Hands-on experience in harddiskrecording

Harddiskrecording (HDR), that's soundprocessing on a PC with a harddisk as a nonlinear storage media. Such systems are in use in professional studios since over seven years. As the prices for memory chips dropped dramatically in the past, this is now also possible for amateurs. Applications are mainly musicproduction, radioproduction and movie- and videosoundprocessing. Many videoamateurs dream already about their own nonlinear workplace for video: fully digital and a true bargain. So, why to spend money for a pure soundprocessing system if one got already everything with the videoprocessing software? This might be true for your vacation movies, but it's not, when it comes to more complex productions. This article will report about some examples of one of the authors (JF).

I'm working with **tripleDat**, a relative cheap harddiskrecordingsystem for the PC (386 and above) since three years. This system is now available for more than three years. The manufacturer is **CreamWare** in Siegburg (Germany). There have been several updates of the software within these three years. CreamWare was always one step ahead compared to the market leaders Macintosh or complete systems by other companies. This is especially true if one takes the price: this system for 2.400 DM (complete version) can be compared to systems of 150.000 DM and more. The entry version (1.700 DM) offers already everything you want. However, a PC with a fast harddisk needs to be added (if not already present). This system is a special kind of a soundboard: with digital inputs, a midiconnector, and now also two analog inputs. Therefore one may connect a keyboard to the midiinterface, can use up to four simultaneous available recording and playback channels, and gets a fast data-backup-program for the PC on top of that, which uses a regular DAT-recoder as digital storage medium. So, this is definitely more than what a simple sound board can offer for 150 DM.

The program is extremely remarkable: several sound tracks can be displayed and can be played back simultaneous and in real time. This depends only on the power of the PC. (my 486 machine with a *2.7 Mbyte/sec* fast harddisk can mix and play 6 tracks at once, my pentium can handle even 12 tracks.) The display of the waveform is always available without jerks and delays. Cutting, copying, and moving operates with the mouse as ergonomic as in modern word-processing - maybe even more intuitive. The motif is "what you see is what you get". That's it, what makes up effective working with such a system! If you want to calculate effects and just play them back this can be achieved by much simpler programs for regular soundboards. I want to commend **Cool95** as a nice shareware program for this purpose.

Requirements for digital dubbing

Here I want to describe my last two projects. However, three special features need to be explained before, which make film- and videodubbing possible!

i: The HDR-system must be synchronizeable <see "timecodemess">. Most systems can be synchronized via the ***M**idi **T**ime **C**ode (MTC)* which is wellknown in musicproduction or at least via the *Midi Clock*. Some systems even have a separate *SMPTE* input. While filmprojectors can be easily synchronized, this is impossible for video (at least for consumer or semiprofessional technique). The Midi Time Code contains a timecode that gives a number to each frame. The *Longitudinal-Time Code (LTC)* is well known in the studio technique. Converters to the MTC are available. Professional SVHS-recorders can record the LTC on an additional longitudinal track. Though a direct coupling between the HDR and the VCR can be achieved. The *VITC* known at SVHS-machines is even better, however, converters are still exotic and to expensive for amateurs.

ii: In addition the option of in- and output of sound files in the *WAV-format* is very useful. It is even a prerequisite for the cooperation with nonlinear video editing systems like the **VideoMachine** and programs like **Adobe Premiere** or **Ulead Media Studio**. In this case one doesn’t need extra synchron-signals.

iii: A third very useful option offers the tripleDat board: it has an infrared (IR) remote transmitter that allows automatic control of data-backup to the DAT-recorder. This transmitter can also be used for video editing. One only needs to know the right control code. I encrypted this code for the Sony equipment family (download: **remote.zip**). A special analyzing program was planned at CreamWare, but wasn’t needed yet. The remote function is part of the basic set, but not with all operation modes. Though, one can send a remote command at every position. The device is picked up from a table, that contains the command to start the video-recorder in the AudioDub-mode at the right time.

timecodemess

You'll need a time-code in order to synchronize (couple in time) two devices. Otherwise even the most accurate devices will deviate. One device will be the master and give the timing for all other devices.

The most important time-code for linear media (tape- and video-recorder) is the ***SMPTE-time-code***. However, SMPTE stands for ***S****ociety of* ***M****otion* ***P****icture and* ***T****elevision* ***E****ngineers* that is an organization, that developed a digital time-address-code. This time-code can be recorded onto a longitudinal track (tape or video) by frequency-modulation and is then called ***LTC (Longitudinal-Time-Code)***. This signal is also called SMPTE time-code. It can be read only at regular playback-speed - not during wind or even stop-mode.

The ***VITC*** *(**V**ertical **I**nterval **T**ime **C**ode)* contains the same information like the LTC and belongs also to the SMPTE-timecodes. He is specially designed for video, because the signal sits in the invisible vertical burst between the half-frames. Therefore this time-code can be read during fast-wind and stop.

Another form is the ***MTC*** *(**M**idi **T**ime **C**ode)*. It contains the same information too. The Midi- Time-Code can be put into a midi-interface <*see "What is Midi"*> by microprozessors. He serves for sychronsation between several keyboards or effect-devices and the recording device.

The midi-commands have also a ***MidiClock*** command. This is a certain data word, which is periodically emitted. However, this is not more than a pure beat-signal - it contains no time-value.

The ***RCTC*** *(**R**ewritable **C**onsumer **T**ime **C**ode)* by Sony is a time code, that can be written afterwards onto the slanted track of Video8 systems. Some recorders emit it via the control-L remote control. This time-code has almost the same information like the SMPTE-code, but isn't standardized. It can be read during wind and stop like the VITC. -back-

Dubbing with VideoMachine and tripleDat

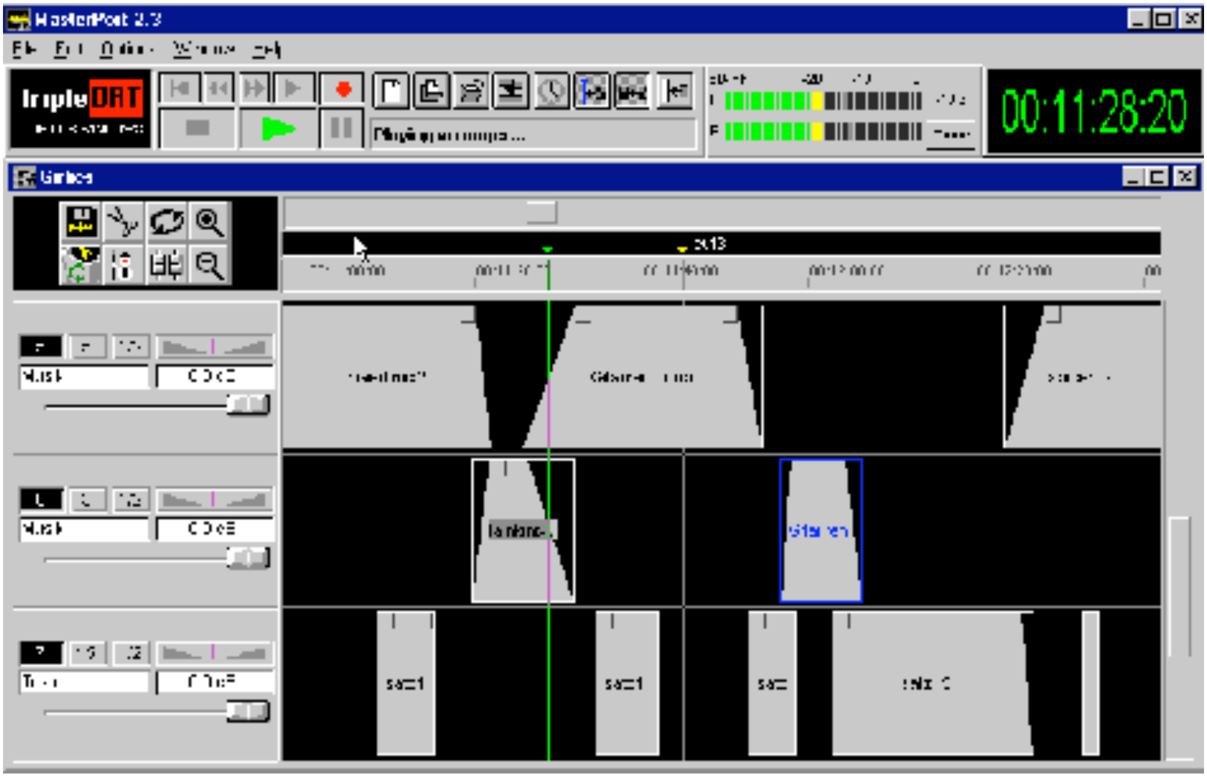
Now the projects: Last fall we made a small video about camera-practice in the zoo. However, the sound was very short. First, the text should be spoken afterwards and second our sound engineer let us down. By using the regular camera microphone, noises made by other visitors can't be eluded. The humoristic texts can't be understood well and though we decided to synchronize our movie. First, we cut the film on VideoMachine with DPR - however, a much simpler nonlinear editing system with Adobe Premiere or Ulead MediaStudio would have done the same job. The only important issue was, that the original sound was digitized with *44.1 kHz* together with the frames and stored as a *WAV-file*. It shouldn't be modified afterwards! This is the only guarantee for a later consistency. This original sound (O-sound) serves as a *master* for the new sound. Although it will never be used, it has all the necessary timing information of the film which can be depicted from the jerks. A *frame and sound flag* were put at the beginning and the end of each take in order to make backloading after the modification easier.

Computer terms and other explanations	
Mbyte/sec	data-transfer speed in ... million characters per second.-back-
ppm	parts-per-million; 50ppm = 0.005% -back-
digital number-stock	16bits are equivalent to a number-stock from 0 to 65535 -back-
WAV-format; WAV-file	this data format for audio-data is common on a PC and supported by all sound-boards. - back *-format- / -back *-file-
44.1 kHz sampling frequency	this sampling frequency is common at CD-players and sound-boards; DAT-recorders use a sampling frequency of 48 kHz because of a concern of pirate-copies (which are nevertheless possible). -back-
export; import	alien files can be written/read by export/import filters and modified into the system-specific format. -back-
backup (media)	(media) for data-storage like computer tapes, DAT-tapes, meanwhile also writable CD-ROMs and
ZIP-drive	floppy-drive with high data density. -back-
reverberation	If a sound is repeated with a short interval, we have the impression of a reverberation. If the interval increases, we call it an echo. If two signals coincide perfectly, we don't hear any reverberation nor echo. -back-
frame- and sound-marker	A test-frame(s) will be inserted together with a test-sound. Wherever the movie ends, the sound has to end abruptly. -back-

The O-sound WAV-file must now be transferred to the HDR workplace. This is easy via [export and import](), if both places are on a PC. If both places are not connected via a network, datacarriers like movable harddiscs, [backup-media](), [zip-drives](), etc. can be used. The only requirement is, that it can be connected to both systems. I'm using an old SCSI harddisc with a movable frame. My friend has a compatible frame on his computer.

I am reading the original WAV-file to my HDR-system and "fix" it first - that is: the soundfile is put to track 1 and stored with the attribute "unmovable". This O-sound-track will not be modified. It serves as my reference and can only be played back. I can hear all cuts by the jerks on this track. It is a meter for the timing of the movie. If a question about the correct frame arises, I can take the start-flag from the picture, count the time with the aid of a Tape counter, and compare it with the timing-scale on the HDR-system.

Track 2 contains a copy of the O-sound. I'm using only the animal sounds, the atmosphere and the sound flags, all the rest is made silent. Track 3 has some looped parts from the background. This was necessary in places where the O-sound was unusable, and neither music nor our speaker could be put. The 4th track holds the music - that wasn't much, because the text and pictures had priority. Within an arrangement only one sample-frequency could be used. The music was taken directly from CD in most cases, therefore the whole arrangement had to be sampled with 44.1 kHz. That is, the text had to be read with just this sample-frequency. However, not every DAT-recorder can do this. Though, several possibilities arise: first, tripleDat comes with conversion program, and second, several construction sets (Elektor 10/96) are available. And, the analog-to-digital converter of the videodigitizing board can be used.



The Soundfiles will be placed at the timeline with the Arranger

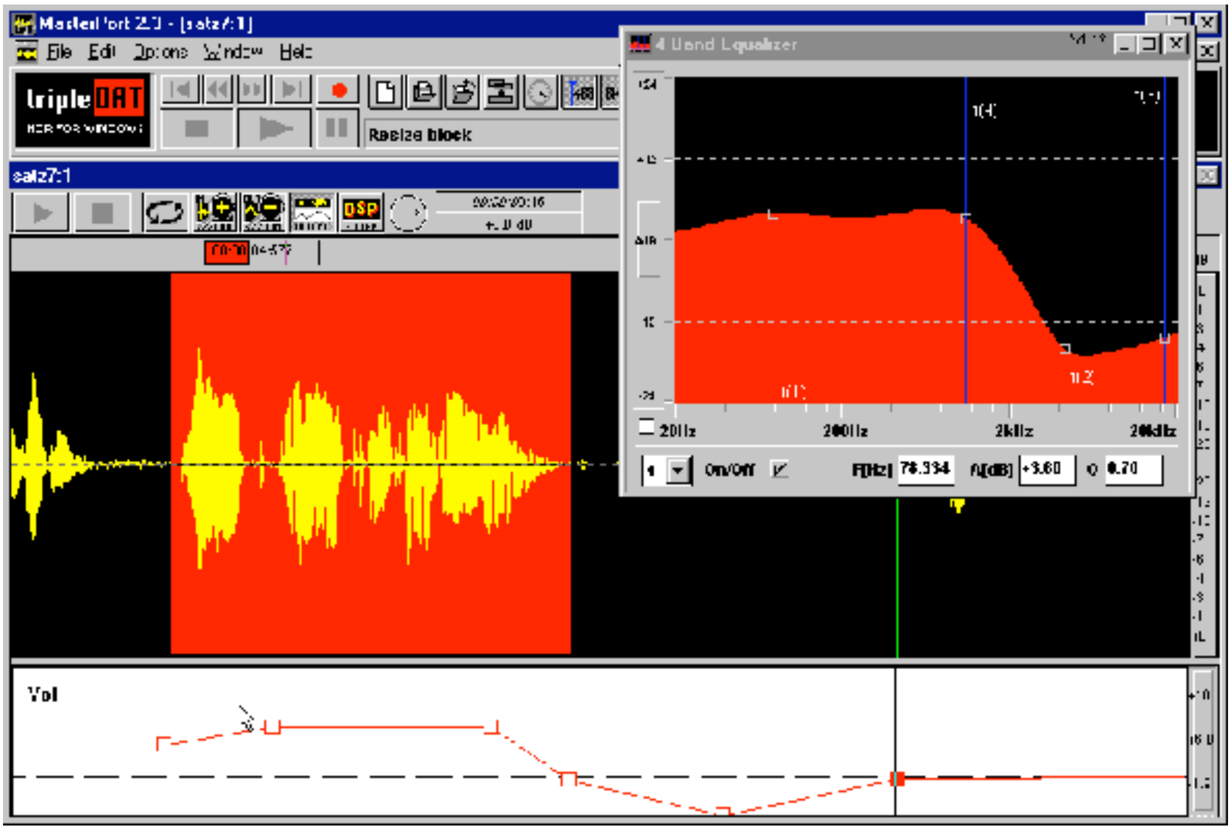
How to dub the sound? This is quite simple, if no translation is required: all you need is a microphone and a headphone. The textsegment will be selected in the O-sound, repeated in a loop and played back via the headphones. The new sound will be recorded on the DAT-recorder or the videorecorder. One should try to speak synchrone to the text from the headphones (as if you are singing). This might require some practice, but after a few trials you will get a satisfying result. The acoustics of the room may raise some problems, because the original place isn't the sound-editing place. The aid of a partner might help, if you move to a balcony with the microphone. A strongly damped room is generally better when dubbing the sound (e.g. you put blankets around the speaker). *(If you plan a commerchal project, contact a local Soundstudio. They have all the Equipment for reasonable costs.)* A reverberation can be produced electronically very easy (tripleDat comes with several options), but it's impossible to remove a reverberation by electronic means. Soundeffects on a PC are available for 1/1000 the price you have to spend for effect-devices (Cool95, SoundForge, etc.) and, of course, are part of the tripleDat programs - now even in real-time! The new sound-passages will be read to the HDR and, if necessary, modified to fit the correct sample-frequency. In order to position the new parts correct, the O-sound track and the dubbing-track will be played simultaneous (either together or one on the left and the other on the right channel). The new part will be shifted in time until no reverberation-effect can be heard. This should be possible for amateurs using their own voice. If the best fit is found, don't forget to protect it against inadvertent shifting!

What is midi

Midi is an interface between electronic music instruments. The midi-interface is standardized and vast spread. It can also be used as a connection to a computer. Every sound-board now has an midi-interface. Often you need a special cable as an accessory. The midi-interface emits commands like: "Sound on/off, pitch, vehemence", but also time-information. The midi-interface is galvanically decoupled (protection against high currents to Ground and protection against hum-noises). The control of the midi-interface is so complex, that only a computer can manage it reasonable. If you restrict yourself to only a few functions, this might also be possible by simple electronics. -back-

If all the text is managed this way, it's worth to listen to the whole within the context, before you'll go into the details. The main part of this will be to find the correct recording level. The right recording level is the most important criterion when working with the digital-technique. Once all the pile of *digital number-stock* is used, your nice sound will be scrambled. On digital systems this limit is reached at a recording level of 0 dB, there is no spare buffer of 3-6 dB like on audio- or videotapes! Furthermore one should comply, that a mixture of two soundtracks of equal level yields a total level that is 3 dB larger. The built-in effects like compressor, limiter, etc., are explained in the HELP window. They can achieve good service in order to find the right level correction. Empirically a level of 6 to 9 dB below maximum is fine for the single tracks, this results in a spare of 3-5 dB for the final result.

If the arrangement is done, it is useful to mix all to a new WAV-file and read it back to the digital video-cutting-program (without new digitalization but via datatransfer). The new mixture will now be placed onto a separate audiotrack. As the cut-programs all have a graphical display, one can compare the new mixture with the O-sound and move it to the right position (with the help of the zoom or magnification option). As a control, one should do a sample run and care for echoes of the sound marker. If necessary, the position can be altered. With this method, the level at the end of the videosequence should be equal - the number of the samples between the markers did not change.

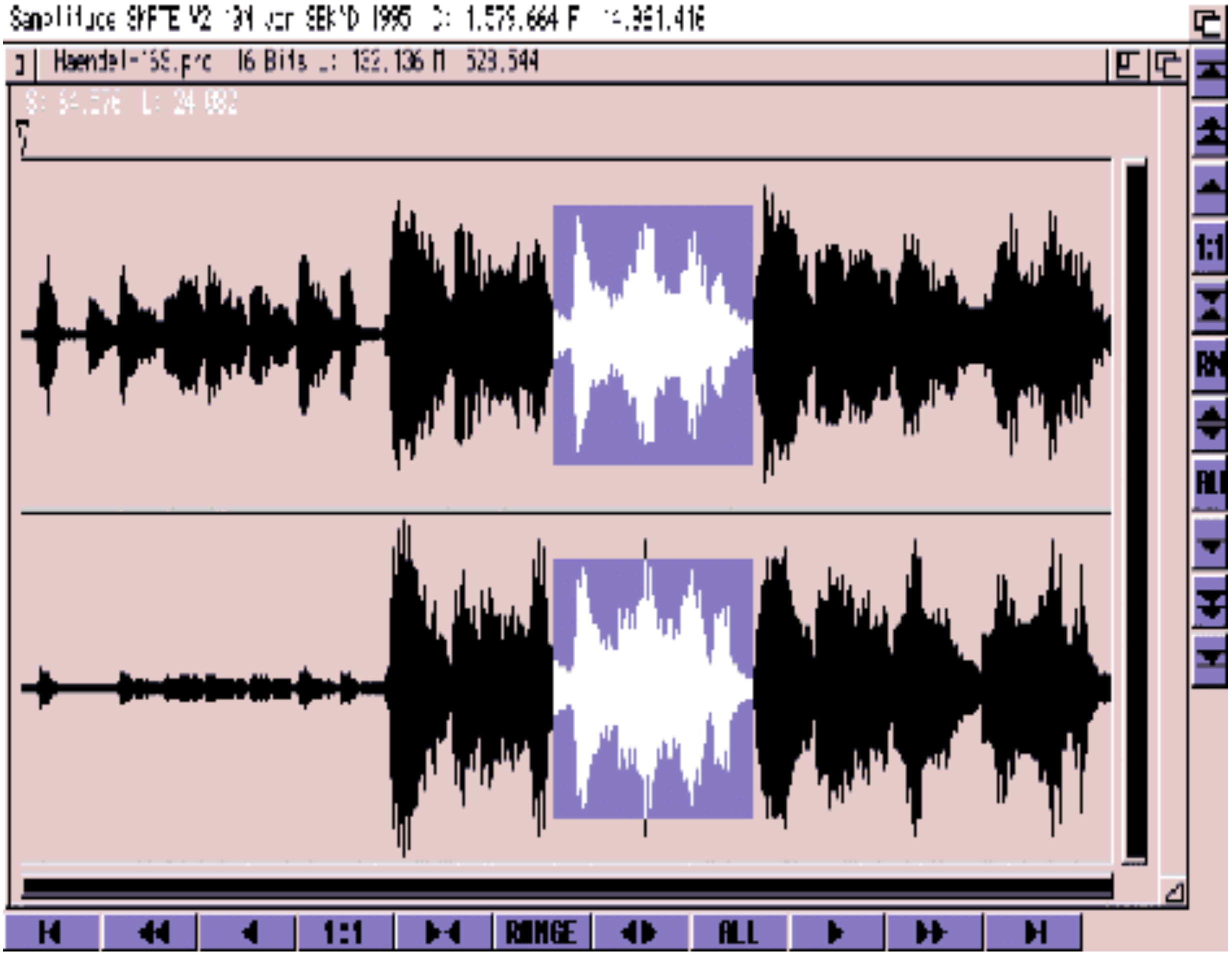


With the Editor you can cut out a whisper. In the left is the 4-Band-Equalizer shown.

The result compensated for our ostentation with this article. It was almost a shame that nobody realized that nearly everything was synchronized and which terrific technique we used!

Two Amiga systems

I like to present here a different procedure with a HDR-system that runs on an Amiga. The system consists out of an AD-converter with SMPTE-input **AD516** by **Sunrise Industries**. It has a programmable mixerinput that can mix the O-sound with stored data in real time. The HDR will be used in conjunction with a professional SVHS cutting side. The already mentioned VITC-SMPTE-converter **TC30RIV** by **Appermann&Velte** will be used to synchronize between sound and frame.



The Amiga-Version of Samplitude: A Song will be adjusted in the duration with a pop-free cut.

The dubbing works like this: the hi-fi sound of the finished video (or the sequence) will be put into the live-input of the HDR-board. If the VITC starts at 1h30min and takes 20min, the base-time of 1h30min must always be subtracted by the HDR-system (SMPTE-offset). If one starts the VCR at an arbitrary position, the belonging sound of the arrangement will be fetched and played back. This time-code couple's sound and frame perfectly. This way the soundmix can be done live during the audio-dub mode of the VCR - with a professional SVHS system even in DolbyStereo.

Another system is available for the Amiga and the PC: **Maestro Pro** by **Macro Systems** comes with the program Samplitude. This is a purely digital cutting-board, that can again be controlled by the *SMPTE-Code (LTC)*. The system has extra effect-modules, that allow for modifications of the sound. Both systems work reliable and their combination is in my opinion the ideal Amiga based system for SVHS sites. However, the VITC-SMPTE-converter is very expensive.

There are no* converters for Sony's TimeCode

I can't use those blessed developments like the *VITC*, because I decided for the Sony appliances (Hi8). The *RCTC* used by Sony is technologically even more interesting; however, I don't know of a system, that converts it to the Midi *Time Code (MTC)*. Theoretically it is even simpler, because the *RCTC* is part of the remote-control signal and not the videosignal. On the other hand, with a Hi8 recorder (at least with another copy) a VITC can be mixed into the videosignal. In that case I could work with my tripleDat likewise comfortable and economically. Unfortunately there is the problem, that the *SMPTE-Code* needs an extra conversion into the *MTC*. But, I'm not eager with another copygeneration!

(*: AX-Systems has a complete HDR-System: AudioMachine-lite with RCTC-Input. Now Roland announced a Sync-Interface! -back-)

Lip-synchrone stereo dubbing on the PCM-track

With my last project - the complete soundtrack for a movie on DVC - that wouldn't have worked out. You cannot save the burst-signal (which contains the VITC-signal) while compressing the videodata. Luckily tripleDat offers another possibility for lip-synchrone dubbing.

The complete movie has a length of about 20 minutes and contains several parts that need for exact synchronisation. The author had a good eye and recognized dubbing-errors of only one frame! Beforehand a small example: the quartz of a consumer DAT-recorder has a frequency-tolerance of ± 50 *ppm*. If the quartzfrequency is unluckily at the upper border (e.g. because of a to high temperature), with a sampling-frequency of 44.1 kHz this causes after 6min40sec a deviation of one halfframe (1/50 sec).

This example explains the synchronisation-problem very drastically. The problem can be defused by using the same quartz-beat for recording and playback. The error averages away and only tolerance-variations between recording and playback enter into. That isn't a problem with only one DAT-recorder, but when playing back from the harddisc the HDR-system gives the timing. TripleDat offers a solution: the timing can be gained back from the digital-input of the DAT. The converter of the tripleDat-board must be used for playback in this case.

Let's return to my last project: it was a reportage with interviews, two announcers, live atmosphere, and some newspaper clippings. The whole movie took about 20 minutes. At first I read the O-sound into my DAT-recorder (right away with a sampling frequency of *44.1 kHz*), and positioned it. This was only used as a timing reference. Track 2 got a copy of the O-sound, of which the volume was adapted and all undesirable noises were removed. The next track contained the cut and modified textpassages of the two announcers. Three more tracks hold the music. Just here you can find an advantage of modern music, that was computer-recorded: with some skill the length of the music can be fit to the desired length for the movie (by some loops). We even shifted one title, because the instrumental solo was better for the beginning. The sound of one announcer got some modifications and reverberations by the effect-modules. These kinds of modifications can be achieved in a few minutes, even if it must be repeated several times in order to find the right sound. Well - I won't conceal that the whole job took a while. The work on 1.8 Gbyte of audiodata scattered over 10 tracks lasted about 18 hours.

My procedure with tripleDat

The huge shiver came after the last mix - does everything stay synchrone?

The whole procedure runs with the already mentioned features of tripleDat. At first the tripleDat-board is synchronized via the optical-digital intersection onto the pulse from the DAT-recorder (which was also used for recording). The Sony DVC-recorder waits in the pause state on AudioDub-PCM-sound. When the playback of the last mixing was started via the internal converter, at a given mark the infrared code for "Pause" will be broadcasted. Though, the Hi8- and DVC-recorder in stereo mode on the PCM track can be started reproducibly. By this procedure I achieve synchrony for more than half an hour.

Postsynchronising in the livingroom in front of a heavy curtain

Nevertheless, the first result was tragic: during recording everything seemed to be synchrone, but when playing back there was a clear shift. The solution to the enigma was a constant shift of 2 frames with the PCM-dubbing. After recognizing this scheme, a perfect dubbing isn’t witchcraft nor a result by chance.

development tendencies

Meanwhile the development didn't remain standing. Professional systems now work with 24 bits (this gives a theoretical range of 144 dB - from the hearing threshold up to a jet-engine in 1 meter distance, and all in hi-fi stereo). More interesting for us is the other direction: the light version of a HDR-system is offered for 100.00 DM, or even less in a bundle with a soundboard; digital converter boards are available for 300 to 800 DM and probably less in the future. And with some chance the RCTC-MTC-converter will be up for sale.

There is an announcement from Roland for a Video-MIDI Sync Interface witch accepts Control L input and sends MTC! The Name is SI-80S. This seemed to be very interesting and might be the solution! -back-

conclusion

Even if you own a digital, nonlinear videocutting-system, a pure sound-workplace isn't a wrong investment. This is especially true, if you consider that with high-merit video-work the movie-portion is about 10-20 times bigger than the sound-portion (2-4 *Mbyte/sec* for the frames; 200 *kByte/sec* for the sound). A program for videocutting is intended for the processing of videodata and therefore not optimal for sound-processing. You wouldn't expect a textprocessing program to be good in picture processing - it is enough, if you have compatible in- and output formats. If you take 2.000,00 DM for the HDR-system, 3.000,00 DM for a Pentium-PC with a huge harddisc and about 1.000,00 DM for a DAT-recorder, you'll have a high-end soundprocessing-system for 6.000,00 DM. This system will be much better than a homestudio with 3 Revox taperecorders, a mixing-unit, several effect devices, and a synchrone-appliance for perfo-tapes. Especially for video-hobbyists, that work in team with like-minded persons, A HDR-system might be an alternative investment. Together you can work on one person's movie and the other person's sound and split the expense for the technology.

Although I have taken my examples from video-processing, the sound-processing options will be mainly used for professional and semiprofessional movies.

Most of the described products are now not available!
