

K-060

番組台本を利用した解説放送用原稿作成支援システム Manuscript production support system for audio description using a TV program script

加藤 直人[†] 清水 俊宏[†]
Naoto Kato Toshihiro Shimizu

[†]NHK 放送技術研究所
NHK Science & Technology Research Laboratories

1. はじめに

解説放送とは、主に視覚障害者を対象として、テレビ番組のナレーションやドラマのセリフとは別に、番組映像を音声で説明する放送サービスのことである。解説放送付きの番組は増えてはいるものの、その数はまだ少なく、拡充が望まれている。

解説放送原稿の作成では、解説文を挿入すべき箇所（タイムコード）を探し、情景描写をするという作業を行う。≠切までの時間的制約も厳しい上に、独特のノウハウも必要とされる。[1]

本稿では、このような原稿作成作業を効率化するために開発している、解説放送用原稿作成支援システムについて述べる。本システムでは、これまで独立であった番組映像、番組台本、解説文入力を1つの作業環境として計算機上にまとめることにより、操作性の向上を図っている。また、「解説文候補の提示」機能も備え、解説文作成を支援する。

2. 解説放送用原稿作成支援システム

解説放送用原稿作成支援システムに番組映像と番組台本を入力すると、図1のインターフェースが立ち上がる。作業者は番組を視聴し、解説文を挿入すべき箇所を発見したら、解説挿入ボタンを押す。すると、解説文入力ボックスが開くので、映像や台本を参照して、そのボックスの中に解説文を書き込む。

本システムの特徴は「解説文候補の提示」機能である。解説文入力ボックスの右側をクリックすると、システムは解説文の候補を提示する。例えば、図1のように、「横で見る塚本瑞恵」という解説文候補を提示する。作業者は、候補の中から利用可能な文を選択し修正することで解説文を作成することができる。

「解説文候補の提示」機能を実現するために、次の2つの処理が必要となる。

- ・タイムコードの自動推定
- ・解説文候補の自動生成

以下では、それぞれの詳細について述べる。

2.1 タイムコードの自動推定

解説文の作成には、台本中のト書きが参考になることが多い。参考となるト書きは解説文挿入箇所の前後の時刻にあるものであるが、台本にはタイムコードが入っていないため、ト書きや台詞にタイムコードを付与する必要がある。本システムでは、音声認識処理と言語処理によって台本中のト書きや台詞のタイムコードを推定した。

タイムコードの自動推定のアルゴリズムを図2に示す。

Step1 の音声認識処理では番組音声を文字化するとともに、そのタイムコードを決定する。Step2 の言語処理では認識結果と台詞との間で類似した文の対応付けを取る。対応付けが取られた場合、その認識結果のタイムコードを台詞に割り当てることで、タイムコードの推定を行う。



図1 解説放送用原稿作成支援システムのインターフェース

Step1: 音声認識による番組音声の文字化

| | | |
|----------|----------|---------------------|
| 01:02.95 | 01:04.49 | 待ってよ。 |
| 01:07.24 | 01:09.83 | ホールインワンだと思ったんだけどなあ。 |
| 01:13.94 | 01:16.04 | 一度出してみたいんだけどなあ。 |
| 01:18.29 | 01:20.73 | ホールインワン。 |
| 01:21.10 | 01:22.84 | もっと強くなれば出せるよ。 |
| 01:23.04 | 01:25.95 | もっと強いかあ。 |
| [タイムコード] | | [認識結果] |

文対応付け

Step2: 言語処理による認識結果と台詞の文対応付け

| | |
|-----|--------------------------|
| ト書き | 日曜日の屋。 |
| ト書き | 公園でゴルフのボールが転がっています。 |
| ト書き | ボールはフラッグの手前でそれました。 |
| ト書き | ボールを打っていたのは石川楓です。 |
| ト書き | 塚本瑞恵が横で見えています。 |
| 楓 | ホール・イン・ワンだと思ったんだけどなあ。 |
| 楓 | 一度出してみたいんだけどなあホール・イン・ワン。 |
| 瑞恵 | もっと強くなれば出せるよ。 |
| 楓 | もっと強いか。 |

[番組台本]

[タイムコード付き台本]

| | | | |
|----------|----------|-----|--------------------------|
| 00:53.18 | 00:54.26 | ト書き | 日曜日の屋。 |
| 00:54.41 | 00:58.01 | ト書き | 公園でゴルフのボールが転がっています。 |
| 00:58.16 | 01:01.39 | ト書き | ボールはフラッグの手前でそれました。 |
| 01:01.54 | 01:04.60 | ト書き | ボールを打っていたのは石川楓です。 |
| 01:04.75 | 01:07.08 | ト書き | 塚本瑞恵が横で見えています。 |
| 01:07.24 | 01:09.83 | 楓 | ホール・イン・ワンだと思ったんだけどなあ。 |
| 01:13.94 | 01:20.73 | 楓 | 一度出してみたいんだけどなあホール・イン・ワン。 |
| 01:21.10 | 01:22.84 | 瑞恵 | もっと強くなれば出せるよ。 |
| 01:23.04 | 01:25.95 | 楓 | もっと強いか。 |

Step3: ト書きや文対応付けできなかった台詞のタイムコードは式(1)、(2)により推定

図2 タイムコードの自動推定

しかしながら、ト書きや一部の台詞は Step2 ではタイムコードを推定することができない。そこで、Step3 で、式(1)、(2)により推定する。

$$T_{Start}[i] = T_{Start}[i+1] - t(u[i+1]) \quad (1)$$

$$T_{End}[i] = T_{Start}[i+1] \quad (2)$$

ここで、

$$t(u) = \begin{cases} t_0(u) & T_0 \geq T \text{ のとき} \\ t_0(u) \times T_0/T & T_0 < T \text{ のとき} \end{cases}$$

$u[i]$: i 番目の文。ただし、 $i=0, n+1$ の文はタイムコードが推定済み、 $i=1, 2, \dots, n$ はタイムコードが未定。

$T_{Start}[i]$: 文 $u[i]$ の開始時刻

$T_{End}[i]$: 文 $u[i]$ の終了時刻

$$T_0 = T_{Start}[i+1] - T_{End}[0]$$

$$T = \sum_{i=1}^n t_0(u[i])$$

$$t(u) = \text{SentLength}(u) \times 0.18$$

: 文長と時間長の関係。実験的に求めた。

$\text{SentLength}(u)$: 文 u の語数

実際の台詞(音声)は台本の台詞と一致する部分も多い。そこで、音声認識の際には、言語モデルを台本に適応化することにより認識率の向上が期待できる。また、文類似度により対応付けを行っているので、音声認識率が100%でなくとも対応付けは可能であり、タイムコードを推定することができる。

2.2 解説文候補の自動生成

解説文の特徴を分析するために、解説音声が入った2番組を観察した。その結果、解説文は台本のト書きを利用して作成される場合が多いことがわかった。さらに解説文には次の2つの特徴があることがわかった。

- ・ 時間や場所を明示する
- ・ 「連体修飾節+名詞」などの独特の表現を使う。

このような特徴を利用して解説文候補を自動生成した。図3にアルゴリズムを示す。

本アルゴリズムを使って解説文候補を自動生成した。対象とした番組はNHK教育テレビ「時々迷々(2010年1月6日放送)」[2]である。

本システムによる自動生成の結果([自動])と実際の解説文([実際])との比較結果の例を次に示す。

例1: 完全に一致

[自動] 「公園」

[実際] 「公園」

例2: 情報が不足(ただし、映像を見ないとその情報は得られない)

[自動] 「横で見る塚本瑞恵」

[実際] 「ベンチで見ているイヌを連れた瑞恵」

例3: 構造変換しない

[自動] 「楓がまた打つと、大きくフラッグを外れるボール」

[実際] 「ボールがコースを外れる」

4. おわりに

解説放送用原稿作成支援システムについて述べた。今後は、作業者の評価や意見を聞き、高性能・高機能化を図りたい。また、解説文候補の自動生成に関しては、解説放送コーパスを作成し分析することで精度を向上させていく予定である。

[参考文献]

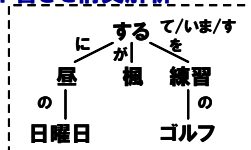
[1]伊藤藤之: 放送のアクセシビリティ向上の取り組み、IBMアクセシビリティフォーラム2009.

[2] <http://www.nhk.or.jp/mayomayo/ja/frame.html>

Step1: 解説文挿入時刻の前後5文のト書きを利用

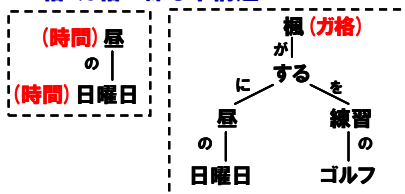
ト書き「日曜日の屋に楓がゴルフの練習をしています」

Step2: ト書きを構文解析



Step3: 構文木中のノードを辿って

- シソーラスで時間・場所表現を特定
- ハ格・ガ格に係る木構造に



Step4: 下位ノードを出力

「日曜日の屋」、「日曜日」、「屋」、「日曜日の屋にゴルフの練習をする楓」

図3 解説文候補の自動生成