

ト書きと動き情報を用いたドラマ映像の解析

6S-7

柳沼良知 和泉直樹 坂内正夫
東京大学生産技術研究所

1 はじめに

ワークステーション上で、映像、音声といったマルチメディア情報が比較的容易に扱えるようになるにつれて、動画像だけではなく、音声等も含めたマルチメディア情報をいかに加工し、編集し、利用するかが重要な問題となってきた。このような動画像、マルチメディア情報の認識、利用の試みとして、様々な研究が行なわれている⁽¹⁾⁽²⁾⁽³⁾⁽⁴⁾。しかしながら、より深い認識、利用のためには、動画像という1つのメディアを用いただけでは限界が生じる。一方、ドラマ映像は、映像、音声、シナリオ文書から構成されており、映像だけではなく、これら関連情報の認識を協調させることにより、より高度な認識、利用を行なえる可能性がある。このため、本稿では、ドラマ映像に同期されたシナリオ文書を用いたドラマ映像の解析について述べる。

2 シナリオ文書を用いたドラマ映像の解析

シナリオ文書には、各シーンの場所、登場する人物の名前、それぞれの人物の台詞等の情報が記述されている。また、「雄大な野山を走る列車」、「〇〇入ってくる」といった、ドラマの状況や人物の説明を行なうト書きが存在する。このため、ト書き情報を用いることで、ドラマ映像のより高度な解

析を行なえる可能性がある。本稿では、ト書き情報の中で、特に動きを表す言葉と、映像中の動き情報を用いたドラマ映像の解析について述べる。

ト書き情報を用いてドラマ映像の解析を行なう場合、問題となるのは、シナリオ文書中のト書きと映像があらかじめ対応付けられているとは限らない点である。このような、シナリオ文書と映像、音声の対応付け手法としては、先に、「台詞」や「シーンチェンジ」といったメディアに依存しないパターンを各メディアから抽出し、それらのパターンをDPマッチングを用いて対応付ける手法の提案を行なっている⁽⁵⁾。ここでは、これらの手法を用いて、シナリオ文書中のト書きと映像が対応付けられているものとする。

ト書きと対応する動きの学習は、以下の手順により行なう。

- 1) シナリオ文書中のト書きから動きを表す言葉を抜き出す。
- 2) 対応する映像から動きの抽出を行なう。映像中からの動きの抽出例を図1、2に示す。図1、2は、ともに、「はいつてくる」に対応する映像であるが、図1では、画像中央のドアが開き、人物が入ってくるため、動き部分が画像の中央部に集まっている。また、図2では、画像の左はじから人物が入ってくるため、動き部分が左はじから、中央に向かって移動している。このように、一般には、1つの動きを表す言葉について画像中の動きパターンは複数存在する。
- 3) 上述の結果を用いて、ト書き中の動きを表す言葉について、画像中の動きパターンの学習を行なう。

TV Drama Image Analysis based on the Scenario Document
Yoshitomo Yaginuma, Naoki Izumi and Masao Sakauchi
Institute of Industrial Science, University of Tokyo

以上の結果得られた学習パターンをもとに、以下の手順により、画像中から特定対象物(動きの主体)の抽出を行なうことができる。

- 1) シナリオ文書中のト書きから動きを表す言葉を抜き出す。
- 2) 対応する映像から動きの抽出を行なう。
- 3) ト書きと対応する動きの学習結果を用いて、ト書き中の動きを表す言葉に対応する映像中の動き部分を抽出し、その部分をト書きに記述されている動きの主体として抽出する。

3 おわりに

本稿では、ドラマ映像に同期されたシナリオ文書を用いたドラマ映像の解析について述べた。今後は、より多数のドラマ映像について本手法を適用し、その性能評価、および、有効性について更に、検討を行なっていく予定である。

文献

- [1] M.Sakauchi and J.Yamane: Realization of fully automated keyword extraction in image database system, SPIE International Symposium on Optical Applied Science and Engineering (1992).
- [2] Y.Gong and M.Sakauchi: A Method for Color Moving Image Classification Using the Color and Motion Features of Moving Images, ICARCV'92 (1992).
- [3] 長坂晃朗、田中譲：カラービデオ映像における自動索引付け法と物体探索法、情報処理学会誌、Vol.33, No.4, pp.543-550 (1992).
- [4] 柴田正啓：映像の内容記述モデルとその映像構造化への応用、電子情報通信学会論文誌、Vol.J77-D-II, No.5, pp.754-764 (1995).
- [5] 柳沼良知、影山誠、坂内正夫：DP マッチングを用いた映像と文書情報の対応付けとその応用、信学技報、IE94-11, pp.9-13 (1994).

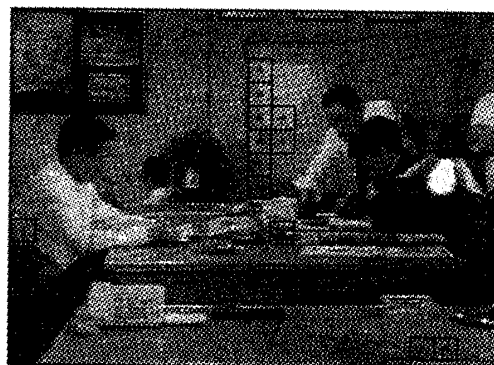


図1 「くる」に対応する動きパターン1



図2 「くる」に対応する動きパターン2