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

This invention relates to an audio processing method, audio processing apparatus, and recording and reproduction apparatus that enables sound of a normal pitch to be output regardless of the reproduction speed from a commercial-use VTR, 6-mm tape recorder, or the like, of which the pitch of output sound would otherwise change in proportion to the reproduction speed.

The pitch of reproduced sound that has been recorded in an analog recording medium such as magnetic tape normally changes in proportion to the reproduction speed of the apparatus. Thus, it has been impossible for a VTR to reproduce sound simultaneously with images that are to be reproduced in slow motion without that sound appearing to have a non-realistic pitch.

The changeable speed reproduction apparatus 1 (Fig. 1) modifies an audio signal that is to be reproduced at a speed different from that at the time of recording. It divides the audio data into blocks, each having a prescribed time length, and if necessary, performs interpolation or thinning thereof, according to a changeable speed ratio as determined by the VTR changeable speed reproduction part (2) and the sound attributes. The sampling frequency conversion part (4) matches the sampling frequency $f_i$ (Hz) of A/D conversion to the sampling frequency $f_o$ (Hz) of D/A conversion. Consequently, high-quality sound with no change in pitch thereof is output in synchronization with the timing of the image signal from part 2.

In particular, A/D conversion of the audio signal is performed using the relation, $f_i = r \cdot f_o$ (Hz), when the sampling frequencies $f_i$ and $f_o$ (Hz) satisfy $f_i / f_o = r$. When $f_i / f_o \neq r$, because $f_i$ and $f_o$ cannot be set to the given values, the audio signal is converted into audio data whereby part 4 does sampling using a sampling frequency conversion coefficient $c = r \cdot f_o / f_i$ (Hz). The series of sound-attribute analyses on the audio signal is performed to divide the audio data into blocks, each having a prescribed time length, and if necessary, the data are interpolated or thinned in units of a block to lengthen or shorten data by $1/r$. After that, D/A conversion of the audio signal at $f_o$. The sound output from this procedure will have no change in the pitch compared with the signal input to part 2, yet will be synchronized to the timing of the image signal output from part 2.

Figure 1