

## シナリオを用いたドラマのシーン検索システム

三浦健仁 中川裕志  
横浜国立大学工学部

映像情報を効率的に扱うために映像情報のデータベース化が重要となっている。しかし、インデックス付けを人手で行なうには膨大な労力が必要となるため、機械による自動インデクシング技術が必要である。本研究ではドラマ映像に対して自然言語処理技術と機械学習を利用し、シナリオを解析してト書きとセリフからシーンの具体的記述と雰囲気インデックス情報として抽出し、シーン記述データベースを構築するシナリオ解析システムを試作した。そしてシーン記述データベースを用いて所望のシーンを検索、提示するシーン検索システムを試作した。

### Scene Retrieval System based Drama Scenes

Takehito Miura Hiroshi Nakagawa

Faculty of Engineering,  
Yokohama National University

{miura,nakagawa}@naklab.dnj.ynu.ac.jp

To build a database of video information is very important to retrieve it efficiently, and auto indexing technique is in great demand because human indexing requires a great deal of labor. This paper describes an information extraction system which extracts concrete descriptions and moods of scenes as index information with natural language process technique and machine learning from a drama scene, we build the database which describes the structure of scenes, and develop a scene retrieval system.

#### 1 はじめに

ビデオ機器やコンピュータの高性能化、さらに記憶媒体の大容量化などが進み、テキスト情報だけでなく、静止画像や音声情報、さらに動画像情報を自由に扱うことが可能となってきた。また、多チャンネル放送や日々リリースされるパッケージメディアなどで映像情報が氾濫している。

しかし、映像情報にはほとんど内容を説明するような情報は付加されていないため、ユーザの求める番組、場面を瞬時に検索することができず、映像情報を効率よく取り扱う機能が必要不可欠となっている。

このためには映像情報をデータベース化し管理することが重要となる。映像情報にキーワードによるインデックスを付与し、検索を容易にすることはよく知られた手法 [2] である。

パッケージメディアにはインデックスが予め付与されている物も多く存在する。だがこれらのインデックス付け [1] は人手によって行われ、膨大な労力と時間を費やさねばならない。また、この作業は複雑で熟練を必要とし、一般視聴者が手軽にできるという物ではない。

よって映像情報を自動解析することに

よりキーワードを抽出してインデクス付けを行い、データベース化する技術が必要である。しかし、映像情報に画像処理を施すことによってインデクスを抽出する方法では自然画像を対象とした場合の認識率は低く、まだ実用化は困難である。

これに対してテキストメディアは現在の自然言語処理技術で計算機による自動解析が可能であり、さらに意味内容を直接反映しているという特徴がある。また、音声メディアは時間軸に対応しているという特徴を持つ。

そこで、テキストメディアや音声メディアから得られた情報を映像から得られる情報と協調させることにより、映像を意味内容に基づいて解析することが可能になると考えられる。

本研究では複数メディア協調の素材としてドラマ映像を用い、シナリオを自動解析し、ト書きやセリフに対応するシーンを記述する情報をインデクス情報として抽出した。そして、得られたインデクス情報を利用してシーン検索システムを試作した。

## 2 シナリオの特徴

シナリオは「シーン名」「ト書き」「セリフとその人物名」の三つの要素から構成されている。また、シナリオを元にドラマは制作されるためドラマの設計図であると考えられる。

### 2.1 シーン名

各シーンには、固有のシーン番号とシーンの場所を表すシーン名を持つ。またシーンの場所に時刻が付属する場合もある。

### 2.2 セリフ

セリフはシナリオを構成する要素の中でもト書きとともに大半を占める。

セリフの機能は

- 事実を知らせる

- 人物の心理、感情を知らせる
- ストーリーを進展させる

の三つである。しかし、全てのセリフがこの三つの機能を全て持っているわけではない。一つだけ、あるいは二つだけしか持たないものもある。

### 2.3 ト書き

ト書きは舞台説明や人物の動作を行なうという役割を担っている。

ト書きの機能は

- 情景の描写と説明
- 人物の行動、心理、性格の描写と説明

の二つに分けられる。

ト書きはできるだけ視覚的に書かれるという性質を持っており、視覚的な表現の可能な限度内で具体的に書かれている。また、ト書きの大きな特徴として、「現在形」で「簡潔に」書かれているということが挙げられ、このことから計算機による自然言語解析が可能なテキストであると言える。

## 3 シナリオの解析

シナリオの解析には「シナリオ解析システム」を用いる。これはシナリオの三つの要素を「内容に立ち入って」自然言語処理の手法で解析し、人物などの行動を構造的に記述したデータベースを作成するものである。

まずシーンを構成する要素として、「場所」「時刻」「人物」「話題」「ムード」の5つを定義する。

また人物、話題、ムードは更にいくつかの属性を持っている。これらの要素を元にシーン記述用のデータ構造を属性と属性値の対の組合せによって定義し、そのデータ構造を図1に示す。

シナリオ解析システムの大まかな流れを図2に示す。

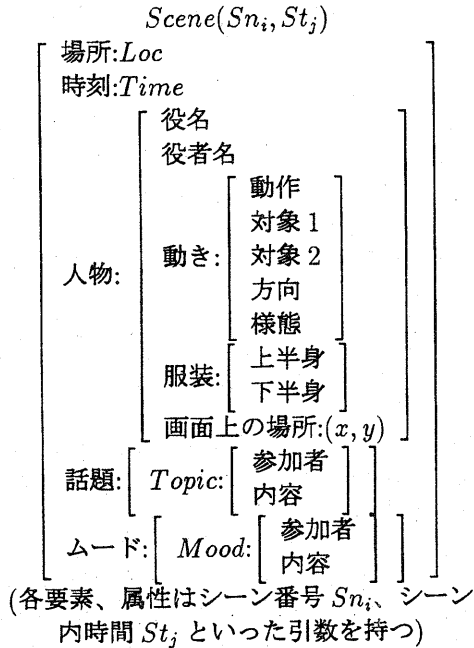


図 1: シーン記述用データ構造

### 3.1 シーン名の解析

シナリオではまずシーンの最初に「シーン番号」、「シーン名」が明記される。シーン番号はそのシーン固有の番号であり、ドラマ内で経過する時間順に付けられる。シーン名は一般的にシーンの場所の名前が付けられ、稀にシーンの時刻が補われることがある。

シーン名はシーンを特定する重要な要素であり、検索する際の有力なキーワードとなりうる。

シーンの最初にある「シーン番号」、「シーン名」をパターンマッチングにより抽出する。また、シーン名は「場所」として扱い、最後に括弧書きがあった場合は括弧内を「時刻」として抽出する。

また、「場所」、「時刻」は同一シーンでは不変であるとする。

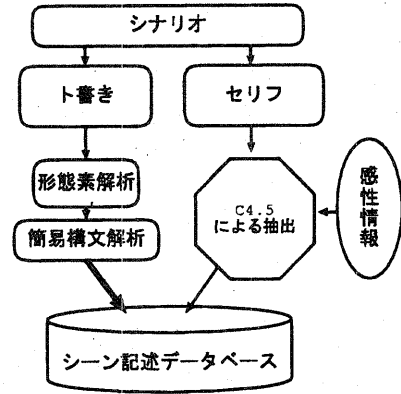


図 2: シナリオ解析システムの流れ

### 3.2 セリフの扱い

セリフはシナリオの中の重要な要素であり、その内容は話しの展開にも大きな影響を与えるので、その内容を理解することが望ましい。しかし、本来は「話し言葉」でありその解析は困難であるため、セリフを「セリフの話者」が「話す」という行動を取っているとした「ト書き」として扱うこととする。

また、通常の方法による自然言語処理とは異なる方法でセリフの内容から雰囲気(ムード)の抽出を試みる研究を行なった。これについては後述する。

## 4 ト書きの解析

次にト書きとト書き化したセリフの解析を行なう。具体的にはト書きを自然言語処理の手法で解析し、人物の動作、様態を抽出、理解する。本システムでは形態素解析と構文解析を行なった。

### 4.1 形態素解析

本システムでは形態素解析に形態素解析システム JUMAN[3] 及び茶釜 [4] を利用した。

まずト書きに形態素解析を行ない、形

態素列として出力する。しかし、形態素解析では主語、述語などの係り受け関係が反映されないという問題があるため、構文解析などにより高度な解析を行なう必要がある。本システムではこの問題を解決するために簡易構文解析を行なった。

#### 4.2 簡易構文解析

[12][13]では主語と述語のみを抽出しているが、ここでは目的語、様態の抽出、省略された主語の補完、関係節・従属節の扱いなどを拡張した。

本システムでは従来のような構文木を出力する構文解析は行わず、先に述べたシーン記述データを埋められるように

- 主語 (人物)
- 述語 (動作・様態)
- 目的語 (対象・方向)
- 形容詞 (様態)
- 形容動詞 (様態)
- 副詞 (様態)

を認識する簡易構文解析を行なった。

ト書き文の構文形態は単文のものでは

1. 主語 + 述語
2. 述語句 + 主語
3. 述語句
4. 主語 + 名詞句
5. 名詞句 + { 「である」 「だ」 }
6. 名詞句

の六種類に分類できる。以下ではこの分類にしたがって主語、述語を抽出する。

##### 4.2.1 主語の抽出

1,4の形態の文は主語が明示されている。よって、

- 名詞 + { 「、」 「が」 「は」 「も」 }

のとき、その名詞を主語とする。

2の形態の文は一種の関係詞節であり、主語は「名詞 + 「。」」の形で明示されている。よって、

- 動詞 (基本形、タ形) + 名詞 + 「。」

のとき、その名詞を主語とする。またここでいう動詞には助動詞や動詞性述語接尾辞を含むものとする。

3の形態の文は1の形の文から主語が省略されたものであり、主語は明示的には示されていないが、連続する二つの文では主語、主題が一致しやすいという性質があり、シナリオのト書き文ではその傾向が顕著である。よって主語が省略されている場合、

- 直前のト書きの主語

を主語として抽出する。

5,6の形態の文では主語の抽出を行なうことはできない。

##### 4.2.2 述語の抽出

1,2,3の形態の文は動詞が存在するのでそれをそのまま述語として抽出する。

また「タ系連用テ形」の動詞は後に続く動詞と並行して行なわれる動作と判断した。

4,5,6の形態の文では述語の抽出は行なわない。

また文の中に関係詞節を含む文も存在する。

- 動詞 (基本形、タ形) + 固有名詞 + 「が」

の時はその動詞を述語として抽出する。ここで固有名詞の後の「が」が他の格助詞の場合は別の文が挿入されているとする。

### 4.2.3 従属節を含む場合

ト書きには単文だけでなく従属節を含む複文も存在する。その場合、省略主語の推測をセンター理論 [6][7][8][9] を用いて検討した。また、センター理論の省略補完をスムーズにおこなうために複文の単文化をおこなっている。接続助詞はその前後の単文での主語の一致・不一致により三つに分類でき [10][11]、その分類によって推測の制約をおこなう。

### 4.2.4 目的語の抽出

ここでいう目的語は「ヲ格」「ニ格」の他に「へ格」を考える。

- 名詞 + { 「を」「に」「へ」 }

のとき、その名詞を

「を」の場合 対象1

「に」の場合 対象2

「へ」の場合 方向

とする。

また、これらの目的語がどの述語に係るかは、これらの目的語の直後の述語をその述語とした。

### 4.2.5 様態の抽出

ここでいう様態とは

- 動作の様
- 人物の心理状態

を示すものであり、「形容詞」「形容動詞」の連用形、「～ように」、「～ふうに」、「～なる」といった表現を抽出した。

更に「驚く」「笑う」「焦る」「ためらう」「あきれる」「慌てる」「急ぐ」といった心理・生理を表す動詞についても様態を表すとみなし抽出した。

表 1: ト書きの解析評価

	適合率 (%)			
	ドラマ1	ドラマ2	ドラマ3	ドラマ4
人物	96.7	98.7	99.1	95.1
動作	98.8	99.3	100	99.2
対象1	92.3	93.2	100	96.7
対象2	83.7	96.4	100	100
方向	100	100	100	77.0
様態	87.5	83.3	92.9	
文数	579	912	980	1025

### 4.3 ト書きの解析評価

本シナリオ解析システムを用いて四つのシナリオに対して簡易構文解析を行なった。評価は全ての文において「人物」「動作」「対象1」「対象2」「方向」「様態」が正しく抽出できているかどうかによって表した。その結果を表1に示す。

## 5 セリフの解析

本来「話し言葉」であるセリフの解析は困難である。そこで、感性的と言語的特徴との対応を機械学習により行ない、場面の雰囲気という面からセリフの解析を行った。

機械学習にはC4.5[5]システム用い、ドラマの各シーンのセリフを順に表示し、そのセリフ一つ一つについて表2の30対の感性語から当てはまるものをユーザーに選んでもらうことにより対応づけする。

### 5.1 決定木の学習

各セリフに対して特徴付けが行われたデータとセリフの言語的特徴データを用いてC4.5を使って決定木を作成することにより、逆に今度は与えられた感性語に対応するセリフの検索が可能になる。セリフの言語的特徴として

- 文末付近に現れるもの  
「終助詞」「助動詞」「受身表現」

表 2: 雰囲気を表す感性語

悲しい⇔楽しい
静か⇔賑やか
安心⇔不安
弱々しい⇔力強い
荒削り⇔洗練
暖かい⇔冷たい
動的⇔静的
退屈な⇔面白い
上品⇔下品
暗い⇔明るい
とげとげしい⇔なごやかな
ドライな⇔ウェットな
深みのある⇔うすっぺらな
緊迫した⇔どかな
きまじめな⇔ユーモラスな
重々しい⇔軽やかな
晴れやかな⇔うれいをおびた
機嫌のよい⇔不機嫌な
親しみやすい⇔親しみにくい
消極的な⇔積極的な
一生懸命な⇔適当な
親切な⇔不親切な
ふまじめな⇔まじめな
かわいらしい⇔にこらしい
やわらかい⇔固い
日常的な⇔非日常的な
対立している⇔仲がいい
シリアス⇔シリアスでない
穏やかな⇔陰悪な
形式ばった⇔くだけた

「記号」「敬語」「否定表現」「勧誘・願望」「動詞の活用(基本形, 言い切り, 連用形, 命令形)」

- 文中どの位置にも現れるもの  
「特定の名詞」「特定の動詞」「特定の形容詞」「呼び掛け」

を用いた。

## 5.2 セリフの解析評価

データそのもののクラス「yes」「no」が分かっているセリフでないとは評価が行えないため、全セリフのうちアンケート

表 3: セリフの解析評価

感性語	適合率(%)	再現率(%)	F
うすっぺらな	56.1	95.8	0.71
くだけた	56.4	93.2	0.70
退屈な	52.9	96.4	0.68
消極的な	56.8	85.7	0.68
穏やか	52.8	95.7	0.68
動的	54.3	90.7	0.68
親切な	52.8	94.0	0.68
かわいらしい	51.9	94.6	0.67
不親切な	50.0	100	0.67
日常的な	54.7	83.0	0.66
重々しい	51.0	91.5	0.65
不安	58.6	73.1	0.65
静か	52.3	85.5	0.65
明るい	61.1	10.7	0.18
シリアスでない	50.0	3.7	0.07

の行なわれたセリフのみに対して評価<sup>1</sup>を行なった。評価はアンケートの行なわれたセリフすべてを使用して行なった。全体を学習データとテストデータに二分し、学習データを用いて作成したルールをテストデータに適用して評価した。その後両者を入れ換えて再度評価を行なった。結果の一部をを表 3に示す。

## 6 シーン検索システム

シナリオ解析システムで構築されたデータベースを利用したドラマのシーン検索システムを試作した。

### 6.1 検索条件入力

検索条件をユーザが入力する際に、無用な負担を与えないようにすることが重要である。そこで検索システムは GUI を用いてユーザが検索条件を提示されたリストの中から選択することで指定することとした。

<sup>1</sup>P を適合率、R を再現率としたとき、両者のトータルでの精度を表す F を  $F = \frac{2PR}{P+R}$  と定義した



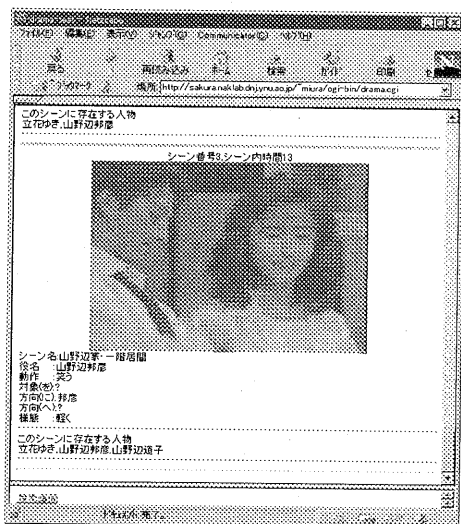


図 5: 検索結果画像表示画面 2

## 謝辞

この研究は文部省科学研究費補助金(課題番号 09NP1401)の援助を受けている。また、東京大学の坂内正夫教授には大変有益な御助言を頂いた。

## 参考文献

- [1] 柴田正啓, 金淵培: 自然言語を用いた映像ハンドリング, 信学技報, pp.23-30, 1996
- [2] R.Weiss,A.Duda,and D.K.Gifford: Composition and Search with a Video Algebra, IEEE MultiMedia Vol.2 No.1, pp.12-25, 1995
- [3] 松本 裕治, 黒橋 禎夫, 宇津呂 武仁, 妙木 裕, 長尾 真: 日本語形態素解析システム JUMAN 使用解説書 version1.0, 1993
- [4] 松本, 北内, 山下, 今, 今村: 日本語形態素解析システム茶筌 version1.0, 1997
- [5] J. Ross Quinlan, 古河 康一 監訳: AIによるデータ解析, 株式会社トッパン, 1995
- [6] Marilyn Walker, Masayo Iida, Sharon Cote: Centering in Japanese Discourse, COLING90, pp.1-6, 1990
- [7] 田村浩二, 奥村学: センター理論による日本語談話の省略解析, 自然言語処理 107-12, pp91-96, 1995
- [8] Barbara J.Grosz, Aravind K.Joshi, Scott Weinstein: Providing a Unified Account of Definite Noun Phrase in Discourse, 21th Annual Meeting of the Association for Computational Linguistics, pp.44-50, 1983
- [9] Megumi Kameyama: A Property-Sharing Constraint in Centering, 24th Annual Meeting of the Association for Computational Linguistics, pp.200-206, 1986
- [10] 吉本啓: 談話処理における日本語ゼロ代名詞の扱いについて, 自然言語処理 56-4, 1986
- [11] 南不二男: 現代日本語の構造, 大修館書店, 1988
- [12] 和泉直樹, 柳沼良知, 中川裕志, 坂内正夫: シナリオ文書の解析による存在・行動マップの作成, 電子情報通信学会論文誌, (D-II), Vol.J79-D-II, no.11, pp.1993-1996, 1996.
- [13] Hiroshi Nakagawa, Yoshitomo Yaginuma, Masao Sakauchi: Scene Direction Based Reference In Drama Scenes, Proceedings of Workshop Sponsored by ACL, pp.52-58, 1997.