Video File System

By Hideki WAKAMATSU*, Itsuo TAKANAMI** Katsushi INOUE**
and Hiroshi TANIGUCHI**

(Received July 15, 1982)

Abstract

We report a realization of an inexpensive system using a microcomputer and a home VTR which can store about 4500 color pictures and its associated comments consisting of maximum 504 characters into a cassette tape. In addition, any desired picture can be searched in practical enough speed under control of the microcomputer which reads the picture numbers and, winds and/or rewinds the VTR in order to search out the desired picture as soon as possible.

1. Introduction

It is needed to be able to file and search a large number of pictures in many fileds, for example, in medicine, CAD, CAI, and so on. We report a realization of an inexpensive system using a microcomputer and a home VTR which can store about 4500 color pictures and its associated comments consisting of maximum 504 characters into a cassette tape. In addition, any desired picture can be searched in reasonable speed under control of the microcomputer which reads the picture numbers and, wind and/or rewinds the VTR in order to search out the desired picture as soon as possible.

2. System configuration

As Fig. 1 shows, the video file system consists of the following components.

(1) A TV camera which takes pictures to be filed.

(2) A monitor TV which displays pictures and its associated comments to be filed and searched out.

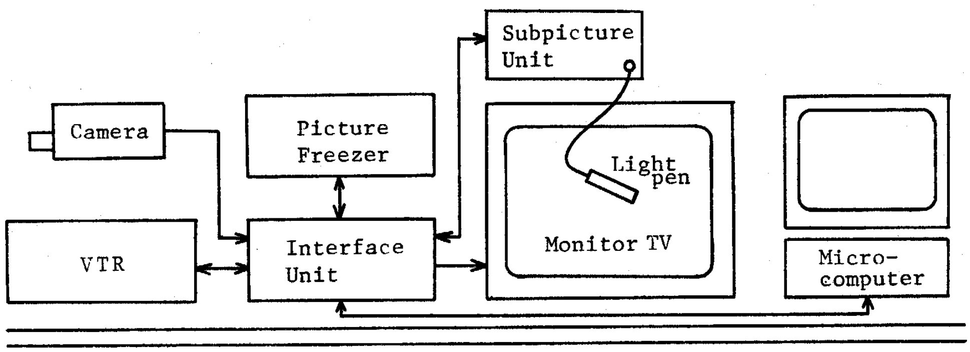

Fig. 1 Video File System.

* Yokogawa Hewlett Packard
** Department of Electronics Engineering

(3) A home VTR which stores a large number of pictures.

(4) A subpicture constructing unit which without injuring a main picture, generates a superimosable subpicture input by the light pen.

(5) A picture freezer which converts analog signals of color pictures searched out from a tape on the VTR to digital signals and stores them into the IC memory.

(6) A microcmputer which controls the whole system. It stores pictures with indices into a tape, detects the indices from the picture data searched out, winds and rewinds the VTR, and outputs a request to freeze the specified picture.

(7) An interface unit which is a hardware to control the VTR and the freezer and, to insert and detect the comments and the picture indices to and from video signal. All the components are connected to the interface unit.

3. Recording format on a tape of VTR

Pictures are searched using the index data which are inserted in part of video signal corresponding to the upper hidden part of the TV screen. Fig. 2 shows the recording

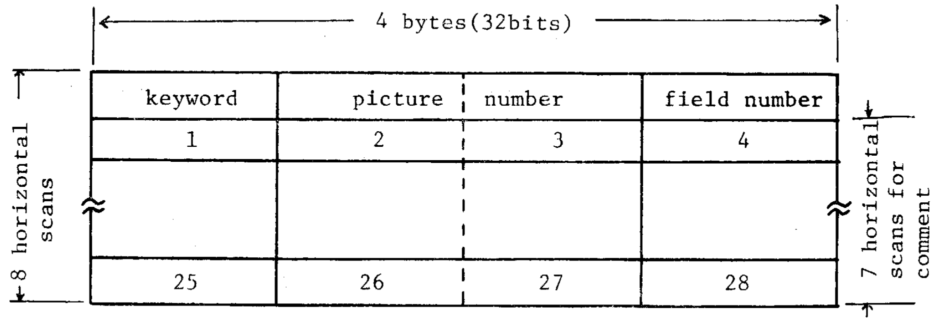

Fig. 2 Recording Format of Index Data.

format of index data inserted in a field of video signal. 8 horizontal scanning intervals in the hidden part are used, each of which has the index data of 4 bytes. Hence, a field contains the index data of 32 bytes, 4 bytes of which are used as information for searching the specified picture. The 4 bytes consist of one byte for making reuse of a tape possible, 2 bytes for the serial number of the picture, and one byte for the field number of the picture. The remaining 28 bytes are used for comment sentences. Since 28 (bytes/field) $\times$ 18(fields) $=$ 504 bytes, the same three contents of 504 bytes(characters) are stored in $18 \times 3 = 54$ fields. The three data are taken by majority and the precise data are restored.

Each field is distinguished from each other by its field number in the index data. Therefore, two fields(i.e, one frame) are sufficient for only reading out a picture from a tape on the VTR. However, in recording pictures, it is difficult to store a picture by two fields in a recording cycle (described in the next chapter) because of the mechanical restriction of a home VTR that is unstable in the rising period. According to our experiment, three or more seconds are needed for a recording cycle in order not to drop quality of the pictures. If several cameras could be used, we could store several pictures in a tape within three seconds.

This gave us a motivation of adding subpictures which explain the main pictures.

Finally, we decided the recording format of the whole data for a picture as shown in Fig. 3. About 220 fields of video signal per picture are stored in a cassette tape on the VTR in a recording cycle, which takes about 3.7 seconds. The main pictures are stored in the front parts of the 220 fields and the subpictures in the rear.

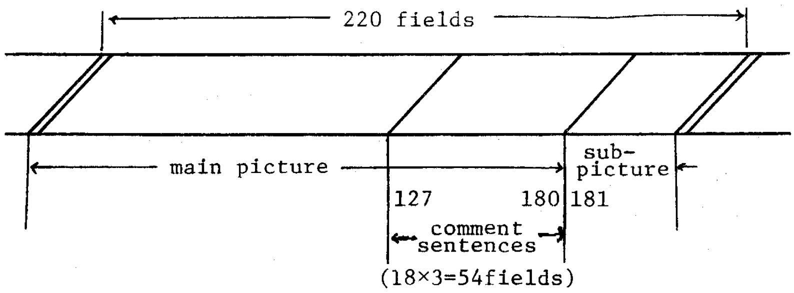

Fig. 3 Recording Format of Whole Data.

4. Operations of system and flows of signals

Fig. 4 shows the flows of the signals to the VTR in a recording cycle and out of the VTR in a playbacking cycle.

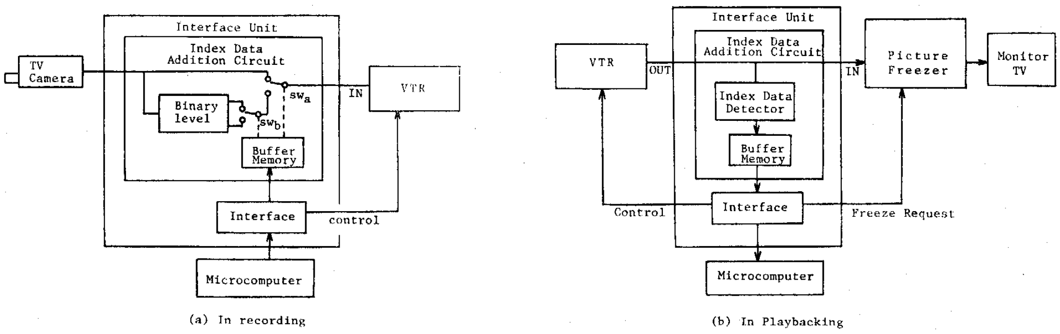

Fig. 4 Flows of Signals.

(a) Recording cycle

The VTR is set to the state of "record". Video signal is supplied to the VTR as follows.

sw_a is switched to the voltage source of binary levels during the period when the binary index data are added to the video signal from the camera. Then sw_b is switched according to 0 or 1 data stored in the buffer memory which have been previously transfered from the microcomputer.

(b) Playbacking cycle

The index data are detected from the playbacked video signal and stored into the buffer memory of the interface unit, through which they are transfered to the microcomputer almost simultaneously. Then the microcmputer knows the picture number and the field number.

Combining the above fundamental operations (a) and (b), pictures are written into and read from a tape on the VTR as follows.

In the case of writing a picture into a tape on the VTR:

(1) The microcmputer puts the system into the playbacking cycle and locates the picture stored at the back end in the tape. Then it tries to find the 200th field of the picture, monitoring the field numbers. On having found it, The microcomputer puts the VTR into the state of "pose".

(2) Next, the VTR is put in the state of "record" and the system is switched to the recording cycle. Then, the pose is released and a new picture and its associated data are recorded during the interval of 234 fields into the tape.

(3) Repeating (1) and (2), each picture of the same field length is sequentially recorded in a tape on the VTR

In the case of reading a picture from a tape on the VTR:

(1) The microcomputer puts the system into the playbacking cycle, reads the picture number from the tape and acertains the present position of the tape. Then, in order to reach the desired position of the tape as soon as possible, the microcomputer selects one of "fast forward", "rewind", "picture search" and the other appropriate opeartions, and executes it for a semioptimal duration estimated from the number of the present picture and that of the desired picture.

(2) The VTR is again put in the playbacking cycle and the picture number is read.

Repeating the above opeartions (1) and (2), the number of the desired picture is found.

(3) Then the microcomputer reads the comments in the index data contained in from the 127th field to the 180th field, and gives to the interface unit an instruction to start freezing the desired picture from the 131st field number.

5. Interface unit

The interface unit is a hardware to play an important role in video filing. As shown in Fig. 4, all the components are connected to the interface unit.

In the interface unit we set a mother board in which the address bus, the data bus and the control bus for the microcomputer, and other signal lines run. The microcomputer, the VTR control board, the freezer control board and all other units, are connected to each other via the mother board.

Index data unit

Fig. 5 shows a block diagram of the index data circuit. As described in Chapter 2, the index data of 32 bytes per field are added to the video signal. Fig. 6 shows the positions where the data are inserted on the video signal. They are the eight horizontal intervals from the third to the tenth for the odd fields and from the fourth to the eleventh for the even fields just after the end of the equalizing pulses. The data are inserted in the 8/11 part of a horizontal interval as shown in Fig. 7 and the video signal is passed for the remaining 3/11 part of the horizontal interval. The effective pictures start at

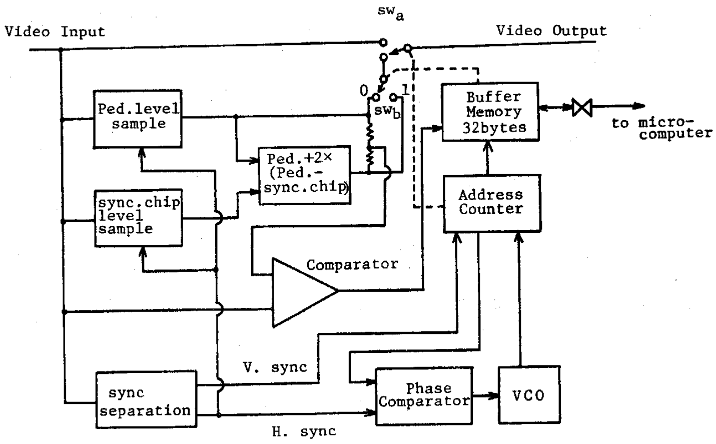

Fig. 5 A block diagram of index data circuit.

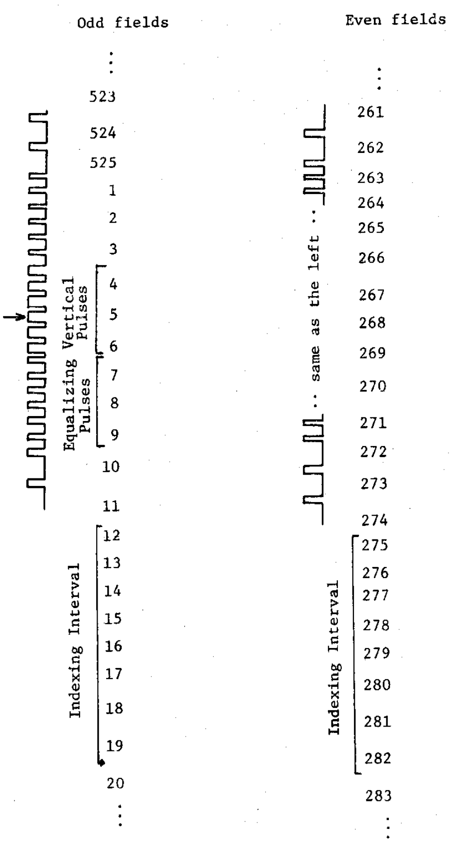

Fig. 6 Positions where the index data are inserted on video signal.

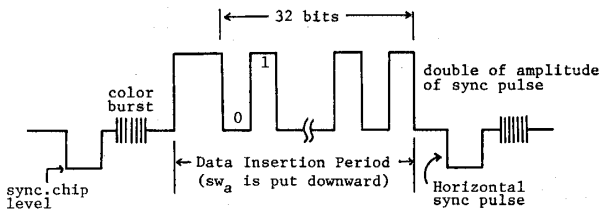

Fig. 7 Period of data insertion.

the thirteenth horizontal interval after the end of the equalizing pulses.

Let A be a pedestal level of the video signal. Then, the low level of the data is A. The high level is A+2×(A−sync. chip level). Since the high level is created by refering to the level from the pedestal to the sync. chip, it can keep the level of video signal almost constant in spite of the amplitude variations of video signal.

Since the index data are detected in the reading cycle by also refering to the level from the pedestal to the sync. chip, the data are detected stably against the amplitude variations of video signal playbacked from the VTR.

The microcomputer communicates with the interface unit during the period except when the index data are inserted or detected. When storing a picture and its associated data in a tape, the communication is done as follows. The index data of 32 bytes to be written into the next field are prepared in the buffer memory. When searching a picture, the data of 32 bytes just read out from the VTR and stored into the buffer memory, are read by the microcomputer.

6. Picture freezer

A picture is recorded in a tape on the VTR for 3.7 seconds. However, the period is only one second when a clear picture is obtained. Hence, a method to display a picture for a long time, is to display a still picture using the pose opearation of the VTR. However, the still picture is not good in quality and since the same track of a tape is scanned by the rotating head for a long time, the tape is apt to be injured.

The picture freezer is a device to realize the above requirement safely. The freezer digitalizes one frame of video signal and stores it into the IC memory. Then the digitalized data read out from the memory is converted to the analog signal and displayed.

6.1 Specifications

(1) Method: to quantize the luminance signal (Y-signal) (dc to 3 MHz), the chrominance signals B–Y and R–Y (dc to 500 kHz) individually, and store them into the IC memory. The NTSC form is adopted on both inputting and outputting.

(2) Sample rate: 9.566 MHz for the Y signal and 2.392 MHz for the chrominance signals.

(3) Quantizing level: the Y signal is quantized by 8 bits as a resultant from summing 4 frames where each pixel in a frame is quantized by 6 bits. Each chrominance signal is also quantized by 8 bits as a resultant from summing 4 fields where each pixel in a field is quantized by 6 bits.

(4) Freezing time: 8/30 second corresponding to 8 frames (4 frames for the Y signal and 2 frames $\times$ 2 for the chrominance signals).

(5) Memory: 256 k bytes for the Y signal ($512 \times 512 \times 8$ bits). 32 k bytes $\times$ 2 for the chrominance signals ($128 \times 256 \times 8$ bits $\times$ 2).

(6) Video input signal: NTSC color composite of from 0.5 to 2 V peak-to-peak. The recommended value is 1 V peak-to-peak. The Impedance is 75 Ω.

(7) Video output signal: NTSC color composite signal of 1 V peak-to-peak. The load impedance is 75 Ω.

(8) Video AGC: two kinds of AGC modes (the peak-to-peak constant mode and the sync. chip amplitude constant mode) are provided.

(9) Digital interface: to transfer the digitalized picture data to and from the host computer. This is executed for the specified parts of a picture by a set of 32×32

pixels. The transfer rate is 100 nsec/pixel.

6.2 Features

The luminance signal and two kinds of chrominance signals are separated from NTSC color composite signal. The signals separated are sequentially A/D-converted and stored into the IC memory. In this process, eight frames (four frames for the luminance and 4×2 fields for two chrominances) are used for quantizing a picture. The digitalized data of four frames for the luminance signal are summed and the data of four fields for each chrominance signal are also summed in each corresponding pixel by the hardware. The summing is expected to decrease random noises in the pictures playbacked from the home VTR.

When quantizing the signals, input levels to the A/D converter are very important. If too small, the maximum allowable level is not effectively used. If too large, saturation occurs. In order to avoid the above situations, the system has two kinds of AGC modes for keeping input levels to the A/D converter constant. One is a peak-to-peak constant mode which keeps the level from sync. chip to white peak constant. The other is a sync. chip reference mode which keeps the amplitude of sync pulse constant. The former stresses contrast of wholly dark pictures automatically.

The picture data of 512×512 bytes for the luminance signal and $128 \times 256 \times 2$ bytes for the chrominance signals can be sent to and/or from the outer computer via the 1 k bytes common memory which can store the data corresponding to 32×32 pixels. The transfer rate is about 100 nsec/pixel.

6.3 Operations

Fig. 8 shows a block diagram of the freezer. In freezing a picture, the digital system is phase-locked to the synchronizing signals in the input video signal.

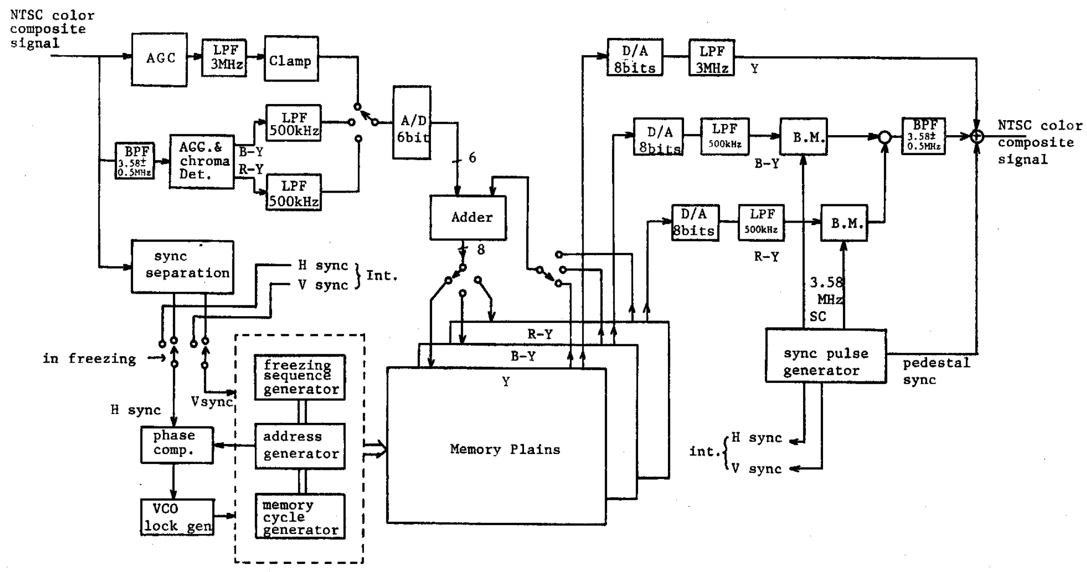

Fig. 8 A Block diagram of picture freezer.

First, the luminance signal (Y signal) is freezed. The first frame is directly written into the memory planes. On and after the second frame, the data at the same memory location as that of the former frames are read out, added to the data of the present frame and restored into the location. This operation is executed by the read-modify-write mode of dynamic RAM. As a result, 6 bits data per pixel are summed up to be 8 bits. The similar process is taken for the chrominance signals.

In defrosting a picture, the digital system is phase-locked to the synchronizing signals generated internally.

7. Discussion

At the present time, the circuit diagrams of all the units are obtained. However, the only interface unit has been completed as a device. Therefore, we tested the fundamental operations of the system using the video freezer of lower grade which had been made previously by us. We used a SONY J9 of β system as a home VTR, and a personal computer APPLE as the microcomputer. A control program is written by BASIC and machine languages so that filing and searching picutures may be done in the conversational style.

In our system, the problem is a time to be taken to search out the specified picture because of a sequential file. We experimented on searching pictures for a file of about 500 pictures. The picture numbers to be searched were generated at random. As a result, any picture was able to be searched within about 10 seconds plus a time to be taken to execute either one fast forward or one rewind. Fig. 9 shows the searching times for several cases.

Since the VTR of β system used can switch from "playback" to "fast forward wind" and "rewind", even keeping head load on a tape, the specified pictures are searched speedily enough by a combination of the fundamental operations of the VTR with the index data inserted in video signal.

In order to search out the specified numbered picture speedily, it is desirable to reach it by one or zero operation of fast forward wind or rewind, if possible. The

picture number		the number of operations				searching time [sec]
present position	destination	F.F	REW	P. S. F	P. S. R	
350	300		1		1	12
300	400	1			1	23
400	500	1				19
500	200	1	1		1	50
200	250	1				11
250	270	1				12
270	272			1		7

P. S. F picture search forward
P. S. R picture search rewind

Fig. 9 An illustration of searching time.

microcomputer computes the duration for which these operations should be executed using an equation. The equation takes account only of the estimated radius of a tape winded around the reels and the inertia at the rising time under the assumption that the angular velocity of reels is constant.

Since the picture number read is decoded under the condition that the picture numbers in the consecutive three fields are identical and moreover the field numbers are increasing one by one, no picture number has been erroneously read.

Since each character in a comment is decoded by majority of three characters which are read separately three times from the same index data of the consecutive fields, errors rarely happen.

Reference

1) Wakamatsu H., Takanami I. and Inoue K., Trial of Inexpensive System to Serach Still Images, Denshi 4 Gakkai Chyugoku Shibu Taikai (1980, 11)