UDC 681 . 3

Desktop Studio and Its Image Quality Evaluation

Masaki KAMIYA*(1) Masao SHOUDA*(1)

Abstract:

Taking the place of the analogue video editing system, a video editing system utilizing computers, the so-called non-linear editing system, is becoming the mainstream for editing images. However, it has been considered that, in the case of non-linear editing, a deterioration of image is generated easily due to the image compression generally adopted, which differs from the case of analogue editing. Nippon Steel has established a highly precise image evaluating method through evaluating the image from which the characteristics have been extracted, instead of applying subjective evaluation, which is the general method for evaluating image quality. In this report, described are the method established and the results obtained by applying it to the evaluation of various images.

1. Introduction

Recently, an editing system utilizing computers, called a non-linear editing system, has been the mainstream for editing and can freely deal with materials stored on a hard disk without deterioration of images during editing/copying from the master tape, a problem that is unavoidable in the analogue editing environment. Specifically, this is because shortening of the process utilizing the random access editing work is possible. Some systems have been developed, which are said to assure NTSC online quality (broadcast quality) using an open system with personal computers as well as using a workstation and a closed system.

However, the non-linear editing system using personal computers employs an image compression technology to reduce data transmission load inside computers that deteriorates the quality along with compression, hence it is said that it is not available for online editing and the system cannot create materials for image synthesis.

On the other hand, an advantage of image compression is enormous from the viewpoint of system cost. In particular, for long-term image editing, the hard disk capacity by a simple calculation is:

Y = 720 × 486 × 1 ≈ 342 Kbytes
Cr = 360 × 486 × 1 ≈ 171 Kbytes
Cb = 360 × 486 × 1 ≈ 171 Kbytes (1024 bytes = 1 Kbyte)

Y + Cr + Cb = 684 Kbytes/Frame

for one frame based on 4:2:2 by ITU-R (International Telecommunications Union, Radio communication division) recommendation BT.601-4 (former CCIR601).

The capacity per second for non-drop frame is:

684 k × 30 = about 20 Mbytes (1024 Kbytes = 1 Mbyte)

The capacity per hour is:

20 M × 60 × 60 = about 70 Gbytes (1024 Mbytes = 1 Gbyte)

Consequently, the hard disk cost to store huge image data exceeds the system cost and the number of management processes increases because of the disk backup maintenance. As a result, the compressed image data can not help being dealt with from the viewpoint of cost.

Compressed material usually leads to deterioration of image quailty. Especially, in the motion JPEG (M-JPEG), which is a compressed form frequently used in video capture boards, peculiar deterioration phenomena occur depending on DCT coding (discrete cosine transformation) used in its compression algorithm. The deterioration includes not only image information but also color information. As for color information, deterioration depends on the process inside the capture board hardware. Therefore, depending on compression with 16-bit signal corresponding to ITU-R BT.601-4 (CCIR601) 4:2:2 or compression with 24 bits to ITU-R BT.601-4 (CCIR601) 4:4:4, the type of deterioration can be changed. Moreover, a computer uses RGB for the internal process, which requires color/space conversion between image signals and color

*(1) Electronics & Information Systems Div.

Table 1 Proposed specification of ITU-R BT. 601-4

Family member	Application	Sample per line Y/G	Cr/R	Cb/B
4:4:4 RGB 4:4:4 YCrCb	Television source equipment High quality signal processing	720	720	720
4:2:2 YCrCb	Standard digital interface International program exchange	720	360	360

Table 2 RGB to YUV color spatial conversion

Y	0.229R + 0.587G + 0.144B
U	0.147R – 0.289G + 0.436B
V	0.615R – 0.515G + 0.100B

signals of the computer. In this conversion, color repeatability is also checked. For reference, **Table 1** shows recommended specification in ITU-R BT.601-4 and **Table 2** shows color/space conversion of RGB To YUV.

The authors evaluated deterioration of image due to compression, correlation between compressed ratio and image deterioration, and the possibility of an alternative for improving the precision of the conventional subjective evaluation method. The authors investigated the possibility of editing NTSC quality online using an open system with personal computers.

2. Evaluation instruments and method

2.1 Evaluation instruments

AVID MCXpress for Windows NT*1 software dealt with in Electronics & Information Systems Division of Nippon Steel Corp. is combined with Truevision TARGA 2000DTX capture board for a desktop image editing system or "desktop studio system". This system uses devices shown in **Table 3** to evaluate image quality in various compressed ratios by actual image incorporation.

Image signals are separated to avoid dependence on characteristics of the image separation circuit of the monitor and board, so that an analogue component signal of image deterioration was used as a connecting medium. The compression ratio was calculated for each capacity per frame referring to the value set in AVID MCXpress for Windows NT for convenience.

2.2 Evaluation method

For evaluation, the images captured under various compression ratio are fetched from AVID MCXpress for Windows NT as static images for subjective evaluation. The subjective evaluation was set in ITU-R.BT.500 but was likely affected by evaluation variations among test subjects, which is very difficult for use in image evaluation accompanied with compression. On the other hand, subjective evaluation by a skilled person sometimes shows high precision exceeding performance of the ordinary devices.

This study revealed that evaluation variation among test subjects for compressed images can be cleared by extracting characteristics in the subjective evaluation. In the applied compression algorithm, characteristics in which compression effect appears in color information rather than in brightness information is mainly reflected for subjective evaluation. In evaluation including color component, noise peculiar to JPEG can be judged as shown in **Fig. 2** (a) through (c) even from images not recognized (see **Fig. 1** (a) through (c)) if utilizing characteristics where compression effect appears in the brightness information rather than in color information. With this method, the authors evaluated images having a nearly unchanging patterned picture, for example, a picture in which one unfigured dish is placed on a blue background (low spatial frequency) and a sharp changing patterned picture in which fine images like check patterns (high spatial frequency) are drawn. In compression, results are different depending on high or low frequencies.

Table 3 Evaluation instruments

	Name of manufacturer	Model No.
Test signal generator	Reader Electronic	425A
Camera	SONY	PVW-637
VTR	SONY	DVW-A500

Generally, JPEG compression used in the non-linear editing system is called non-reversible compression in which not all the information that existed before compression is reproduced, but a part of the information is disposed of. This leads to common image deterioration. In particular, there is an image deterioration peculiar to DCT signals used in the JPEG compression method. This means that each inputted animated image is divided into pixel blocks to cause characteristic image deterioration. Therefore, a great deal of image information is disposed of during compression, thereby producing further image deterioration.

Images used in this evaluation employ a 75% SMPTE color bar as a sample with less display information than test signal generator (Leader Electronics 425A) and a line sweep display as a sample with a great deal of display information. Alternately, as an image having intermediate information volume, actual animated image information is captured with the shooting camera SONY PVW-637 (BETACAM*2 SP) often used for original editing work for the non-linear editing device to be reproduced and taken in by the SONY DVW-A500.

Images were evaluated based on the quality set display of AVID MCXpress for Windows NT as a standard. This produces some errors between written frame capacity and actual capacity, but the authors judged it to be no problem for estimation of the correlation between compressed ratio and image quality. The above described three types of images (see Fig. 1) were respectively captured as animated images for evaluation on the image capture display in AVID MCXpress for Windows NT at 50 (Fig. 2), 100 (**Fig. 3**), 200 (**Fig. 4**), 300 (**Fig. 5**), and 400 KB/Frame (**Fig. 6**). From these images, the animated frame was searched so as to be the same scene at the respective compressed ratios and the quality was evaluated by picking up the portion of the scene as a static image.

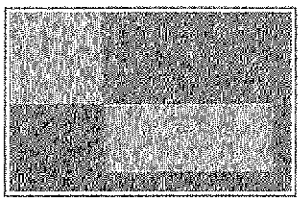
3. Evaluation results

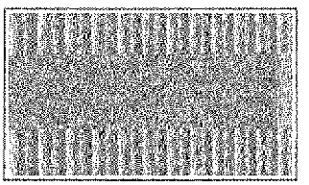
3.1 Evaluation by 75% SMPTE color bar

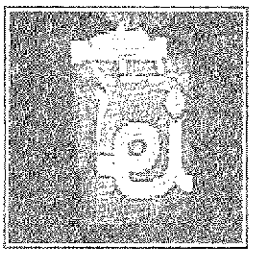
Edges in 50 to 100 KB/Frame were rough and the block noise as a typical symptom of image deterioration peculiar to JPEG was also recognized clearly. However, at more than 200 KB/Frame, little difference was confirmed as a result. This noise appeared significantly at an area where color information changed, which might be generated due to rapid increase in color information. In addition, at more than 200 KB/Frame, repeatability was improved to the extent that any change in image is difficult to find. The reason

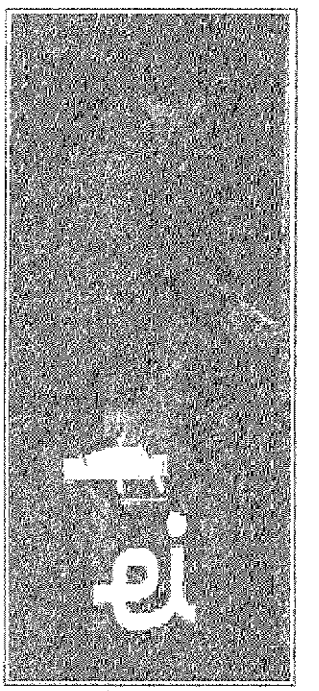
*1 Windows NT is the trademark of the U.S. Microsoft Corporation in the United States and other countries.

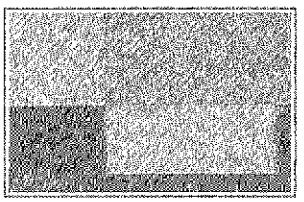
*2 BETACAM is the trademark of Sony Corp.

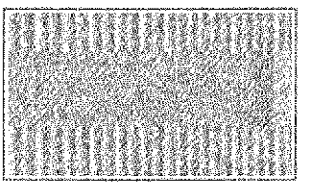

75% SMPTE color bar
Fig. 1 (a)

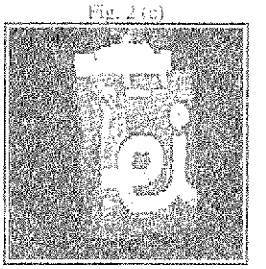

Line sweep
Fig. 1 (b)

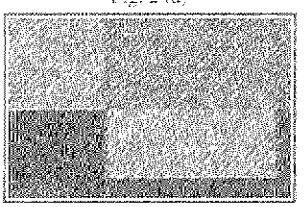

50KB/Frame
Fig. 2 (c)

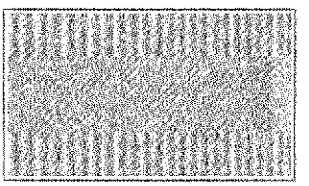

Real image
Fig. 1 (c)

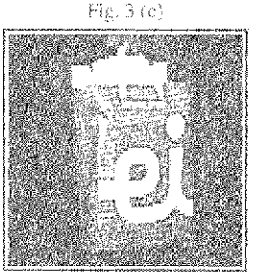

50KB/Frame
Fig. 2 (a)

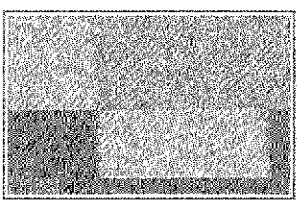

50KB/Frame
Fig. 2 (b)

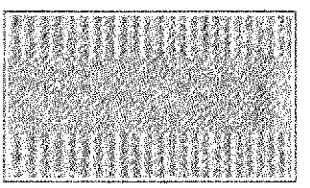

100KB/Frame
Fig. 3 (c)

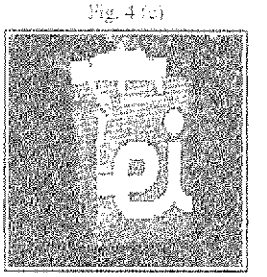

100KB/Frame
Fig. 3 (a)

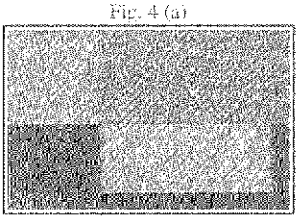

100KB/Frame
Fig. 3 (b)

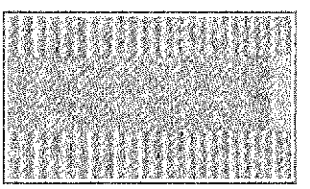

200KB/Frame
Fig. 4 (c)

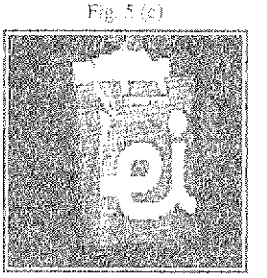

200KB/Frame
Fig. 4 (a)

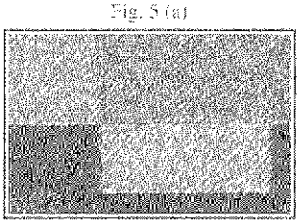

200KB/Frame
Fig. 4 (b)

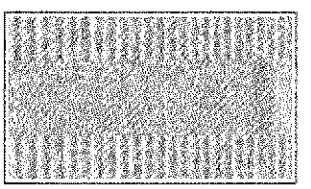

300KB/Frame
Fig. 5 (c)


300KB/Frame
Fig. 5 (a)


300KB/Frame
Fig. 5 (b)


400KB/Frame
Fig. 6 (c)


400KB/Frame
Fig. 6 (a)


400KB/Frame
Fig. 6 (b)

for this is assumed to be that the display information is comparatively little on the 75% SMPTE color bar, so that less information is disposed of after compression.

In this result, capture at 200 KB/Frame (1/3 compression) is sufficient for online editing of the images with less display information.

3.2 Evaluation for line sweep display

For capture from 50 KB/Frame (1/12 compression) to 300 KB/Frame (1/2 compression), image edges are rough and ink stain-like image deterioration is observed. In addition, block noise peculiar to JPEG can be observed. With the capture rate increased from 50 KB/Frame to 300 KB/Frame, the edge becomes smooth and there is less ink stain. It can be proved that the image deterioration

is restricted, the block noise itself cannot be observed anymore and the image deterioration due to compression is relatively lower. At 400 KB/Frame (1/1.5 compression), a level is achieved such that neither rough edge, ink stains nor block noise are confirmed, thus ensuring the online quality.

From the results, it can be assumed that capture at 400 KB/Frame is needed for online editing in the case of image having a great deal of display information. Moreover, the capture on the order of 400 KB/Frame is also necessary for images requiring high quality at the edge (changing point of color information) such as key synthesis.

3.3 Evaluation with real images

The block noise as a symptom of image deterioration peculiar to JPEG can be observed with images of 50 KB/Frame and 100 KB/Frame (1/6 compression), and ink stains can also be observed at the outlined white area on a colored background of paper bag and hand. This suggests that data unnecessary for exact image reproduction are generated by quantum conversion due to compression of the image. On the other hand, from 200 to 400 KB/Frame, despite evaluation with expansion of images, neither measurable problems such as image deterioration nor generation of unnecessary data could be detected.

In any event, the substantially equal result as evaluation with 75% SMPTE color bar of an example of less image information was introduced. Specifically, it can be assumed that the online editing of the real image level is available in quality for 200 KB/Frame or more. However, where the image with high quality at its edge such as key synthesis is required, the edges should be kept clear, so it is desirable to capture the image at 400 KB/Frame for creating a high quality synthetic image.

4. Conclusions

The above evaluation demonstrates that users requiring online images as high as BETACAM SP class by the non-linear editing system should construct the system with a capturing capacity of more than 200 KB/Frame even though depending on VTR types of image recording. If the key synthesis is frequently applied, the system should be constructed with a capturing capacity at 400 KB/Frame. The image can be useful even at 50 KB/Frame, which is practicable with the current limits of hard disk capacity, for example, in the case of off-line.

Finally, the authors present the correlation between the compression ratio and hard disk capacity TNTSC (640 × 480). The hard disk is an important cost element in constructing the non-linear editing system. The hard disk price accounts for a large portion in this type of system. The price of an image capture card varies with wider bandwidth. This means that capture card and hard disk capacity to be used or the system cost are fixed depending on how long to edit the image at which quality level.

Points for constructing the system include:

(1) Which quality is specified for the target image? (for broadcasting, for multimedia use such as CD-ROM)

(2) What is material? (digital BETACAM, business use BETACAM, DVC, S VHS[*3], etc.)

(3) Whether off-line editing or online editing, and so on.

Table 4 shows a transmission rate (KB/Frame) and capturing time per 1 GB of hard disk for a rough estimation.

Table 5 shows the maximum capture rate of Truevision TARBA series with AVID MCXpress for NT as an experimental value.

In the future, the authors aim to establish an objective evaluation method without evaluation variations due to the test subjects such as feature extraction method and picture difference method, and will establish a fine image evaluation for next generation compression techniques such as the spatial prediction algorithm (spectral selection) using reversible compression and wavelet conversion.

Table 4 Transmit rate and capturing time per 1 GB

Transit rate (KB/Frame)	Compression rate	MB/s	min/1GB	Quality	Remarks
60KB	10:1	1.8	10min	Off line	
100KB	6:1	3	6min	S VHS	
120KB	5:1	3.6	4min30s	DVC	
150KB	4:1	4.5	4min	Business use	
200KB	3:1	6	3min	BETACAM	
300KB	2:1	9	2min	Digital BETACAM	
400KB	1.5:1	12	1min30s		MCXpress maximum value

Table 5 Capture rate of TARGA series

TARGA	Maximum capture rate	Image input/ output	Remarks
1000	160KB/Frame	S, composite	-
1000Pro	160KB/Frame	Component or S, composite	CCIR601 compatible
2000	160KB/Frame	S, composite	-
2000Pro	160KB/Frame	Component or S, composite	CCIR601 compatible
2000DTX	400KB/Frame	Component, S, composite	CCIR601 compatible
2000RTX	400KB/Frame	Component, S, composite	CCIR601 compatible

[*3] S VHS is the trademark of Victor Co. of Japan, Ltd. Names of Corporation and products written in this papers are the trademarks or registered trademark in each corporations.