

FLVQ を用いた複数短文音声の感情認識に関する研究

飯高 智美 (日大工・情報) 金子 正人 (日大工・情報)  
 武内 惇 (日大工・情報) 藤本 洋 (日大工・情報)

1. はじめに

本研究では、最も身近なコミュニケーション手段である“言葉(音声)”から、個性や感性を抽出することを目的とする。多くの言葉には感情が含まれており、その感情はあいまいでコンピュータには扱いにくい。そこで、あいまいな情報を扱うことができるファジィ学習ベクトル量子化法(以下、FLVQと略)<sup>[1]</sup>を用いて、音声からの感情認識を試みた。

これまで、1つの短文のみという条件での感情認識や、学習済みである複数句における感情認識の実験を行い、約8割の認識結果を得ることができた<sup>[2]</sup>。

そこで本稿では、複数の短文を学習させ、未学習短文による感情認識を試みた実験結果について報告する。

2. 音声の感情認識実験

2.1 目的

実験の目的は、未学習である短文からの感情認識であり、学習に用いる短文の数の変化による認識率の変化を見る。学習短文の数の増加に伴い、認識率が向上することがわかれば、複数の短文を学習することは有効であるといえる。実験の手順について、次に記載する。

2.2 発話実験

実験で使用する音声サンプルを採取する。今回は、表1に示す5種類の短文(以下、短文A~Eとする)に、5つの感情「平静」「喜び」「悲しみ」「怒り」「嫌々」を付加したものをそれぞれ発話し、感情情報を含む音声サンプルを採取する。

今回は、成人男性6名、成人女性2名の計8名の音声サンプル(5短文×8人×5感情=200サンプル)を採取した。

表1. 実験に使用する5つの短文

A	「ああ、雪だ。」
B	「はい、わかりました。」
C	「ええ、聞いています。」
D	「もうたくさん。」
E	「本当だよ。」

2.3 聴取実験

発話実験で得た音声サンプルの妥当性を示すために聴取実験を行う。実験の内容は、被験者に「平静」の音声を聞かせ、その後各感情音声をランダムに聞いてもらう。感情ごとに感じた印象を0から1の間に自由に設定できるファジィ形式のアンケートに記入させる(図1)。被験者の数は30名とし、それにより、各音声それぞれの妥当性が示される。

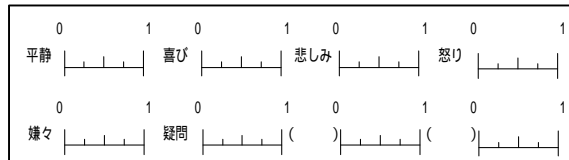


図1. 聴取実験における感情評価記入欄

2.4 音響パラメータの解析

音声サンプルの波形をそのまま使用すると膨大な情報量になるため、音声の特徴的な部分だけを抽出し感情認識実験で用いる。抽出する音響パラメータとして、ピッチ周波数や発声時間の解析を行う(図2)。

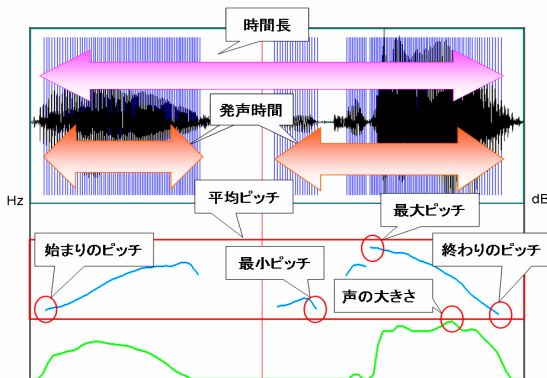


図2. 音響パラメータ解析

2.5 感情認識実験

音声サンプルの音響パラメータの解析結果から、FLVQプログラムを用いた感情認識実験を行う。以下に実験条件を示す。

学習及び認識の対象となる感情は、「平静」「喜び」「悲しみ」「怒り」「嫌々」の5種類であり、音声に含まれている感情を学習、認識させる。

感情認識実験では、入力ベクトルをピッチ周波数の“最大ピッチ”、“最小ピッチ”、“平均ピッチ”、“始まりのピッチ”、“終わりのピッチ”さらに、“声の大きさ”、“時間長”、“発声時間”の8次元に設定する(図2)。

学習回数を500回に設定する。

短文A~Eまでの5つの短文を用いる。学習させるデータを、1つから4つの短文に変化させた場合について、学習済み短文、未学習短文それぞれにおける感情認識の実験を行う。

音声サンプルとして使用する学習データは、学習の効果を高めるために、聴取実験を行ったアンケート結果において、その短文内の上位4名の音声を使用する。

以上の実験条件により、学習に用いる短文の数による認識率の変化を見る。

### 3. 感情認識実験の結果と考察

複数の短文を学習させ、未学習短文からの感情認識を試みた各感情ごとの結果を図3~7までに、図8に5感情平均の値を示す。これより、全体的に学習短文を増やすことで認識率が向上している。また、学習済み短文と未学習短文の認識率の差が縮まっていることがわかる。図3~7より、未学習短文については、学習短文を増やしただけ認識率が向上している。ただ、学習短文では、感情ごとに認識率にばらつきがみられた。

悲しみ、怒り、嫌々（図5~7）では、学習短文の数による認識率の変化はみられない。また、平静の学習短文数による認識率の変化は、特に大きかった。これは、平静にも何らかの感情が含まれているためだと考えられるが、なお検討が必要である。

5感情平均（図8）から、全体的に認識率が向上していることがわかるが、学習短文が3つの場合と4つの場合とでは、それほど認識率の変化がみられない。これ以上学習する短文を増やしても、認識率が約8割で収束されるのではないかと推測される。

### 4. おわりに

本稿では、学習短文を増やすことによって、未学習である短文の認識率が向上することを述べた。つまり、複数の短文を学習させることは、未学習である短文からの感情認識に有効であると判断でき、あらゆる短文についての感情認識への適応が期待できる。

今後の課題としては、短文による、発声時間等の時間差の影響を調べるとともに長文の感情認識についても検討する。また、短文をさらに細かい単位である文節に分け、この文節単位で、学習・認識を行うことにより、認識率の向上を図る。

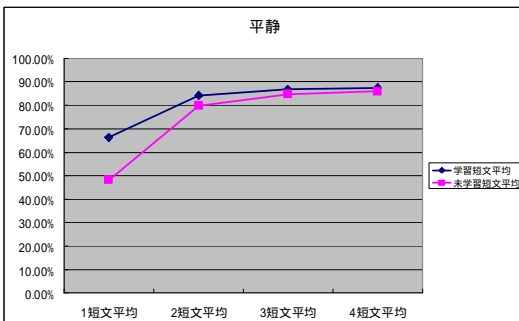


図3. 学習短文の数と認識率（平静）

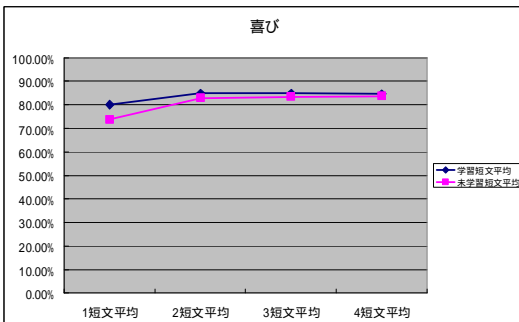


図4. 学習短文の数と認識率（喜び）

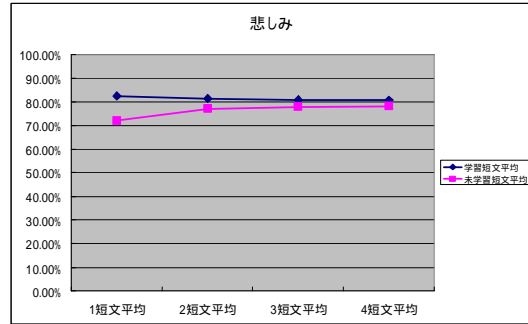


図5. 学習短文の数と認識率（悲しみ）

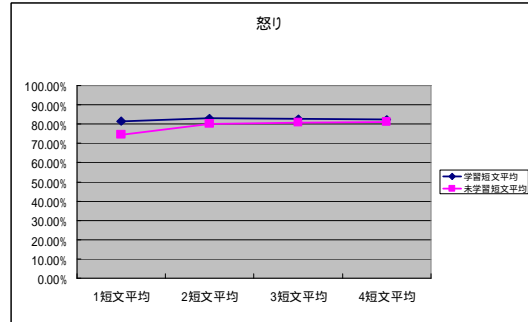


図6. 学習短文の数と認識率（怒り）

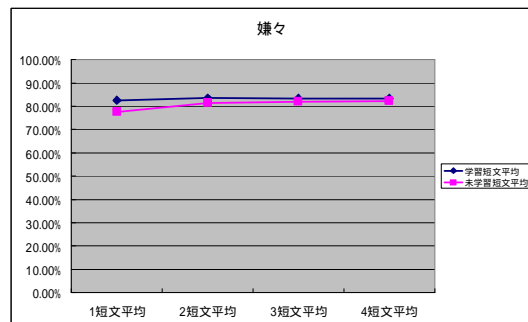


図7. 学習短文の数と認識率（嫌々）

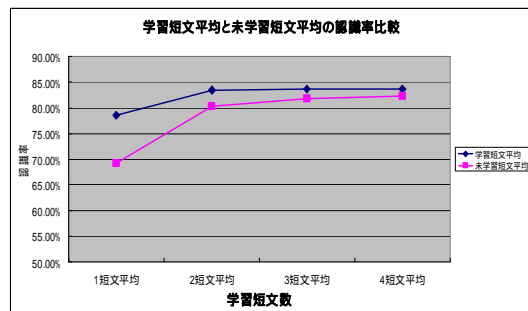


図8. 学習短文の数と認識率（5感情平均）

#### 【参考文献】

- 櫻庭，中本，森泉：“ファジィ理論を用いた学習ベクトル量子化法”，電子情報通信学会論文誌 Vol. J-73-D No11, pp. 1863-1871 (1990)
- 斎藤，金子，武内，藤本：“ファジィ学習ベクトル量子化法を用いた未学習音声データの感情認識”，電気関係学会東北支部連合大会論文集 2H-19, pp. 303 (2005)