

音声情報を用いた医療コミュニケーションにおける不適切発話の検出手法

栗栖 道久 目良 和也 黒澤 義明 竹澤 寿幸

広島市立大学情報科学部

1 はじめに

近年、日本でも患者に対する医師のコミュニケーションのあり方について関心が高まり、医学教育の場でも医療コミュニケーションについての実技指導や評価が導入されている。医師は医学教育や研修を積み専門家として一人前になる過程で専門的知識や技術を身につけ、その過程を通じて先輩医師や看護師などの「やり方」や現場の「習慣」に染まっていく。さらに現代医学という大きな枠組みによる独特の世界観を持つに至り、これらがときに医師側からみた場合と患者側からみた場合での隔たりとなる。医師と患者の間でのコミュニケーションギャップの要因の一つとして医師が言った言葉が患者に十分に理解されないということがある。

我々の研究グループではテキストから医師の不適切な発話の特徴を計11項目定め、SVMを用いて不適切発話の検出を行った[1]。しかし、自信の有無や心がこもった発言か否かなどテキスト情報だけでは判断することが難しい特徴もある。そこで本研究では、テキスト情報からでは判断できないこれらの心的状態を音響的特徴により判別する手法を提案する。

2 SVM

SVM[2]とは、Support Vector Machineの略で統計的学習理論の枠組みで提案された学習機械のことである。SVMは線形入力素子を利用して2クラスのパターン識別器を構成する。学習には訓練サンプルから各データ点との距離(マージン)が最大になる分離超平面を求めるマージン最大化という基準で線形入力素子のパラメータを学習する。SVMの基本的な構造は図1に示すように線形しきい素子である。図では x_1 と x_2 の2つの素性からなる入力データを分類している。しかしこれでは線形分離不可能なデータに適用することができ

ない。そこで非線形な分類を可能にする方法として高次元化が挙げられる。これは非線形写像によって元の入力データを高次元特徴空間に写像し、特徴空間において線形分離を行うという方法である。そうすることによって結果的に元の入力空間においては非線形な分類を行っていることになる。

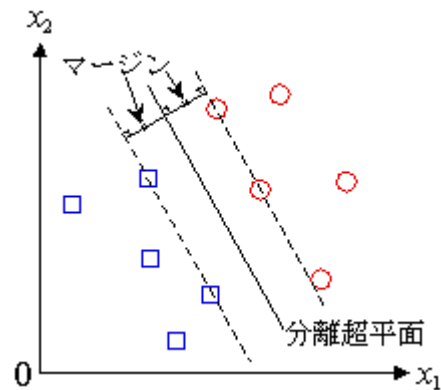


図1 SVMのイメージ

3 音響的特徴

発話者の心的状態や発話の意図を判別するため、本研究では音声データから以下に示す8項目の音響的特徴を用いる。

- ①ピッチ最大値： $f0_{max}$
- ②ピッチ平均値： $f0_{mean}$
- ③ダイナミックレンジ： $f0_{range}$
- ④パワー最大値： $power_{max}$
- ⑤パワー平均値： $power_{mean}$
- ⑥話速： $speaking-rate$
- ⑦ピッチの勾配比率： $f0_{slope}$
- ⑧パワーの勾配比率： $power_{slope}$

特徴値の算出方法を以下に示す。まず録音した音声からサンプル番号、サンプリング時間[sec]、音声パワー[dB]、ピッチ周波数[Hz]

を抽出し、ピッチ周波数の値を基に無音・無声区間を除外する。その後有音区間からピッチの最大値と最小値、平均値を決定する。ダイナミックレンジとは最大値と最小値の比率であり、ここでは最大値を最小値で割ることにより算出する。パワーの最大値と平均値もピッチと同様に求める。話速は、無音区間と有音区間の境界から発話の開始時間と終了時間を求め、その差をとることで発話時間を求め、発話時間を発話の文字数で割ることで話速を算出する。ピッチの勾配比率とパワーの勾配比率は抽出したピッチとパワーの値が直前と比べて変化している場合を勾配とし、下勾配のサンプル数で上勾配のサンプル数を割ることにより算出する。図2は勾配の例で、△は上勾配のサンプル、□は下勾配のサンプル、○は変化がない・頂点・始端終端を表している。この図では、上勾配は4、下勾配は8であることから勾配比率は0.5である。

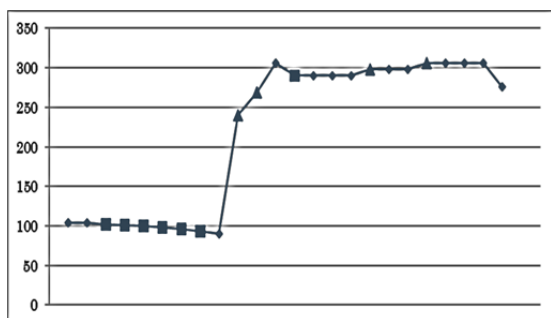


図2 勾配のサンプル数の比率

4 評価実験

医師から問診を受けた患者へのインタビューから、患者が好む医師の印象として、“温かく、共感的な人”、“自信を持っていそうな人”、“話しやすい人”、“同じことを繰り返さない人”、“患者の話に耳を傾け反応を持ってくれる人”が挙げられる[3]。これを参考にし、本実験では以下の3種類の発話を実験データとした。

- ・疑問と納得の意図の判別をするための発話、「そうですか」

p : 疑問の意図を込めた発話

n : 納得の意図を込めた発話

- ・気持ちがこもっているか否かを判別するための発話、「ありがとう」

p : 気持ちがこもっている発話、

n : 気持ちがこもっていない発話

- ・自信があるか否かを判別するための発話、「検査の結果、風邪です」

p : 自信がある発話

n : 自信がない発話

男子大学生10名に納得の意図を込めた「そうですか」と疑問の意図を込めた「そうですか」を各5回ずつ発話してもらい、「ありがとう」と「検査の結果、風邪です」についても、同様の手順で音声データを収集した。収集した音声データの判別を被験者3名(男性3名)に行ってもらい、3名全員が正しく判定できたもの、または2名が正しく判別し1名がどちらでもないとしたもののみを評価実験に用いた。評価用音声データ数の内訳を表1に示す。

表1 評価用音声データ数

	p	n	合計
そうですか	42	51	93
ありがとう	27	39	66
風邪です	47	34	81

素性値の計算の前に、素性値の標準化を行っている。例として $f0_{max}$ の標準化の手順を以下に示す。jは発話者を表す変数、iは一つの発話を指す変数、 $f0_{max_{ij}}$ は発話者jのi番目の発話の $f0_{max}$ を表し、iとjはそれぞれ1から10までの値をとる。まず発話全体の平均を求める(式(1))。次に1人毎の平均を求める(式(2))。そして全体平均から個人平均の差をとり、その値を個人毎の値に加えることで標準化した $f0_{max'_{ij}}$ を得る(式(3))。他の素性、発話も同様の手順で標準化を行う。

$$\overline{f0_{max}} = \frac{1}{i \cdot j} \sum_i \sum_j f0_{max_{ij}} \dots (1)$$

$$\overline{f0_{max}}_j = \frac{1}{i} \sum_i f0_{max_{ij}} \dots (2)$$

$$f0max'_{ij} = f0max_{ij} + (\overline{f0max} - \overline{f0max_j}) \cdot \dots (3)$$

素性 8 項目の全組み合わせである 255 組の多次元データを作成し SVM による判別を行った。学習及び評価は表 1 のデータに対して Leave-One-Out 方式で行った。表 2 に各発話の分類結果を示す。それぞれの発話で最も良い素性の組み合わせは“疑問と納得”で疑問が (f0slope, powerslope), 納得が (f0max, f0mean, powermax, powermean, speaking-rate, f0slope, powerslope), “気持ちがこもっているかない”では共に (f0max, powermax, powermean, speaking-rate, f0slope), “自信があるかない”では共に (powermax, speaking-rate, f0slope, powerslope) であった。図 3, 4 に疑問と納得の発話を示す。疑問の後半部でピッチ(点線)とパワー(実線)が上がるという特徴が、(f0slope, powerslope) という素性によって抽出できていることが確認できた。

表 2 各発話の実験結果

	適合率	再現率
疑問	0.795	0.833
納得	0.843	0.843
気持ちがこもっている	0.786	0.815
気持ちがこもっていない	0.868	0.800
自信がある	0.898	0.936
自信がない	0.906	0.852

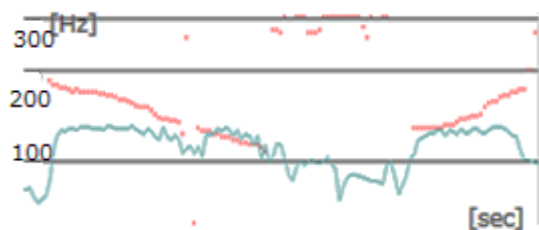


図 3 正しく分類された疑問の発話

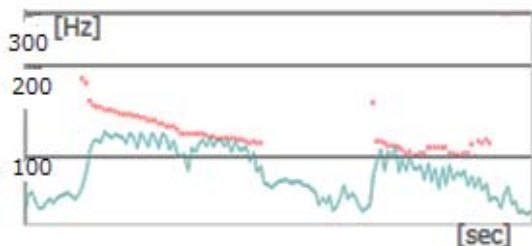


図 4 正しく分類された納得の発話

5 まとめ

本研究では, SVM を用いてテキスト情報だけでは判別できない発話の検出方法について提案した。実験の結果, 納得と疑問の意図の判別で F 値は納得: 0.814, 疑問: 0.843, 気持ちがこもっているか否かの判別で気持ちがこもっている: 0.800, 気持ちがこもっていない: 0.857, 自信の有無の判別で自信がある: 0.917, 自信がない: 0.879 の結果が得られた。

今後の課題として, 音声前半部と後半部の特徴が打ち消しあったため誤分類された等の発話を正しく分類するために“発話を分割し特徴値を算出する”, “ピッチやパワーが最大となった位置”等の新しい素性の追加や, 実際の医療現場での対話データや医学部生教育のためのロールプレイを録音したデータを用いて実験を行うことを予定している。

- [1] 和田龍之介, 目良和也, 黒澤義明, 竹澤寿幸, SVM を用いた言語情報に基づく医療コミュニケーションにおける不適切発話の検出手法, 2011 IEEE SMC Hiroshima Chapter 若手研究会講演論文集, pp.61-64, 2011
- [2] Vapnik, V.N.: Statistical Learning Theory, A Wiley-Interscience Publication, 1998
- [3] Margaret Lloyd, Robert Bor, 事例で学ぶ医療コミュニケーション・スキル—患者とのよりよい関係のために—, 山内豊明(監訳), 西村書店, 2007