

Audio Processing Method, Audio Processing Apparatus, and Recording Reproduction Apparatus Capable of Outputting Voice Having Regular Pitch Regardless of Reproduction Speed

Patent No.6,360,198(U.S.), Patent No.3,073,942(JP), Patent No.303913(KR), Patent No.ZL98801333.9(CN), Patent No.2271463(CA)

This patent pertains to the invention of an audio processing method and a recording reproduction system that are capable of outputting sound with a normal pitch regardless of the reproduction speed of the recorded audio data. The method and system especially provide high-quality audio synchronized with video data at any desired reproduction speed in an application to a VTR apparatus.

Conventional audio reproduction at a speed that differs from the original recording speed, such as in slow-motion playback or fast-forwarding a VTR apparatus or other equipment, usually results in limited listening comprehension due to the reproduced sound having a fluctuating pitch in proportion to the reproduction speed. This new invention realizes a high-quality sound output with stable sound pitch, by synchronizing the sound presentation timing to that of video data obtained by changeable-speed reproduction.

The figure 1 shows a system overview for a sound-equipped VTR changeable-speed reproduction apparatus that employs the new audio processing method. By controlling the VTR changeable speed reproduction (part 2) for sound signal reproduction at a speed different from that of the original recording speed, this apparatus performs an A/D conversion of the audio signal at a sampling frequency of f_i , which realizes $f_i = r f_0$ (Hz), if $f_i / f_0 = r$ can be satisfied (the f_i (Hz) is the sampling frequency at A/D conversion, f_0 (Hz) the sampling frequency at D/A conversion, and r speed ratio of the VTR changeable speed reproduction part). If $f_i / f_0 \neq r$, the apparatus performs a conversion into audio data sampled at a sampling frequency conversion coefficient $c = r f_0 / f_i$ (Hz), using its sampling frequency conversion (part 4). It analyzes and divides the audio data based on an attribute possessed by the audio data, such as voiced sound, unvoiced sound, or silence, to divide it into blocks, each having a prescribed time width. A D/A conversion is performed with this data block unit at a sampling frequency f_0 (Hz), after successively using a ratio of $1/r$ for the data time width, through interpolation or thinning out, if necessary. This made it feasible to take sound that has had its pitch changed through a changeable speed reproduction at a speed that falls within a practical use range of changeable speeds, such as a reproduction speed of $1/2$ or 2 times, and restore its original level of sound clearness within a range of ± 1 octave in a state where the sound is synchronized to a highly accurate degree with the picture.

Regarding this patent, the speech rate conversion technology is implemented in a radio receiver (Figure 2). The system is useful for not only elderly listeners, but also for people listening to foreign language broadcasts. Thus, it contributes to the uses for a wide range of broadcast services.



Figure 2: Radio receiver using speech rate conversion technology

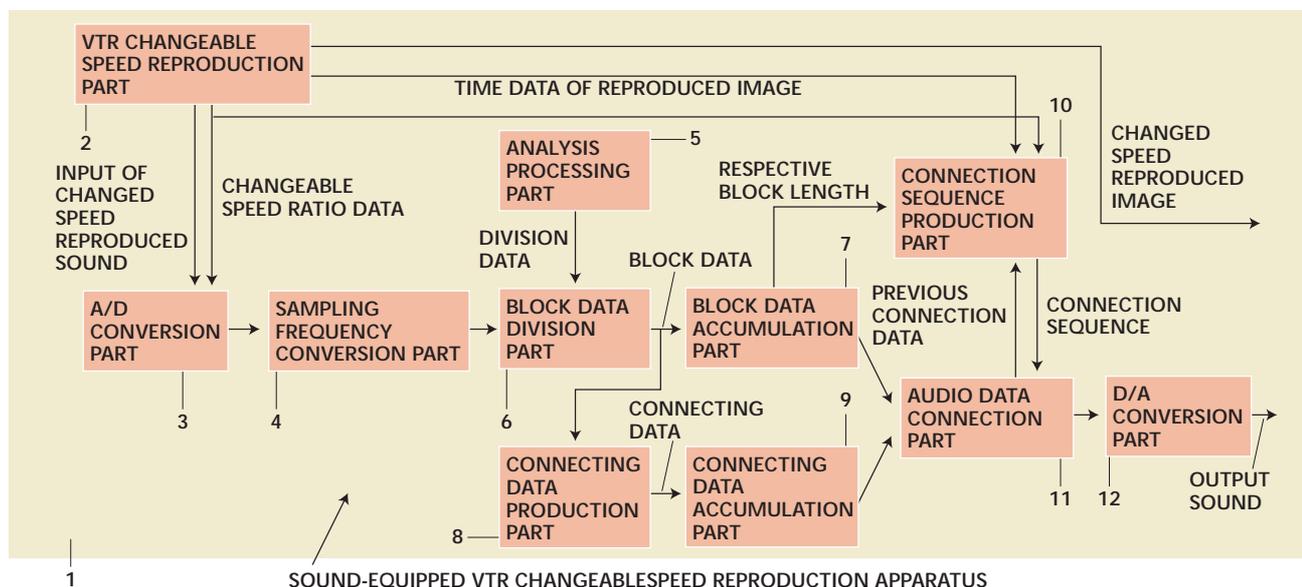


Figure 1: Block diagram of sound-equipped VTR changeable-speed reproduction apparatus