ISSN: 1411-1284

Audio Video Processing on Tapeless On-Air Television System

Dodi Sudiana* and Mohamat Sudiantoro*

* Dept. of Electrical Engineering, Fac. of Engineering, University of Indonesia,
Kampus Baru UI, Depok 16424
Tel. +62-21-727-0078, fax. +62-21-727-0077 email : dodi@ee.ui.ac.id

Abstract-**Digital on-air television system has been rapidly developed in the digital-era nowadays. The analog TV system was broadcasted in 1941 for the first time (B/W, NTSC), followed by colour NTSC in 1953. PAL and SECAM were introduced in 1960 and in 1982 broadcast engineers developed the digital video for the first time. HDTV as a new television standard was initiated in 1990, MPEG1, JPEG, DBV and MPEG2 are launched in 1991, 1992 and 1994, respectively. Up to now these system are still in development and new technologies grow rapidly. In Indonesia, most of TV stations have recorded video in analog format and converted them to digital, however, the broadcasting system transmits the video in analog format again due to the receiving TV systems belong to the audiences. Since the video is recorded and edited on magnetic tapes, it could be damaged due to scratch, dirt, fungus, and other physical disturbances. Therefore, the quality of both video and audio will be degraded. To solve this problem, the analog audio and video signal could be converted to digital and compressed to MPEG2 format to save the allocation space. The whole editing system is performed using computers and no tapes are involved (tapeless system). This paper will discuss the conversion of analog audio and video signals into digital format and evaluate its performance in a real video production on a tapeless on-air television broadcasting system.**

Keyword: *video audio processing, tapeless on-air television system, MPEG2*

I. INTRODUCTION

The advanced technology of digital audio-video processing has been changed the television broadcasting industry. TV stations have to attract more audiences in very tight competition and TV programs have to be zero mistaken. One of the solution to have high quality in audio and video broadcasting is to convert the analog data to digital format, eventhough the broadcasting is still in analog due to the widely used analog TV receiving systems belong to the audience. Since all of recordings are converted to digital format, they are stored in computers storage, so the operation cost can be reduced because there are no video tapes needed. Previously, the magnetic tape in Betacam format is the major recording storage in many TV broadcasting companies. There is a major problem in using magnetic tapes, i.e. the tape player/recorder. Mechanical rotation in the tape and direct contact between the head and the tape could be resulted in scratches on the tape surfaces. Tropical high temperature and humidity could also introduced fungus in the tape. Stacks due to mechanical problem in the player/recorder could also cut the magnetic tape itself. Since all of these problems could be occured anytime, the production cost will rise and the risk of broadcasting failure, which is the most fatal problem, will be worst. More vulnerable digital audio/video format will solve these problems, although still has some weakness in storage system.

Digital on air and *tapeless TV production* have been developed as a new trend in the world TV broadcasting industries nowadays, including Indonesia. Several TV stations have applied this system in their production lines. Using digital communication, *broadband network* and *digital storage*, the efficiency in space, time, cost and high quality Audio/Video should be achieved in the production process. New broadcasting companies in TV industry could also be built with less initial cost.

This paper will analyze the digital audio/video conversion process and the performance of digital on-air and tapeless TV production system in term of quality of data and efficiency in time and storage space. The sample data was taken from one of TV broadcasting company in Jakarta.

II. BASIC THEORY

Image is a 2-dimensional data, which is formed from the light reflectance of an object captured by an imager. The imager could be human eyes, camera or scanners. The intensity of light in the image could be expressed as $f(x,y)$, where:.

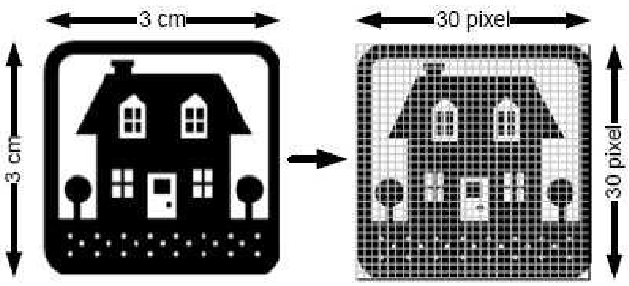
(x,y) : cartesian-coordinate in 2-dimension

$f(x,y)$: light intensity at point (x,y)

The intensity function is a multiplication of the light received $i(x,y)$ and the reflectivity of the object $r(x,y)$. Light has color characteristic which is described as color temperature. In a clear day, the sun gives light energy of 9000 fc (footcandle), 1000 fc in a cloudy day and 0,01 fc of a full moon in the night. The reflectivity r(x,y) depends on the capability of the object reflects the signal, which is 0 for a perfect absorbance objects and 1 for perfect reflective objects. The intensity *f(x,y)* of a black and white picture is stated as gray level.

$$i_{min} < f < i_{max} \quad (1)$$

Color images is spectral images, which consist of three basic color components per picture point (pixel). Color intensity at a pixel point is a composite of the three color intensities. In a TV system, color addition system is used to compose any color, which is red, green and blue. Other system is subtractive color mixing which is widely used in printing and drawing. The pigment colors are: yellow, red and blue. The digital imageries are converted from analog images through sampling and quantization. For example, an analog image in 3×3 cm size is sampled to 30×30 pixels which is stored as a matrix in the computer memory as shown in Fig. 1. The more detailed the sample is taken, the higher resolution of digital image is achieved. Nowadays, the sampling rate of analog imageries acquired using commercially available camera have reached 5 Mega pixel and still increasing for various size of images. Therefore, the quality of digital images almost the same as the analog ones.



(a) (b)

Figure 1. a. Example of an analog image, b. Sampling process of an analog image with a matrix 30×30 size

TV cameras record moving objects through the lenses and a prism block to refract into red, green and blue, convert the light energy into electrical signals using the Charge Coupled Display (CCD) as shown in Fig. 2. The output of CCD signal is analog and modulated by an encoder.

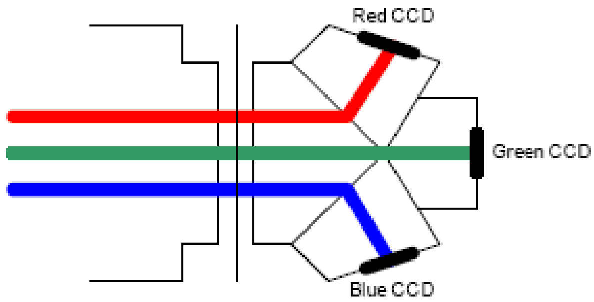


Fig. 2. Prism block and the CCD

There are 3 color systems that generally used in the world TV industry: Phase Alternating Line (PAL), National Television System cCommittee (NTSC) and SECAM (Séquentiel couleur à mémoire, French which means sequenced color with memory) as shown in Fig. 3. PAL has 625/50 Hz scanning line per frame, 25 frame per second and bandwidth 5,5 MHz, while NTSC has 525/60Hz scanning line per frame, 30 frame per second and 4,5 MHz bandwidth.

Indonesia, Australia and most of countries in Asia have adopted PAL system. NTSC is used in the US, while SECAM mostly used in the European Union. In the PAL coloring system, the red, green and blue signals from CCD are converted to YUV matrix. The block diagram of PAL encoder system is shown in Fig. 4. NTSC and SECAM signals are converted to Luminance-Chrominance and YdbDr matrices, respectively.

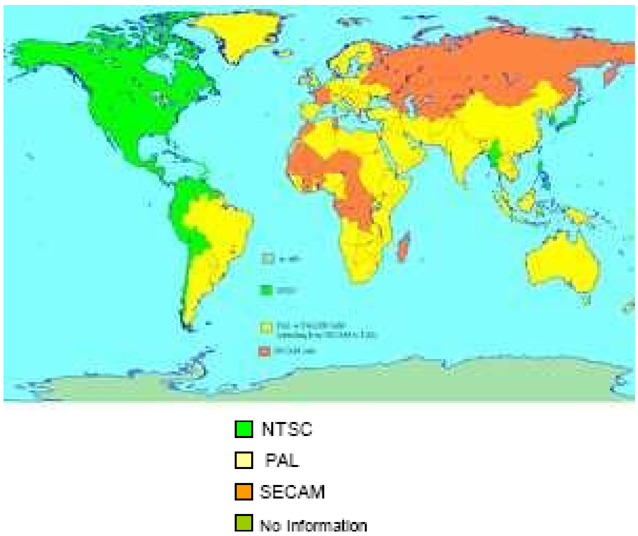


Figure 3. Color System used in the world TV system

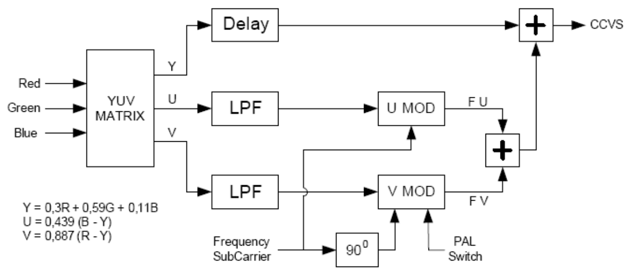


Figure 4. Block diagram of PAL encoder

Y signal is the Luminance or moving images without color (black and white), U and V are the color component of PAL system with a certain composition as shown in the Figure. The encoder uses the Quadrature Amplitude Modulated Subcarrier as modulation method which resulted in Color Composite Video Signal (CCVS) or simply called as video. This analog video is then converted into uncompressed digital video using ADC (Analog to Digital Converter). Video compression is used to reduce the digital video bitrate and get more efficiency in memory or storage space and resulted in the reliability in transmission both audio and video. There are several compression methods, such as AVI, WAV, MP3 and MPEG. In the case study at one of TV broadcasting company, MPEG compression is selected.

The video format is 4:2:0, PAL with 25 frame per second bit rate, so the 122 Mbps video could be compressed to only 12 Mbps.

III. VIDEO COMPRESSION METHOD

The compression of audio and video is performed by reducing the data redundancy in transmission process and recover the original data in the receiver side. Pair of transmiter and receiver is called the codec (coder and encoder). A video codec is a device or software that enables video compression and/or decompression for digital video. In the video signal there are two redundancies: (a) temporal and spatial redundancy and (b) psychovisual redundancy. Temporal redundancy could be defined as pixels in two video frames that have the same values in the same location. Exploiting temporal redundancy is one of the primary techniques in video compression, while spatial redundancy is the elements that are duplicated within a structure, such as pixels in a still image and bit patterns in a file.

Psychovisual redundancy deals with lossy compression, so the information could be lost to some extent. Since the human eye does not respond to all visual information with equal sensitivity, some information is simply of less relative importance. This information is referred to as psychovisual redundant and can be eliminated without introducing any significant difference to the human eye. The reduction of redundant visual information has some practical applications in image/video compression. Since the reduction of psychovisual redundancy results in quantitative loss of information, this type of reduction is referred to as quantization. The most common technique for quantization is the reduction of number of colors used in the image, thus color quantization. Since some information is lost, the color quantization is an irreversible process. Even if this method of compression is lossy, in situations where such compression technique is acceptable the compression can be very effective and reduce the size of the video considerably.

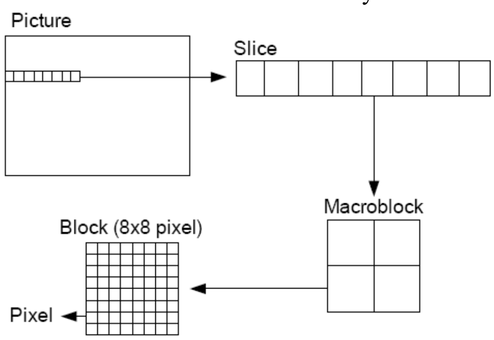


Figure 5. Hirarchy of digital frame

MPEG2 describes a combination of lossy video compression and lossy audio compression (audio data compression) methods which permit storage and transmission of movies using currently available storage media and transmission bandwidth. It is widely used as the format of digital television signals that are broadcast by terrestrial (over-the-air), cable, and direct broadcast satellite TV systems. It also specifies the format of movies and other programs that are distributed on DVD and similar disks. Basically the structure of the frame in MPEG2 is the same as digital frame shown in Fig. 5. Since there are redundancies in the video frames, several frames could be predicted from the previous or the next one. Therefore, not all of the frames are transmitted so the bandwidth becomes smaller. MPEG-2 specifies that the raw frames be compressed into three kinds of frames: intra-coded frames (I-frames), predictive-coded frames (P-frames), and bidirectionally-predictive-coded frames (B-frames). An I-frame is a compressed version of a single uncompressed (raw) frame. It takes advantage of spatial redundancy and of the inability of the eye to detect certain changes in the image. Unlike P-frames and B-frames, I-frames do not depend on data in the preceding or the following frames. Briefly, the raw frame is divided into 8 pixel by 8 pixel blocks as shown in Fig. 4. The data in each block is transformed by a discrete cosine transform. The result is an 8 by 8 matrix of coefficients. The transform converts spatial variations into frequency variations, but it does not change the information in the block; the original block can be recreated exactly by applying the inverse cosine transform. The advantage of doing this is that the image can now be simplified by quantizing the coefficients. Many of the coefficients, usually the higher frequency components, will then be zero. The disadvantage of this step is the loss of some subtle distinctions in brightness and color. If the inverse transform is applied to the matrix after it is quantized, the output image looks very similar to the original image but that is not quite different. Next, the quantized coefficient matrix is compressed. Typically, one corner of the quantized matrix is filled with zeros. By starting in the opposite corner of the matrix, then zigzagging through the matrix to combine the coefficients into a string. Next step is substituting run-length codes for consecutive zeros in the string, and then applying Huffman coding to that result. The matrix will be reduced to a smaller array of numbers. This array is the information that is broadcast and in the receiver side, the whole process is reversed, enabling it to reconstruct, to a close approximation of the original frames of video. Typically, every 15th frame or so is made into an I-frame. P-frames and B-frames might follow an I-frame like IBBPBBPBBPBB(I), to form a Group Of Pictures (GOP).

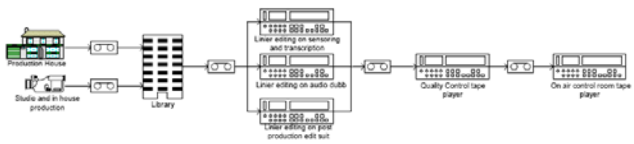


Figure 6. The production process of tapeless on-air system

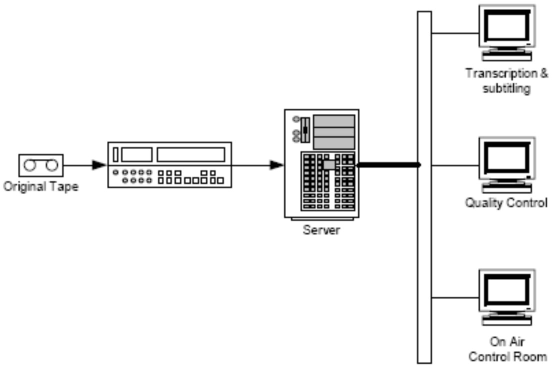


Figure 7. The production process of movie program: "Layar Emas" using tapeless on-air system

In the TV broadcasting company, digital video is stored, edited, and transmitted as a production process which is depicted in Fig. 6. For a movie production, "Layar Emas" for example, the tapeless on-air production involves 3 clusters of server node as shown in Fig. 7. Each node has 8 harddisks with capacity of 36 Gbyte (for cluster 1 and 3) and 146 Gbyte (for cluster 2). The total capacity of video server is then (2 nodes × 36 Gbyte × 8 × 3) + (1 node × 146 Gbyte × 8 × 3)= 5,232 Gbyte. The storage format used in these servers is RAID5, therefore 1/3 of the total capacity of the servers is used to store RAID5 data format. The advantage of this system is keeping the server up even there is an error in one of the disk. The memory allocation of the video server is 1.744 Gbyte (1/3 of the total system capacity) for system format RAID5, 1,152 Gbyte for commercials and 2,336 Gbyte for TV broadcast program. Audio and video data which has already broadcasted will be erased from the disk and replaced by a new program. Since November 2006, the station has been applied the tapeless on-air system comprehensively. All of programs such as dramas, movies, cartoons and news (in package) are broadcasted using a playlist which contains the audio and video stored in the server. Although magnetic tape player/recorder (VTR) is still prepared in the control room to handle any trouble on the servers. VTR is also used when there is a delay of program recordings delivery to the station, so there is no time to ingest (converting, editing, etc.) the video.

IV. ANALYSIS RESULTS

Since it was launched in November 2006, the tapeless on-air system is evaluated based on errors occur in the production process. The data is sampled from editing room, libray and master control before and after the new system is applied. The process monitoring is performed manually based on complaints from the editing operators or editors. Operational hours of editing room is 16 hours, while the library and master control rooms is 24 hours, respectively.

Based on the number of complaints, the errors are categorized as: VTR error (such as jammed), NLE & Server (hang, hardware error), mechanics (roller, tape cut-off), electronics (power line, synchronization), installation (driver corrupts), software, human (skill).In the period of 20003-3006, the total number of errors and monthly averaged errors occur in the production process is shown in Fig. 8 and Fig. 9, respectively.

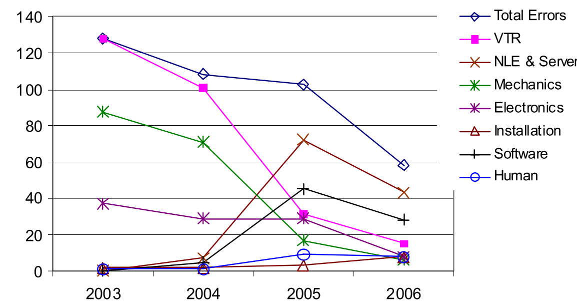


Figure 8. Total number of defects and errors in the production process

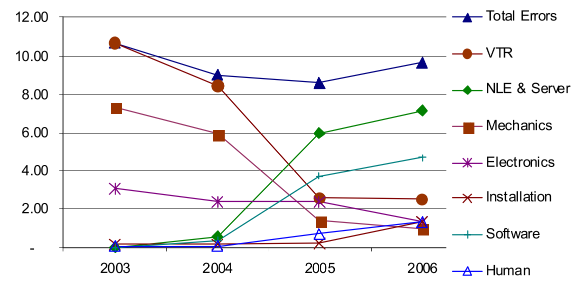


Figure 9. Monthly averaged number of defects and errors in the production process

For comparison, the total number of errors in tape-based and tapeless on-air system in the production process are also presented in Fig. 10 and 11, respectively. In these figures, the trend of tape-based errors is decreasing since the application of tapeless system since Nov. 2006. VTR is the most erroneous part in the system since it involves mechanical movement of magnetic tapes, however, these errors are decreasing as the tapeless on-air system is applied in 2006.

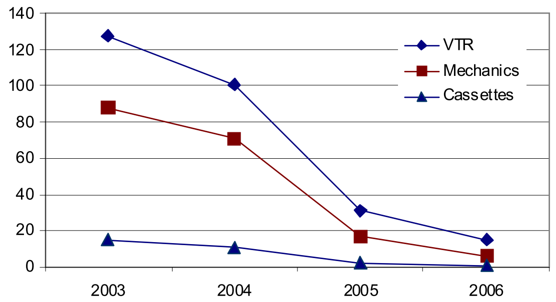


Figure 10. Total number of errors on tape-based on-air system

In the tapeless on-air system, the most erroneous part is the NLE-servers, since they had initial errors since the first intallation, followed by software due to the same reason. The installation error is increasing due to the new handling of software installation that needs skillful operator. As shown in Fig. 9, the lack of skillful operator increases the error caused by human factor.

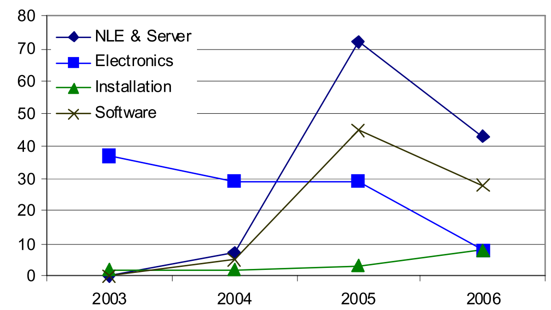


Figure 11. Total number of errors on tapeless on-air system

Table 1. Observation results of audio and video quality in one of the TV program

Date	**Time**	**Program (Movie Title)**	**Error Type**		**Err. Frames**
			Video	**Audio**	
14/11	22:24	Twin Effect 2	Boxes on moving object (BoMO)	-	2-3
	22:43		BoMO	-	2-3
	23:38		BoMO	-	2-3
15/11	22:15	Independence Day	Distributed crystal box (DCB)	-	2-3
16/11	22:45	Swordsman 2	BoMO	-	2-3
17/11	-	Setahun Samsons	-	-	-
18/11	22:49	Lovely Luna	DCB	-	2-3
	23:58		DCB	-	2-3
19/11	22:56	Ih Mama Capek Deh!	DCB	-	2-3
	23:08		DCB	-	2-3
	23:24		DCB	-	2-3
20/11	-	OB Spesial	-	-	-
	-	Delik	-	-	-
	-	Tracking	-	-	-
21/11	22:45	Armor of God	Freeze 1 frame	-	2
	23:00		DCB	-	2-3
22/11	-	The Peace Maker	-	-	-
23/11	22:45	Born to Defence	DCB	-	2-3
24/11	-	Jaka Tingkir	-	-	-
25/11	22:55	Tak Biasa	Lined Video	-	25
	23:40		DCB	-	2-3
	23:55		DCB	-	2-3

To evaluate the quality of video broadcasting, the observation is performed for a month on a program (layar emas) and other programs in the same time segment (22:00-24:00). The program is observed using a TV unit (SHARP fineCrystal 21inch). The main observation is the quality of digital audio and video, not the strength of the signal received.. If the error occurs in the quality of audio/video, then the source of errors resulted from production process or computer storage. If the errors occur in the receiving unit, then the source may be from the broadcasting or TV receiving system. The observation result is shown in Table 1.

V. CONCLUSIONS

The implementation of tapeless on-air system has reduced the total errors in TV broadcasting system. From the data samples obtained, during 2003-2006, the decreasing of errors by implementing tapeless on-air TV system could reach 21% (2003-2004), 69% (2004-2005), and 17% (2005-2006), respectively. Human and installation error is increasing due to the lack of human resources. The quality of video observed mainly due to the production process and errors averaged at 2-3 frames, except a program with lined video which exceeds 20 frames.

REFERENCES:

1. P. N Tudor, MPEG-2 Video Compresion, Electronics & Communication Engineering Journal, December 1995.
2. The Free Encyclopedia, PAL, Wikipedia, May 2006.
3. The New Paradigm for Digital Video Storage & Delivery, SeaChange Broadcast MediaCluster.
4. Technical On Air Control Room Section dan Maintenance Section, Daily Report, RCTI, 2006.