Digital Video/Audio NR+RR Monitoring System Based on Motion Compensated Interframe/Intraframe Objective Parameters

Takahiro HAMADA

KDD Media Will Corporation
2-1-23 Nakameguro Meguro-ku Tokyo 153-0061, Japan
TEL: 81-3-3794-8174, FAX: 81-3-3794-8179, E-MAIL: ta-hamada@kddi.com

Abstract

We have developed Digital Video/Audio Monitoring System. The basic component is 1 U rack size unit, which has one video and two audio inputs and data outputs. This unit can monitor video/audio quality with No Reference (NR) signal based on the MC interframe / intraframe objective parameters, and results are clearly displayed at each local point. Further, with use of NR monitoring results, this unit can also produce essential signals as the Reduced Reference (RR) signal for video/audio quality relative measurement, which can be coded with very limited number of bits (less than 64kbits/s). By connecting these units via common network such as LAN and IP, RR signals at remote monitoring points can be gathered into a central control site. At this site, each RR signal, which corresponds to each monitoring point are compared against each other and quality parameters for each video link can be easily calculated at real-time base: This is accomplished by the sample-less MSE approximation algorithm. By the quality parameters, we can immediately locate the problem link in case of service failure. In addition, we can accumulate data related to video/audio quality by logging parameters for long period.

keywords: Monitoring System No Reference, Reduced Reference, Sample-less MSE Apporoximation.

1. Introduction

In the era of analog television service, video/audio quality monitor was conducted by mankind for the limited number of service channels. In the era of digital video/audio however, quality monitor by human operators are no longer practical because too many service channels have become a reality such as digital TV broadcasting, video over IP, DVD and many more.

Considering these backgrounds, we propose solutions such as the Digital Video/Audio Monitoring System. The basic component is 1 U rack size unit, which has one video and two audio inputs and data outputs. This unit can monitor video/audio quality with No Reference (NR) signal, and results are clearly displayed as video/audio failure such as frozen, disappearing and collapse video and audio mute and collapse. Also, degradation of video quality by coding artifacts and resolution loss is given. With use of NR monitoring results, the NR unit can also produce essential signals as the Reduced Reference (RR) signal for video/audio quality comparisons. RR signal can be coded with very limited number of bits (less than 10kbits/s), therefore each unit can be connected easily via common network such as LAN and IP. Under this configuration, RR signals at remote monitoring points can be gathered into a central control site. At this site, each RR signal, which corresponds to each monitoring point are compared against each other and quality parameters for each video link can be easily calculated by PC software at real-time base. By the quality parameters, we can immediately locate the problem link in case of service failure. In addition, we can accumulate data related to video/audio quality by logging parameters for long period.

This paper describes a general system outline, block diagram of hardware for NR measurement and software for RR analysis and system specifications.

2. System outline

Fig.1 shows an external view of the basic NR unit. The NR unit is 1U rack size hardware, to which one video and two audio channels can be fed in. Using this unit and with control PC, monitoring system can be configured as shown in Fig.2 and Fig.3. These two figures show the configuration for the assessment by NR and that by RR, respectively. In Fig.2, the video input into the NR unit is that compressed and regenerated by the encoder and the decoder. The NR unit measures the video quality only from these input signals, and displays the results on PC in a graphic form. On the other hand in Fig.3, the results of the assessment are regarded as a kind of the Reduced Reference of the reference video at the originating point of transmission, and the RR is sent via a separate route to be gathered on PC at one location. By comparing the reduced references, it detects and extracts video/audio failure and video loss respectively.

The followings are menus of quality measurement.

NR (No Reference) quality measurement at a monitoring point (Fig.2)

1) Video Botch: Sudden and short time period of frozen video
2) Video Freeze: Frozen picture for long time
3) Video Loss: Sudden and short time period of video disappearing
4) Black Out: Video disappearing for long time
5) Coding Artifacts: Block shaped distortion by MPEG
6) Resolution Loss: Contour blurring and edge dull
7) Video Collapse: Severely distorted video signal
8) Stripe Noise: Limited number of lines (stripe) are severely damaged
9) Audio Mute: No audio signal
10) Audio Collapse: abnormal audio signal

RR (Reduced Reference) quality measurement of a link (Fig.3)

1) Video Botch and Freeze
2) Video Loss and Black out
3) Video Collapse and Stripe Noise
4) Audio Mute and Collapse
5) Increase (Decrease) of Coding Artifacts
6) Increase of Resolution Loss
7) Pseudo SNR of Video and Audio

The NR unit supports Component SDI, NTSC/PAL analog and 1080I HDTV SDI for video, and analog and AES/EBU digital audio for SDTV and embedded digital audio for HDTV (Table 1).

3. Block diagram of NR quality monitoring unit hardware

Fig.4 shows the block diagram inside the NR quality monitoring unit hardware.

The inputs of video signal and two lines of audio signals are possible. When video signal is in SDTV format, audio signals become AES/EBU, and when video signal is in HDTV, the audio signals are embedded audio.

First, when synchronization pulse (Trigger pulse) of video/audio signals are not detected, the NR unit outputs “Sync loss.” Then, for video signals, the NR unit provides a memory for the current, previous and two-orders-ahead frames, and measures signal difference between the previous frame with respect to the motion compensated frames, before (two-orders-ahead) and after (current). This signal difference indicates the interframe correlation (ñ) for each stripe in a frame, by which an abnormal signal inserted can be detected as collapse and stripe noise. In addition, with respect to the video data of the previous frame, it conducts the video level (L)/activity (S) calculation as well as the video CZC (cyclic zero cross) calculation (A). Based on the results of these calculations, the equipment outputs, using functions Fv (1) - Fv (6), objective parameters relating to various types of video quality

In parallel with this, the NR unit conducts the audio level/activity calculation after the frame delay with respect to audio signals, and based on the results of the calculation it outputs, using functions F_A (1) and F_A (2), objective parameters relating audio.

4. Block diagram of PC software for RR comparisons

Fig. 5 shows block diagram of PC software for RR comparisons. Each signal RR A, RR B and RR C at remote monitoring point is gathered into the PC. In the PC, four functional blocks are applied; those include Delay calculation, Video failure detection, Audio failure detection and Video loss detection for link A-B and B-C. In the Delay calculation, frame activity and level in gathered RR signals are correlated each other on temporal domain among A, B and C, and the frame delay which gives the highest correlation is searched. This search is conducted periodically to refresh frame delay to keep measurement as accurate as possible. The obtained frame delay is given to the other blocks so that comparisons can be performed by delay compensated signals. Two blocks are used to detect failures of video and audio. This is simple logic wherein difference between RR signals indicates failure as OK or NG. On the other hand, video loss is calculated as continuous value using video activity, level and taking difference of coding artifact and resolution loss of A-B and B-C. All results of RR comparisons and also NR results are graphically displayed by sophisticated GUI software.

5. Algorithm descriptions

The system is designed to catch every error of video, even better than human eyes. This is accomplished by adopting a mixture of sophisticated algorithms for NR and RR quality monitoring, which are described as function blocks in Fig.4 and Fig.5.

Among these, NR algorithms of MC (Motion Compensation) based noise detection and RR comparison algorithm to measure link performance measurement is also described.

5-1. No reference MC based noise detection algorithm

Fig.6 shows motion compensated interframe difference calculation, where minimum value D is selected from backward and forward field at 8 × 8 block base.

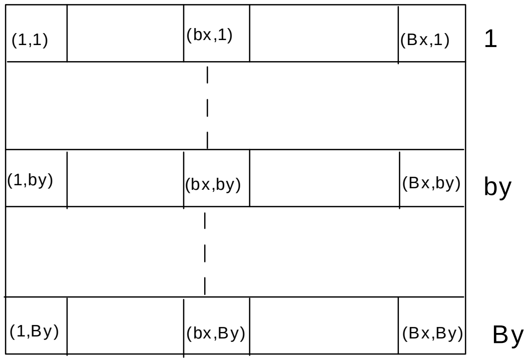


Fig.7 Block, stripe and field

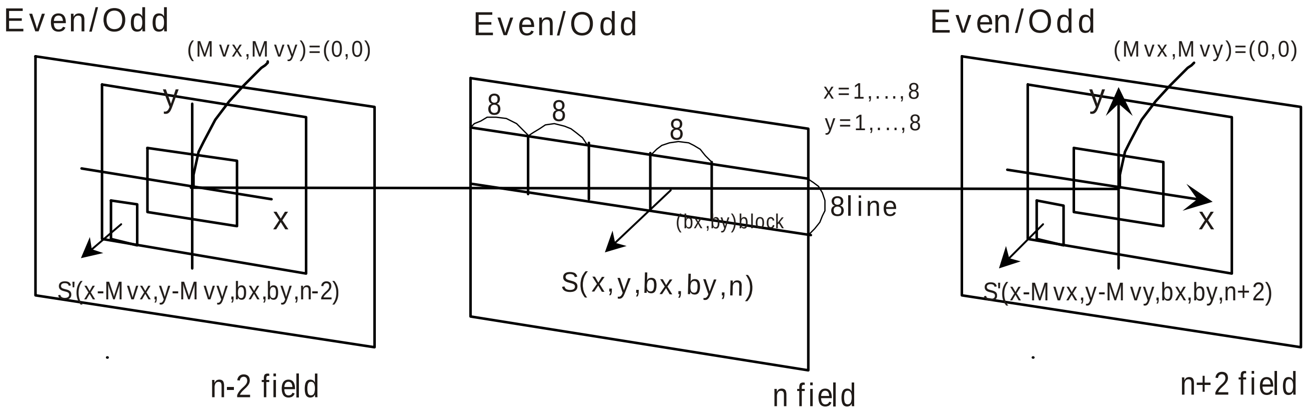


Fig.6 MC interframe difference calculation

Thus,

$$D(bx,by,n)=\frac{1}{64}\sum_{x=1}^{8}\sum_{y=1}^{8}$$

$$\left|S(x,y,bx,by,n)-S'(x-Mvx,y-Mvy,bx,by,n+p)\right| \quad (1)$$

where p=2 or –2,
and
Dmin (bx, by, n) for

$$-16\,pel \le Mvx \le +15\,pel$$

$$-8\,line \le Mvy \le +7\,line$$

Let Dmin (by, n) be defined as an averaged of Dmin over stripe by
Thus,

$$D_{\min}(by,n)=\frac{1}{Bx}\sum_{bx=1}^{Bx} D_{\min}(bx,by,n)$$

At each stripe, correlationñ(by, n) can be obtained as follows,

$$r(by) = 1 - \frac{D^2{}_{\min}(by,n) + D^2{}_{\min}(by,n-1)}{2S(by)} \qquad (2)$$

where S (by) is a stripe based activity.

$$S(by) = \frac{1}{64Bx}\sum_{bx=1}^{Bx}\sum_{m=1}^{8}\sum_{n=1}^{8} \left\{x(bx,by,m,n) - \bar{x}(bx,by)\right\}^2 \qquad (3)$$

We set the determination of “noise strip” with the following condition.

$$r(by) \leq Thr \qquad (4)$$

5-2. Sample-less MSE approximation for RR comparison

We have a signal model as Fig.8, where group of samples V (x, y, z, t) are processed and U (x, y, z, t) are reproduced.

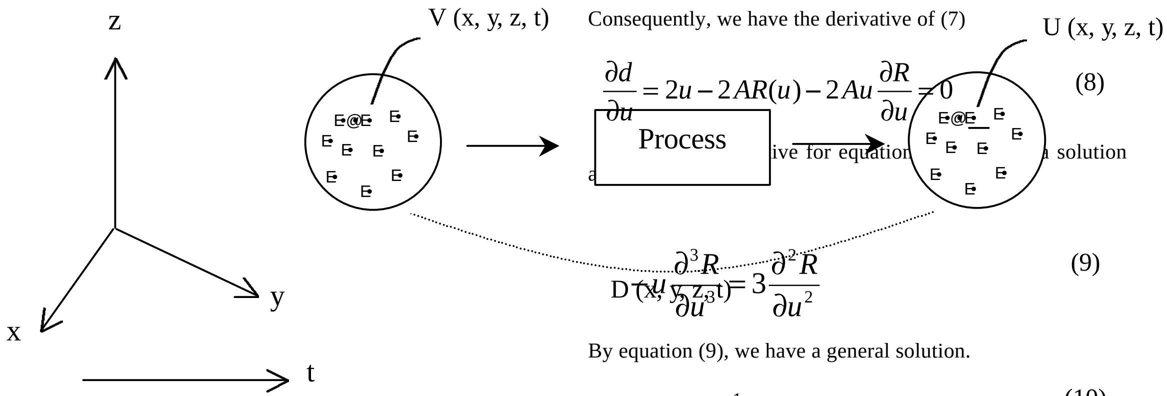


Fig.8 Signal processing model

Between V and U, we can define the signal difference D (x, y, z, t). D is easily introduced as follows.

$$\begin{aligned}\overline{D^2} &= \overline{|V-U|^2} \\ &= \frac{1}{XYZT}\sum_{x=1}^{X}\sum_{y=1}^{Y}\sum_{z=1}^{Z}\sum_{t=1}^{T} \{V(x,y,z,t) - U(x,y,z,t)\}^2\end{aligned} \qquad (5)$$

To obtain $\overline{D^2}$ it is required that every sample value V (x, y, z, t) and U (x, y, z, t) is available at the same site, that is, Full Reference (FR) calculation is needed. But, if we can calculate $\overline{D^2}$ using only averaged scholar value such as $\overline{V^2}$ and $\overline{U^2}$, a greatly reduced reference is accomplished and a lot of beneficial things we can expect.

We can re-express the equation (5) as follows.

$$\overline{D^2} = A^2 + u^2 - 2AuR \qquad (6)$$

Where, A is an averaged V as given value (original single).

And u is an averaged U but a variable by a signal process and R is a correlation of V and U.

On equation (6), $\overline{D^2}$ is varied by a signal process, which is expressed by u and R.

Therefore, we can define $\overline{D^2}$ and R as a function of u. Then, we have,

(3)

$$d(u) = \overline{D^2} = A^2 + u^2 - 2AuR(u) \qquad (7)$$

Practically, the signal process is designed, controlled and operated so that $\overline{D^2}$ is given as minimum as possible.

(4)

Consequently, we have the derivative of (7)

$$\frac{\partial d}{\partial u} = 2u - 2AR(u) - 2Au\frac{\partial R}{\partial u} = 0 \qquad (8)$$

ive for equation a solution a

$$-u\frac{\partial^3 R}{\partial u^3} = 3\frac{\partial^2 R}{\partial u^2} \qquad (9)$$

By equation (9), we have a general solution.

$$R(u) = au^{-1} + bu + c \qquad (10)$$

With equation (7), we have

$$d(u) = (1-2bA)u^2 - 2cAu + A^2 - 2aA \qquad (11)$$

To determine a, b and c, we introduce a model and conditions as shown in Fig.9.

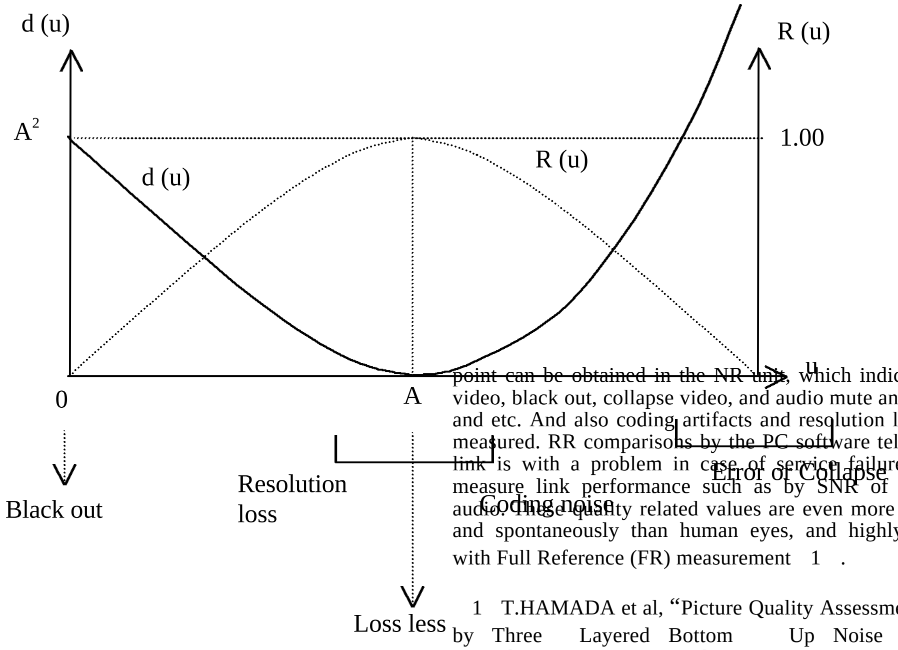


Fig.9 Model and conditions

Then, we separate u A and u A

( ) u A

u 0 , R 0 & d A^2

give a=0, b=1/A and c=0

Then,

$$d(u) = A^2 - u^2 \tag{12}$$

( ) u A

u , R 0 give b=0 and c=0,

And u=A , d=0 give a=A

Then,

$$d(u) = u^2 - A^2 \tag{13}$$

6. Conclusion

For the purpose of quality monitor solution in the era of digital video/audio services, we propose the video/audio NR and RR monitoring system based on motion compensated interframe/intraframe objective parameters. The key component of the system is an 1U rack size NR quality monitoring unit, and the system is configured by connecting NR units to controlling PC, wherein a software conducts RR comparisons based on the sample-less MSE approximation algorithm. In this system, essential parameters for quality at monitoring point can be obtained in the NR unit, which indicate frozen video, black out, collapse video, and audio mute and collapse, and etc. And also coding artifacts and resolution loss can be measured. RR comparisons by the PC software tell us which link is with a problem in case of service failure and also measure link performance such as by SNR of video and audio. These quality related values are even more accurately and spontaneously than human eyes, and highly correlate with Full Reference (FR) measurement [1].

[1] T.HAMADA et al, “Picture Quality Assessment System by Three Layered Bottom Up Noise Weighting Considering Human Visual Perception, SMPTE Journal, January 1999

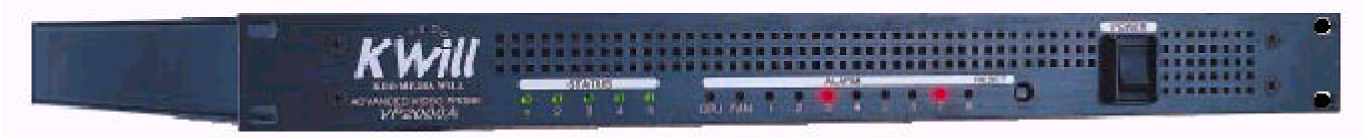
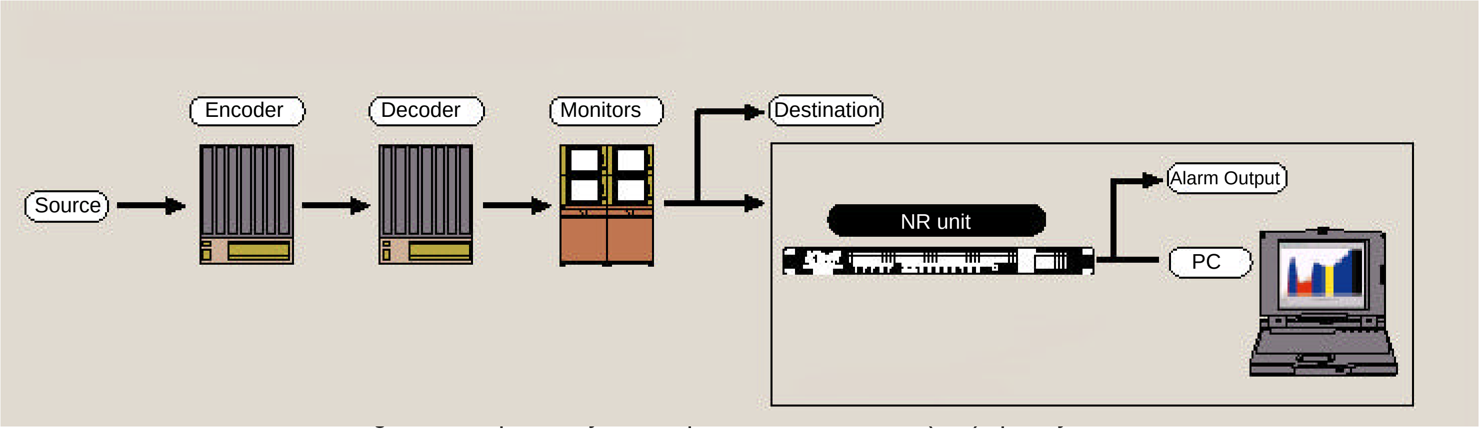


Fig.1 External view of basic NR unit of video/audio monitoring system



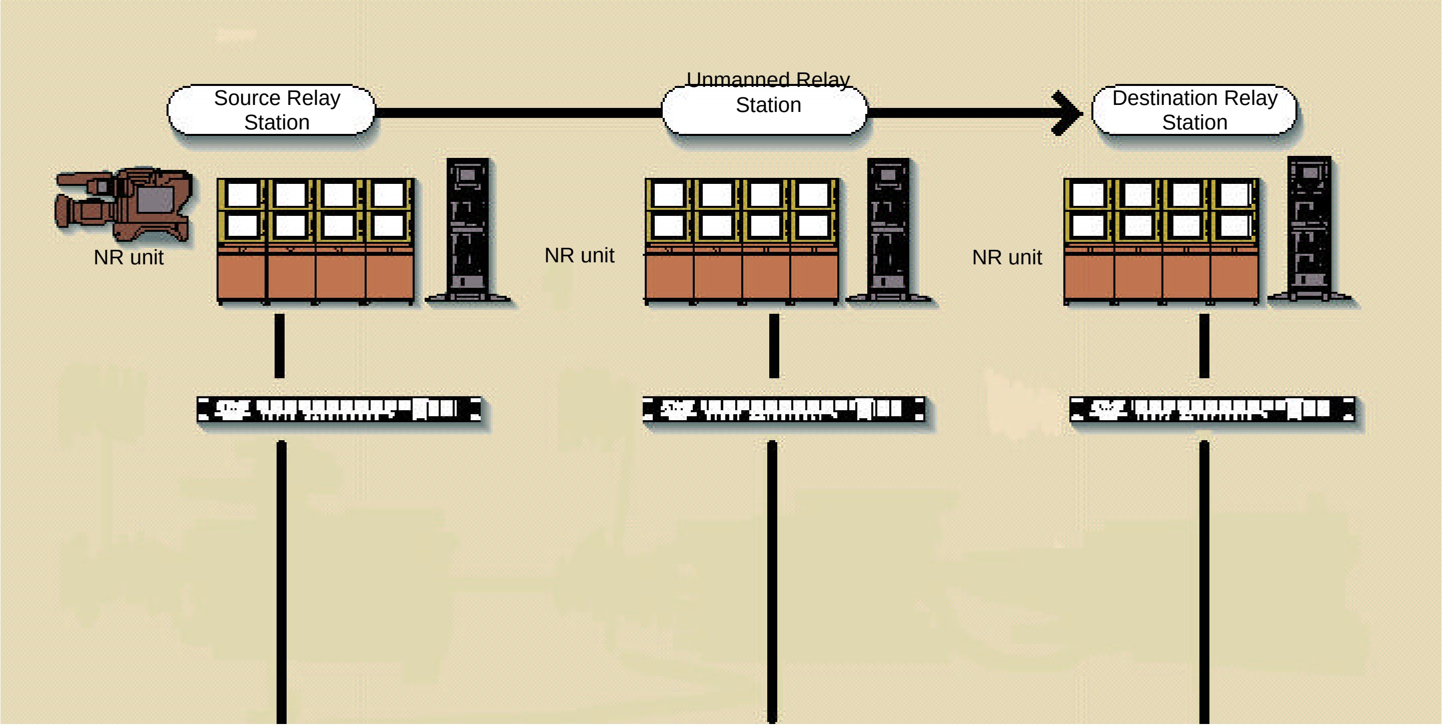


Fig.3 Example of network configuration for Reduced Reference (RR) quality monitor

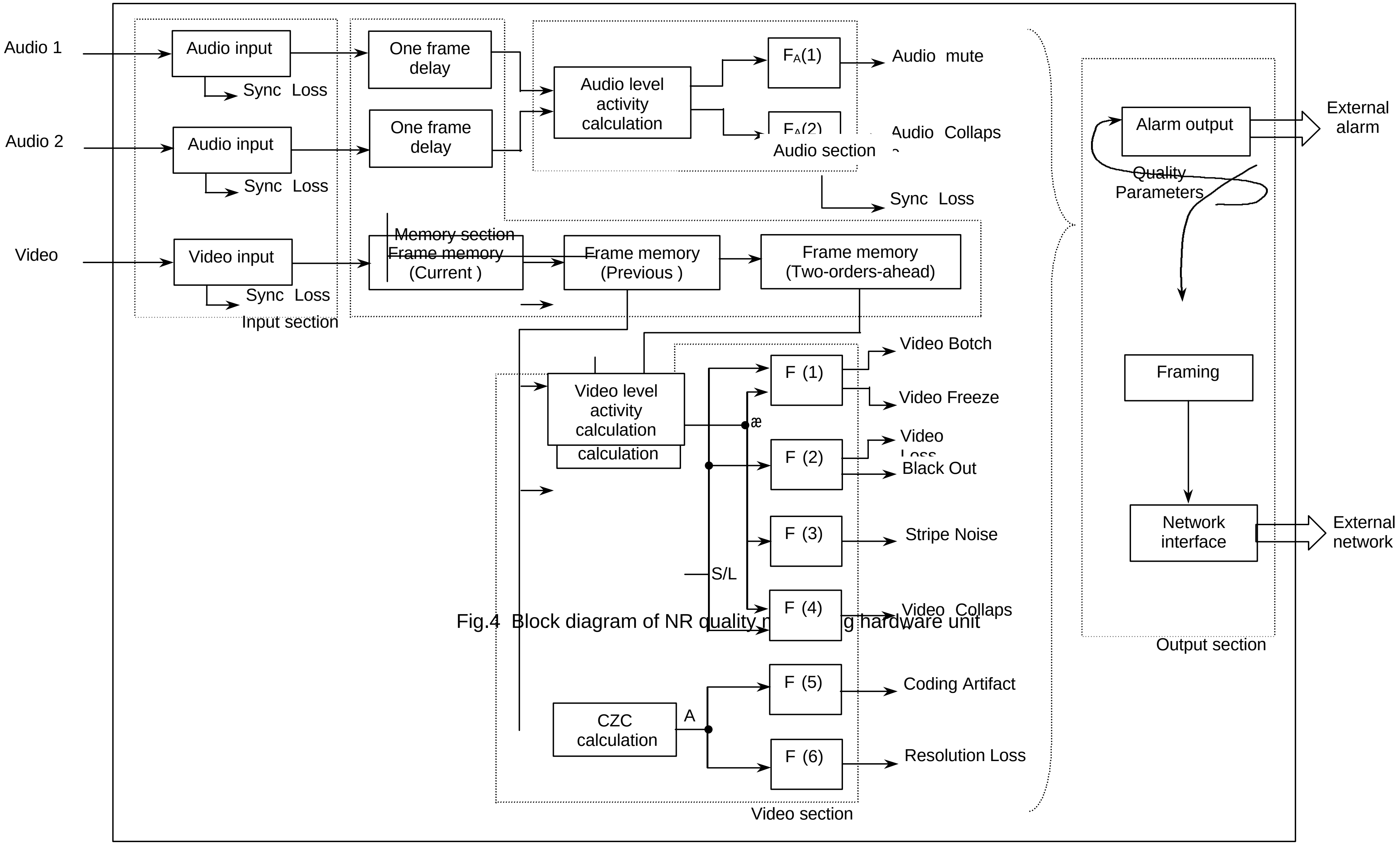


Fig.4 Block diagram of NR quality r g hardware unit

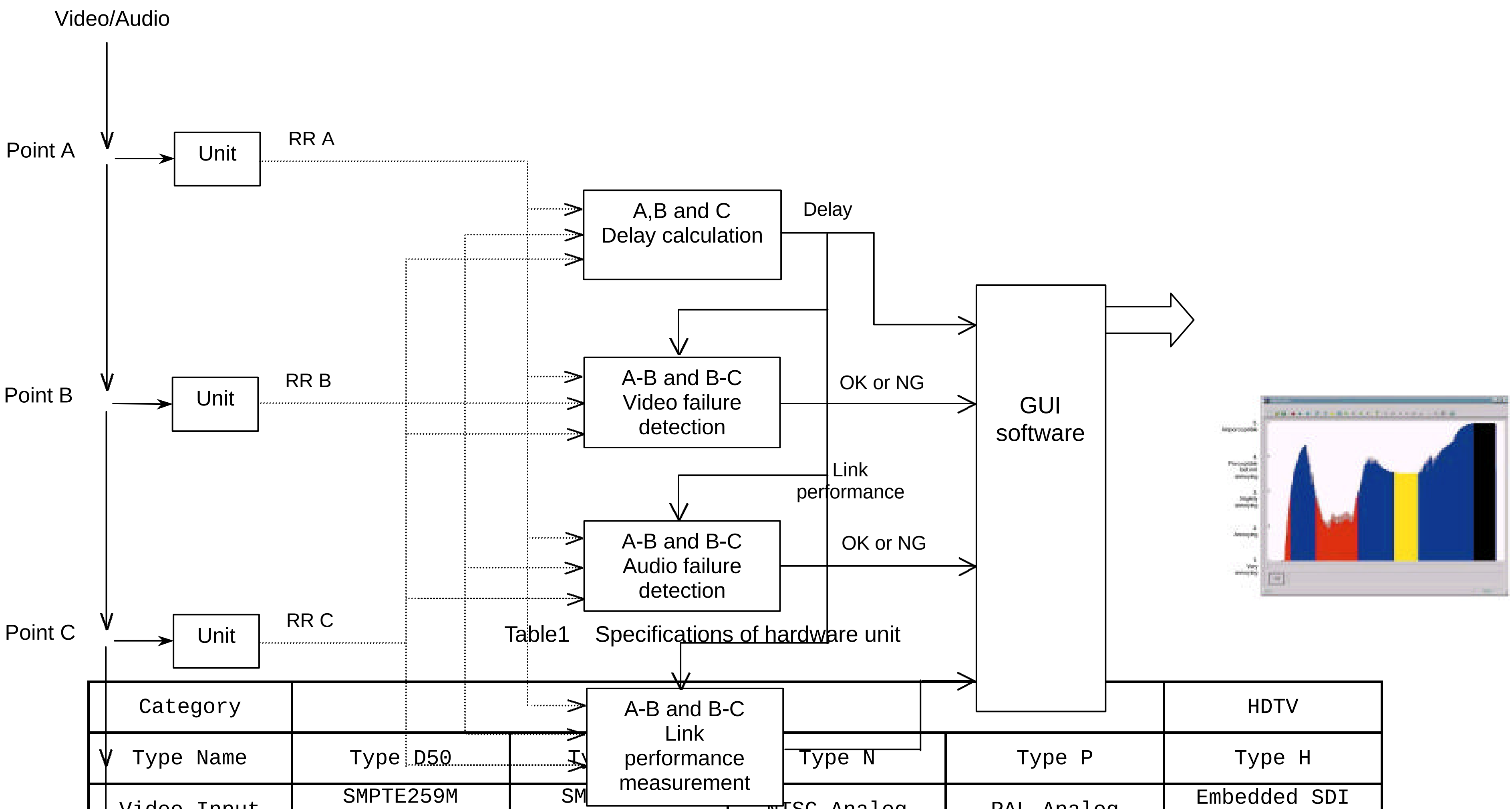


Table1　Specifications of hardware unit

Category					HDTV
Type Name	Type D50	Ty	Type N	Type P	Type H
Video Input	SMPTE259M (4:2:2 –50Hz)	SM (4:2:2 –60Hz)	NTSC Analog	PAL Analog	Embedded SDI (1080)

Fig.5 Block diagram of RR Comparison software

Video Connector	BNC	
Audio Input	AES/EBU Digital	Embedded
Audio Connector	XLR connector (female)	none
Data Interface	Network (UDP/IP) 10BASE-T(RJ45) , Control 4:2:2 Serial interface(DB9)	
Alarm Line	Relay contact output (Dsub-25pin inch) (female)	
Size	430W×480D×44H mm 1U	
Power Consumption	150VAC (100V-240V 50/60Hz)	