be used to compress an image. Most methods are designed to reduce the spatial redundancy so that an image can be represented using a smaller amount of data. The most popular technique used for image compression is the Joint Photographic Experts Group (JPEG) standard (Gonzalez & Woods 2002). In JPEG-based compression, an image is divided into small, square blocks, and each block is transformed using the two-dimensional DCT (2D-DCT). Similar to audio data compression, the energy of a smooth image is concentrated in low-frequency components of the 2D-DCT coefficients. By allocating more bits to a small number of low-frequency components than to a large number of high-frequency components, we can achieve effective data compression.

In the JPEG2000 standard, the 2D discrete wavelet transform (2D-DWT) is used for image compression. In this method, an image is decomposed into several sub-bands, and different quantization schemes are used for different sub-bands. The 2D-DWT usually performs better—that is, it can provide a higher quality for similar compression ratio or a higher compression ratio for similar image quality than the 2D-DCT.

The JPEG and JPEG200 standards are usually used to provide lossy compression with a high compression ratio, although they can also be used for lossless compression. In lossy compression, the decompressed image is only an approximation of the original one, and as a result the image may appear to be blocky and blurred. The quality of an image from lossy compression can be improved using a number of techniques (Liew & Yan, 2004; Weerasingher, Liew, & Yan, 2002; Zou & Yan, 2005). In lossless compression, we can reconstruct or recover the original image exactly. Commonly used lossless compression methods include graphic interchange format (GIF), tagged image file format (TIFF), and portable network graphics (PNG). A GIF image can only show 256 colors and is especially useful for logos, buttons, borders, and simple animation on Web pages. GIF is a patented technology. In TIFF, image information is associated with different tags, which can be defined by users. TIFF is widely used for scanned office documents. PNG, which can support true colors, has been developed to be a royalty-free alternative to GIF. Similar to GIF, PNG

Related Content