

リアルタイム話速変換装置とその応用  
—デモンストレーション—

今井篤、清山信正、都木徹、宮坂栄一

NHK放送技術研究所

〒157 東京都世田谷区砧1-10-11

E-mail: (imaiatsu, seiyama, takagi, miyasaki)@str1.nhk.or.jp

あらまし 高齢者などにも優しい音声放送サービスを目的として「話速変換装置」の開発を進めているが、本来の目的に加え、歌声や楽音への適用、また放送機器への応用について紹介する。例えば、歌声や楽音などのテンポを任意にマニピュレートすることにより、新しい効果が得られる。また、放送機器への応用として、VTRの変速再生時に、映像と音声を完全に同期して0.3倍速から3.0倍速まで原音と同じピッチで高品質に再生するシステムを紹介する。高速再生時の付加機能として、呼気段落の冒頭のみ「ゆっくり」に変換し、「ま」（休止部分）を効果的に短縮することで、音声を聞き取りやすくする手法についても、実際のデモを交えて紹介する。

キーワード 話速変換 映音同期 ピッチシフト 放送機器 エフェクタ VTR

A Speech-Rate Converter and its applications.

Atsushi IMAI, Nobumasa SEIYAMA, Tohru TAKAGI and Eiichi MIYASAKA

NHK Science and Technical Research Laboratories

1-10-11, Kinuta, Setagaya-ku, Tokyo 157-8510, Japan

E-mail:(imaiatsu, seiyama, takagi, miyasaki)@str1.nhk.or.jp

**Abstract** Some applications are demonstrated with the Speech Rate Converter. The feature of this system is to change speech rate as desired on real time, with invariance in pitch. We introduce some trials to apply this system to singing voice or musical sounds as well as to the broadcasting equipments. A new application of this system to a broadcasting VTR system is also demonstrated. It can maintain the pitch of audio signals when playback speed is changed between  $\times 0.3$  and  $\times 3.0$ . For rush watching of the contents, the converted speech can be easy to heard. This effect is produced by both making speech rate slower at every onset of a breath group and shorting pause interval between the breath group.

**Key words** Speech rate converter, AVsynchronous, Pitch changer, Broadcasting use, effector, VTR

## 1. はじめに

これまで開発を進めてきた話速変換装置を、歌声や楽音へ適用した事例について紹介する。さらに、本システムを、業務用VTRの可変速度再生時のピッチ補正方式に応用した事例についても紹介する。話速変換方式については、実用化に即した改良<sup>1)</sup>により、標準化周波数の上昇(32KHz)をはじめ、処理に伴う遅延時間の短縮、話速を大幅に変化させた場合の自然性の向上、話速変換倍率と変換された音声の伸張幅の正確な対応などが実現された。これにより新しい用途への可能性が広がった。

## 2. 話速変換方式を応用した放送用VTRの可変速再生システム

今回、VTRの可変速度再生時に原音声のピッチを保持し、速度制御系とリンクしたリアルタイム映音速度制御システムを試作した。図1にVTR可変速度再生時のピッチ補正方法の概略を示す。

VTRの音声出力は、再生速度に比例してピッチの変化するアナログ信号を用いる。

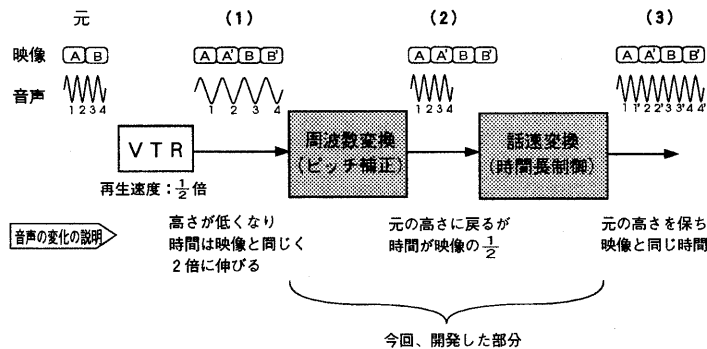


図1 VTR可変速度再生時のピッチ補正方法

### 2-1 本アルゴリズムの特徴

1: 大幅にピッチを変化させた場合の音質の保持

放送で実際に使われているVTRの可変速度再生倍率は、例えばスポーツの決定的瞬間などは0.3倍速程度で再生されることがある一方、編集前の素材テープの内容試写では2、3倍速、またそれ以上で再生される場合もある。これに伴うピッチ変化は、カラオケなどで一般に用いられているピッチチェンジャの補正

範囲を大きく上回る値である。現在、 $\times 0.3 \sim \times 3.0$ の任意の値で設定可能である。

### 2: 映像の再生倍率と音声の伸張倍率の高精度な対応

映像の再生速度変化タイミングに音声のピッチ補正が逐次応答するためには、映像の再生倍率と音声の伸張倍率が高精度に対応する必要がある。本装置では、処理に伴う遅延時間の短縮に加え、入出力音声データの時間監視を常時に行なうことで、映像信号のタイムコードなどを利用すること無しに高精度な映音同期を実現している。

### 2-2 装置の機能

以下の機能を独立に、または同時に実現。

1: VTRの変速再生時にピッチの変化しない高品質の音声の再生が可能

可変速度レンジ:  $\times 0.3 \sim \times 3.0$

2: ピッチの微調整 1のピッチを  $\pm 25\%$  以内で微調整可能。且つこの微調整は、再生速度と連動して逐次変化させることも可能

3: 呼気段落の冒頭のみを「ゆっくり」にし、適宜「ま」を短縮することで、時間的な伸張を抑制しながら音声の聞き取りを補助する機能を付与  $\rightarrow$  2~3倍速の早見 (VTR編集素材のラッシュ試写など) に有効

## 3. おわりに

話速変換の応用技術の一端について紹介した。VTRシステムについては、視聴覚に係わる研究用ツールとしての応用も考えている。

## 文献

1: 清山、今井、都木、宮坂 "リアルタイム話速変換装置の実用化について" 音講論1-P-19, pp.341-342 (1997)

2: 今井、清山、都木、宮坂 "放送用VTRの可変速度再生用音声付加方式の検討" 音講論 3-P-17, pp.613-614 (1997)