

# 感情を付加した朗読システムの為の文章の場面分割と感情分類

浦野 慧介<sup>†</sup> 杉本 富利<sup>†</sup> 米山 正秀<sup>†</sup>

<sup>†</sup>東洋大学工学部情報工学科

## 1 はじめに

今日、音声合成技術の発達により比較的自然的な合成音を得られるようになったが、感情を表現することに関しては未解決な問題がある。本研究ではコンピュータに感情豊かな音声で文章を読ませることを最終目標としている。感情表現豊かな音声合成を実現する試みは過去にいくつか行われている。それらの手法の多くは、飯田ら [1] の研究に代表されるように感情ごとの音声のコーパスを用いることによって、感情豊かな音声合成を実現するというものである。しかし、これらの研究では、あらかじめ文章の感情がわかっていなければ音声合成が行えない。つまり、任意の文章を与えたときに、全自動で感情表現豊かな朗読システムを実現するには問題がある。そこで、本研究では感情表現を自動的に行う為に小型辞書並みの語彙がある感情語辞書を作成し、その辞書に基づいて感情を分類することを試みる。図 1 に示すように本システムでは、まずコンピュータに入力された文章を場面で区切り、場面毎に感情を決定する。次に決定された感情を基に、韻律パラメータを決定して合成音に感情を付加する。

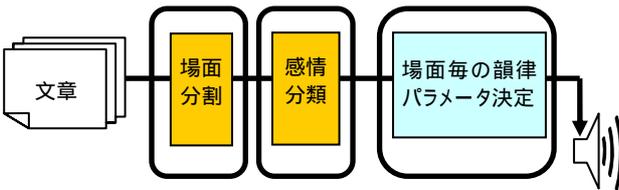


図 1 感情付きテキスト音声合成システムの概要

## 2 研究の概要

### 2.1 文章の場面分割

場面分割とは、文章を適切な句読点で分割して内容の等しい複数の場面に分けることである。本研究では、場面分割に Hearst らの提案した Text tiling algorithm を使用する [2]。Text tiling algorithm では、まず文章を最小単位の単語で区切

る。次に、ある単語数の窓と呼ばれるものを連続して二つ設けて左右の窓とする。左右の窓間の内容の類似度を式 (1) により求める。 $b_l$ 、 $b_r$  がそれぞれ左窓、右窓を表し、 $w$  が単語の反復の数を表す。

$$sim(b_l, b_r) = \frac{\sum_i w_{b_l} w_{b_r}}{\sqrt{\sum_i w_{b_l}^2 \sum_i w_{b_r}^2}} \quad (1)$$

設けられた左右の窓をある単語数の刻み幅で文章の最初から最後までずらしながら類似度の計算を繰り返す。類似度が特に低くなった場所を極小点とし、その極小点に一番近い句読点を場面の分割点とする。過去の研究では、同一単語の反復数から類似度を求めたが、本研究では、類義語の反復数によって、類似度を求める。類義語には大日本図書出版の分類語彙表の一番下の階層を用いる。この場合、ある単語を平均で約 10 個の類義語と同一視することができる。

### 2.2 場面の感情分類

分割された場面に現れる感情語を抽出し、その感情の強さを累積する。そして、喜びの強さの累積値がある値以上であれば快の場面、悲しみの値が大きい場面は不快の場面、両方の値とも小さい場合は無感情の場面とする。

### 2.3 感情語辞書作成

前節で使用した感情語辞書は、前述の分類語彙表から抽出した形容詞および副詞 988 語について被験者アンケートによって 6 感情 (喜び、悲しみ、怒り、驚き、嫌悪、恐怖) の強さを求めたものである。

## 3 小説の場面分割の最適パラメータ推定

被験者による場面分割結果と Hearst 法による場面分割結果を比較して Hearst 法の最適なパラメータを推定する。

### 3.1 最適パラメータ

Hearst 法は、本來說明文の場面分割を行うことが目的であり、物語のような文章を場面分割する場合、窓幅、刻み幅のパラメータの最適な値を推定する必要がある。本研究では、単語を類義語に拡張して類似度を求めているのでその点からも、この最適パラメータの推定が必要である。

Scene Division and Emotional Classification of Text for a Reading System with Emotion  
Keisuke Urano<sup>†</sup>, Futoshi Sugimoto<sup>†</sup>, Masahide Yoneyama<sup>†</sup>  
<sup>†</sup>Dept. of Information and Computer Sciences, Faculty of Engineering, Toyo University

表1 一致度の被験者平均

	F 値の被験者平均
窓幅 100 刻み幅 10	57%
窓幅 100 刻み幅 20	48%
窓幅 100 刻み幅 30	75%
窓幅 120 刻み幅 10	55%
窓幅 120 刻み幅 20	57%
窓幅 120 刻み幅 30	39%

窓幅、刻み幅の単位はそれぞれ単語数である。

### 3.2 Hearst 法と被験者実験の比較

被験者に壺井栄の「あしたの風」を読んでもらい、句読点により場面分割をしてもらう。被験者は学生5名である。被験者による分割点と Hearst 法による分割点の一致度が高くなる窓幅と刻み幅のパラメータを最適パラメータとして推定する。一致の誤差は前後2文とし、その範囲内であれば分割点が一致していると考え。一致度は、式(2)(3)に示す再現率と精度を用いて計算される式(4)のF値によって判定する。

$$\text{再現率} = \frac{\text{被験者とHearst法の場面境界が一致した数}}{\text{被験者の場面境界の数}} \quad (2)$$

$$\text{精度} = \frac{\text{被験者とHearst法の場面境界が一致した数}}{\text{Hearst法の場面境界の数}} \quad (3)$$

$$\text{F値} = \frac{2 \times \text{再現率} \times \text{精度}}{\text{再現率} + \text{精度}} \quad (4)$$

### 3.3 結果と考察

Hearst らの研究結果では、F 値 58%以上で十分な一致度が得られていると判定されているので、この実験で得た窓幅 100、刻み幅 30 の場合の F 値 75%は十分な一致度が得られており、窓幅 100、刻み幅 30 が最適パラメータと推定できる。

### 4 被験者による場面の感情分類

3.2 節で行った被験者による場面分割の後に、それぞれの場面の感情を快、不快、無感情に分類してもらった。また、2.2 節で述べた方法によって、各被験者が分割した場面の感情を感情語辞書に基づいて感情分類した。この辞書による感情分類と被験者の行った感情分類の一致度を表2と表3にまとめた。一致度は式(5)によって求めた。被験者 C 以外は一致度がほぼ 70%以上あるので、感情語辞書による場面の感情分類が妥当であるといえる。

$$\text{一致度} = \frac{\text{被験者の快(不快) + 無感情}}{\text{辞書による快(不快)}} \quad (5)$$

表2 快場面における被験者の感情分類

	辞書による分類	被験者の分類			一致度
		快	無感情	不快	
被験者 A	5	3	1	1	80%
被験者 B	9	5	3	1	89%
被験者 C	9	4	0	5	44%
被験者 D	7	4	1	2	71%
被験者 E	8	3	2	3	63%

表3 不快場面における被験者の感情分類

	辞書による分類	被験者の分類			一致度
		快	無感情	不快	
被験者 A	2	0	0	2	100%
被験者 B	4	1	0	3	75%
被験者 C	4	2	0	2	50%
被験者 D	3	0	0	3	100%
被験者 E	4	1	1	2	75%

### 5 まとめと今後の課題

Hearst 法による場面分割は、パラメータを窓幅 100、刻み幅 30 で非常に満足のいく結果が得られた。しかし、今回は一つの物語のみでしか検証しておらず、更に試行錯誤的にパラメータを組み合わせおこなった。今後は遺伝的アルゴリズム等を用いて最適なパラメータを求める方法を考慮すべきである。文章の感情分類に関しては、ある程度満足のいく結果が出たように思われるが、今後は感情語辞書の単語を形容詞と副詞のみではなく、名詞や動詞まで拡張することによって更なる感情分類の精度の向上が期待できる。

#### 参考文献

- [1] 飯田朱美, ニック・キャンベル, 安村通晃(1999): 「感情表現が可能な合成音声の作成と評価」, 情報処理学会論文誌, Vol.40, No.2, pp. 479-486
- [2] M.Hearst; Multi-Paragraph Segmentation of Expository Text. In Proceedings of the 32<sup>nd</sup> Annual Meeting of Association for Computational Linguistics, pp.9-pp.16. 1994