社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2007－ＳＬＰ－69 (25)
2007／12／20

音声認識との統合によるシステム要求検出

佐古 淳† 山形 知行†† 滝口 哲也†† 有木 康雄††

† 神戸大学大学院自然科学研究科 〒657-8501 兵庫県神戸市灘区六甲台町1-1
†† 神戸大学大学院工学研究科 〒657-8501 兵庫県神戸市灘区六甲台町1-1
E-mail: †{sakoats,yamagata}@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

あらまし 音声をインターフェイスとして用いる際，システムに対してなされた発話か，周りの人間に対してのものかを判別する必要がある．この問題に対し，柔軟な発話を受理可能なものとして，音声認識結果をブースティングによってシステム要求か雑談かを判別する手法の提案を行ってきた．しかし，音声認識結果には認識誤りを含む場合があることから，認識誤りを原因として，システム要求と雑談の判別を誤る場合があった．本稿では，システム要求検出を音声認識の定式化に組み込むことにより，認識仮説まで用いたより高精度な要求検出について述べる．システム要求検出には従来と同様ブースティングを用いる．ただし，ブースティングの出力スコアは確率ではないため，sigmoid関数を用いて疑似確率化することで，音声認識との統合を行った．実験により，従来の認識結果から識別する手法よりも再現率が改善し，適合率0.98，再現率0.94，F値0.96を実現した．

キーワード システム要求判別，ブースティング，sigmoid，音声認識

System Request Discrimination Based on AdaBoost

Atsushi SAKO†, Tomoyuki YAMAGATA††, Tetsuya TAKIGUCHI††, and Yasuo ARIKI††

† Guraduate School of Science and Technology, Kobe University Rokkodaicho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan
†† Guraduate School of Engineering, Kobe University Rokkodaicho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan
E-mail: †{sakoats,yamagata}@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

Abstract It is necessary to discriminate system requests from human-human conversation speeches for speech user interfaces. We had proposed the boosting method that discriminates system requests from chats based on 1-best result of speech recognition system. This method can retrieve various expressions due to boosting algorithm. However it causes discrimination error when speech recognition results includes keyword mis-recognition. In this paper, we propose the system request detection method that can consider not only 1-best result but also speech recognition hypotheses. The proposed method is formulated incorporating system request detection into speech recognition. Boosting method is employed as system request discrimination model, however its output score is not probability. Thus boosting score is converted into pseudo probability based on sigmoid function in order to integrate system request discrimination and speech recognition. The experimental results showed that 0.98 of precision, 0.94 of recall and 0.96 of F-measure.

Key words System Request Detection, Boosting, sigmoid, Speech Recognition

1. はじめに

近年，音声による機器操作インターフェイスが実用化されつつある．特に，ロボットとのコミュニケーションや，カーナビ操作などの手を使うことが困難な機器への適用がなされている．音声インターフェイスを用いる際，

システムへの要求発話とそれ以外の発話を区別する必要がある．このため，物理的なスイッチを用いることで区別を行う手法や，発話者の視線方向を用いる手法[1]，発話の音響的な特徴を用いる手法[2],[3]，ネットワーク文法を用いて自動的に区別を行う手法[4]が提案されている．しかしながら，視線方向を用いた手法は，ロボットと発話者の距離が離れており視線方向検出が困難な場合や，カーナビ操作のように視線をシステムへ向けることが危険な場合には利用することができない．音響的な特徴を用いた手法では，そもそも音響的な特徴だけでは判別が困難な場合が存在することや，システムが高度化・高性能化するにつれて，さらにはユーザーがシステムに習熟するにつれて，システムに対する発話が人に対する発話に近づき，判別が困難になるといった問題がある．ネットワーク文法を用いた手法は，効果的にシステム要求を検出することが可能であるが，一方で，人手によって詳細なネットワーク文法を構築するには大きなコストがかかることや，システム要求発話の柔軟性・多様性が失われてしまい，決まった発話をユーザーが暗記しないとシステムを操作できない，などの問題があった．

この問題に対し，我々も，音声認識結果に対してブースティングを用いることで，システム要求発話の柔軟性・多様性を保持したまま自動的に区別を行う手法の提案を行ってきた[5]．ただし，音声認識結果を用いることから，認識誤りがシステム要求の検出に悪影響を与えるという問題があった．認識誤りの影響を抑えるためには，1-bestの認識結果だけではなく，認識仮説まで含めた情報を用いてシステム要求検出を行うことが効果的であると考えられる．ただし，認識仮説を用いてシステム要求検出を行う場合には，システム要求検出のスコアと認識仮説のスコア（音響スコア・言語スコア）を統合する必要がある．本研究においても，従来と同様，システム要求検出にはブースティングを用いた．ブースティングの出力スコアは確率ではないため，そのまま認識仮説のスコアと統合するには問題がある．そこで，本研究では，sigmoid関数を用いてブースティングの出力スコアを擬似的に確率に変換することにより認識仮説との統合を行い，統一的な尺度で認識仮説の選択，及びシステム要求か雑談かの判別を行った．

以下，次章では本研究で用いたシステム要求判別タスクの概要について述べ，3. 章で提案手法について述べる．4. 章で評価実験と考察について述べ，5. 章でまとめる．

2. システム要求判別タスク

本章では，本研究で用いたシステム要求判別タスクの概要について述べる．本タスクでは，まず，二人以上の人間とシステムが同時に存在することを想定する．これは，ロボットを操作する際に周囲に人がいる場合や，カーナビを操作する際に助手席に同乗者がいる場合のように，自然な状況であると考えられる．二人以上の人間が互いに会話を行いながら，任意にシステムへの要求発話を行う．本研究では，“システム”として，図1のようなロボットを用いた．ロボットの機能を表1にまとめる．典型的な利用方法としては，少し離れた場所から「こっちに来て」とロボットを呼ぶ，「写真を撮って」と写真を撮ってもらう，などがある．

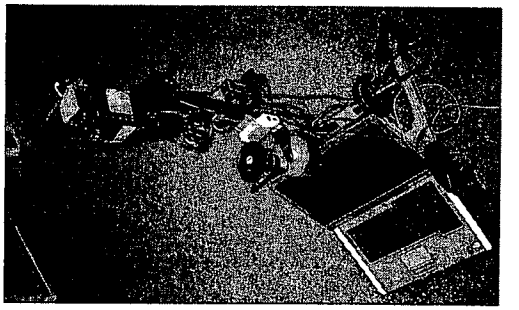

図1　本研究で用いたロボット
Fig. 1　An image of the robot.

現状でロボットが受理できるコマンドは表1の通り決まった文章のみである．ただし，本研究のために収録した音声には，「こっち来て」「早く来いよ」「あっち行って」のように，ロボットは受理できないものの，同じ動作を期待する表現も含まれており，これらもシステム要求発話であるとのラベルを付与した．システム要求ではない発話としては，通常の雑談に加え，「こっちに来て，とか言うと……」「こっちに来て，向こうへ行ってだけでは……」のようにシステムへの要求発話を含むような発話も含まれている．

収録は，二人の発話者それぞれの胸元に取り付けたマイクで行った．発話数は330で，内49発話がシステム要求発話であった．書き起こしは，ひらがな，及びかな漢字混じりの二種類作成した．ひらがなのものは音響モデル適応用である．

ロボットを用いた従来研究では，視線情報を用いるものも存在する．ただし，本研究では，発話者とロボットとの距離が離れており視線方向の推定が困難であったこと，また，発話者がロボットを見ずにシステム要求発話を行う場合もあること，将来的にロボットではなく，カーナビでの利用も想定していることから，システム要求発話の判別に，視線情報は用いない．

3. 提案手法

本研究では，システム要求と雑談の判別を音声認識の中に組み込む手法を提案する．これにより，1-bestの認識結果ではなく，認識の仮説を含む多くの情報をシステム要求検出に用いることが可能となる．また，システム要求検出の結果と音声認識の結果の整合性をとることも可能となる．システム要求か否かを $s \in (request, chat)$，

表 1　本研究で用いたロボットの機能
Table 1　Abilities of the robot.

機能	CSP による音源到来方向推定 音源方向/反対方向への移動 障害物の回避 アームによるボトルの設置 写真の撮影
コマンド例	こっちに来て 向こうへ行って 写真を撮って ついて来て ボトルを置いて

音声認識単語列を $\mathbf{W} = (w_1, \cdots, w_n)$，観測音声特徴系列を $\mathbf{O} = (o_1, \cdots, o_t)$ とすると，音声認識と統合されたシステム要求検出は以下のように定式化される．

$$(\hat{s}, \hat{\mathbf{W}}) = \underset{(s,\mathbf{W})}{\operatorname{argmax}} P(s, \mathbf{W}|\mathbf{O}) = \underset{(s,\mathbf{W})}{\operatorname{argmax}} P(\mathbf{O})^{-1} P(s, \mathbf{W}, \mathbf{O})$$

ここで，$P(s, \mathbf{W}, \mathbf{O})$ をベイズの定理により，以下の二通りに展開できる．

$$P(s, \mathbf{W}, \mathbf{O}) = P(s) \cdot P(\mathbf{W}|s) \cdot P(\mathbf{O}|\mathbf{W}, s) \quad (1)$$

$$P(s, \mathbf{W}, \mathbf{O}) = P(\mathbf{W}) \cdot P(\mathbf{O}|\mathbf{W}) \cdot P(s|\mathbf{W}, \mathbf{O}) \quad (2)$$

式 1 の定式化は，言語モデル・音響モデルが s に依存するようなモデルを用いる手法となる．ただし，本研究では，音響モデルの s への依存は無視し，s に依存する言語モデルのみを考慮した．すなわち，

$$P(\mathbf{W}|s) = \prod_i P(w_i|w_{i-1}, \cdots, w_{i-N+1}, s) \quad (3)$$

という s に依存した N-gram を言語モデルとして用いた．システム要求，及び雑談それぞれに対して言語モデルを構築し，それぞれを用いて認識を行った結果，確率の高い s を採用する手法となる．

式 2 の定式化は，言語モデル・音響モデルは通常のものを用いる．加えて，認識仮説 $\mathbf{W}$ や観測音声 $\mathbf{O}$ から直接 s を推定するような確率モデルが存在する．$\mathbf{W}$ や $\mathbf{O}$ から直接 s を推定するという意味において，このモデルは識別的なモデルである．識別的なモデルとしては，Support Vector Machines (SVM) やブースティングが考えられる．特に，我々は従来，ブースティングを用いてシステム要求と雑談の判別を行う手法について提案を行ってきた [5]．このことから，このモデルとしてブースティングを採用することを考える．しかし一方で，ブースティングは確率に基づく手法ではないため，音響モデル，及び言語モデルとの整合性をとるためには，ブースティングの出力結果を確率化する必要がある．本研究では，完全な確率とは言えないものの，ブースティングの出力スコアを sigmoid 関数を用いて疑似確率化して用いるものとした．sigmoid 関数は図 2 に示すような関数であり，最小値 0，最大値 1 となる．また，識別境界付近を詳細にモデル化できる特徴を持つ．ブースティングの出力スコア $f(\mathbf{W}, \mathbf{O})$ を sidmoid 関数で疑似確率化することにより，モデル $P(s|\mathbf{W}, \mathbf{O})$ は，

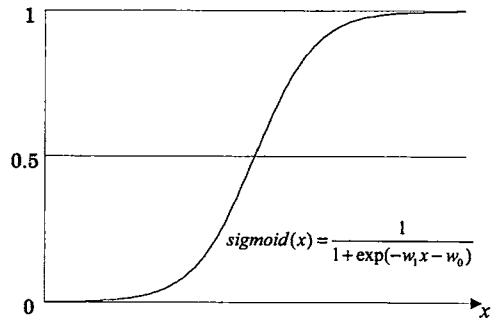

図 2　sigmoid 関数
Fig. 2　sigmoid function.

$$P(s = request|\mathbf{W}, \mathbf{O}) = sigmoid(f(\mathbf{W}, \mathbf{O}))$$

$$P(s = chat|\mathbf{W}, \mathbf{O}) = 1 - sigmoid(f(\mathbf{W}, \mathbf{O}))$$

と定式化できる．ここで，w_1 及び w_0 は sigmoid 関数における重み係数であり学習により推定する．ただし，本研究では，$\mathbf{O}$ は用いず，$\mathbf{W}$ のみを用いた．次節で，ここで用いたブースティングアルゴリズムについて述べる．

3.1　ブースティング

本研究では，ブースティングの手法として AdaBoost を用いた．AdaBoost は，いくつもの識別器を組み合わせてひとつの高度な識別器を構成する *ensenble learning method* のひとつである．Schapire ら [6] が提案している学習のアルゴリズムを図 3 に示す．図中，I は，$I(true)$ ならば 1，$I(false)$ ならば -1 となる．ϵ_t が 0.5 未満の弱学習器を見つけ続けることができれば，学習誤差 0 の最終学習仮説を生成できる．また，未知のサンプルに対する汎化誤差も小さくできることが実験的に報告されている [7], [8]．一方，雑音を有するサンプルの場合，過学習を起こすことが報告されている．これに対しては，AdaBoost の学習過程をマージン最大化ととらえ，SVM における Soft Margins の概念を導入した手法も提案されている [9], [10]．本研究では，認識結果を扱うため，サンプルには多くの雑音が乗っているものと考えられる．このことから，通常の AdaBoost ではなく，Soft Margins 付きの AdaBoost を用いることとした．

AdaBoost を用いたテキスト分類手法としては，文献 [6], [11] などが提案されている．これらの文献では，テキスト分類のための弱学習器として，Decision Stumps が用いられている．Decision Stumps とは，ある素性の有無に基づいて分類を行う単純な手法である．素性には，単語や単語 bi-gram，ラベル付き順序木などが用いられ

Input: n examples $Z = \{(\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_n, y_n)\}$

Initialize: $w_1(\mathbf{z}_i) = 1/n$ for all $i = 1 \cdots n$

Do for $t = 1, \cdots, T$,

(1) Train a base learner with respect to weighted example distribution w_t and obtain hypothesis $h_t : \mathbf{x} \mapsto \{-1, 1\}$

(2) Calculate the training error ϵ_t of h_t:

$$\epsilon_t = \sum_{i=1}^{n} w_t(\mathbf{z}_i) \frac{I(h_t(\mathbf{x}_i) \neq y_i) + 1}{2}.$$

(3) Set

$$\alpha_t = log \frac{\epsilon_t}{1 - \epsilon_t}$$

(4) Update example distribution w_t:

$$w_{t+1}(\mathbf{z}_i) = \frac{w_t(\mathbf{z}_i) \exp\{\alpha_t I(h_t(\mathbf{x}_i) \neq y_i)\}}{\sum_{j=1}^{n} w_t(\mathbf{z}_i) \exp\{\alpha_t I(h_t(\mathbf{x}_i) \neq y_i)\}}.$$

Output: final hypothesis:

$$f(\mathbf{x}) = \frac{1}{||\alpha||_1} \sum_{t=1}^{T} \alpha_t h_t(\mathbf{x}).$$

図 3 AdaBoost のアルゴリズム
Fig. 3 AdaBoost algorithm.

る．学習時には，学習サンプルを最もうまく分類するような“素性”を選択し，その際の重みを得る．識別時には，学習によって得られた全ての素性について，サンプル中にその素性があれば，クラス y に重み α の投票を行うということを繰り返し，最終的に重みの大きかったクラスと判別する．

3.2 ブースティングと音声認識の統合

前述の通り，ブースティングによる出力スコアを sigmoid 関数に当てはめることにより疑似確率化し，音声認識との統合を行う．弱識別器の数を T 個，弱識別器を $h_t(\mathbf{W})$，弱識別器の重みを α_t とすると，

$$P(s|\mathbf{W}, \mathbf{O}) = sigmoid(\sum_t \alpha_t f_t(\mathbf{W})) = \frac{1}{1 + \exp(-w_1 \sum_t \alpha_t f_t(\mathbf{W}) - w_0)}$$

となる（$s = request$ の場合）．sigmoid 関数のパラメータ w_1 及び w_0 は勾配法により学習する．ただし，学習をしすぎてしまうと sigmoid 関数の識別境界付近の勾配が急峻になりすぎてしまい，0 か 1 かの二値に近づいてしまうため，ある程度学習された時点で推定を止めるものとした．

ブースティングによるシステム要求検出と音声認識（音響モデル・言語モデル）を統合することにより，1-best の認識結果のみからシステム要求検出を行うのではなく，複数の認識仮説を利用できるようになる．また，その際，どの仮説からのシステム要求検出結果を用いればいいかについて，確率の枠組みの中で統一的に解を選択することが可能となる．例として，図 4 のような場合を考える．ここでは，「こっちに来て」「こっちにして」の 2 つの仮説が考えられる．次に，それぞれの仮説に対し，ブースティングによってスコアを求める．例として，「こっちに来て」は高いスコアでシステム要求，「こっちにして」は低いスコアで雑談と判定されるものとする．このスコアと音声認識の確率との統合を行う．ここで，ブースティングによる出力スコアは確率ではないため（値域も 0〜1 ではない），sigmoid 関数を用いて疑似確率化を行う．例では，「こっちに来て」は高確率でシステム要求，「こっちにして」は中確率で雑談，のようにブースティング出力スコアが（疑似）確率に変換される．統合後，最も確率の高い仮説が選択される．例えば，2 つの仮説の音声認識確率が同程度であれば，システム要求発話として自然な「こっちに来て」が選択される．逆に，2 つの仮説のうち，「こっちにして」の音声認識確率が顕著に高ければ，ブースティングによる出力スコアの違いを超えて「こっちにして」が選択される．システム要求検出と音声認識を確率の枠組みで統合することにより，認識仮説をシステム要求検出に利用することが可能となり，また，統一的な基準を基にして認識仮説，及びシステム要求か雑談かを判別することができる．

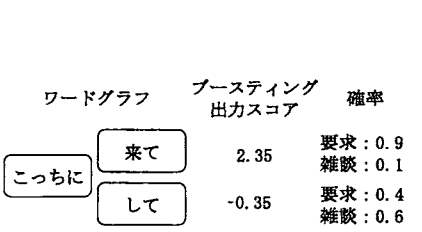

図 4 ブースティングによる出力スコアの確率化例
Fig. 4 An example of conversion of boosting score to probabilities.

4. 実　　験

本節では，提案手法を用いたシステム要求検出実験について述べる．実験のタスクとして，2. 節で述べたものを用いた．ブースティングの学習は，音声認識器による 1-best の認識結果を学習データとして行った．また，このとき学習されたブースティングのモデルと，同一の学習データから sigmoid 関数のパラメータの学習を行った．その後，提案手法による認識を行った．実験はすべて，10 folds のクロスバリデーションによって行った．評価は，システム要求検出の再現率・適合率・F 値によって行った．

比較手法として，式 1 に基づく要求依存言語モデルを用いた手法，従来提案してきた認識結果を用いてブースティングを行う手法を用いた．また，システム要求発話のみで trigram を構築し，認識結果の単語信頼度の平均を閾値で区別する手法も比較に用いた．

次節で，実験で用いた音声認識器について述べる．

4.1 音声認識条件と結果

ベースラインの音響モデルは，日本語話し言葉コーパス (CSJ: Corpus of Spontaneous Japanese) モニター

表 2　音響分析条件と HMM の仕様
Table 2　Condition of acoustic analysis and HMM specification.

音響分析	サンプリング周波数	16kHz
	特徴パラメータ	MFCC(25 次元)
	フレーム長	20ms
	フレーム周期	10ms
	窓タイプ	ハミング窓
HMM	タイプ	244 音節
	混合数	32 混合
	母音 (V)	5 状態 3 ループ
	子音+母音 (CV)	7 状態 5 ループ

表 3　音声認識結果の単語正解精度.
Table 3　Word accuracy of the speech recognition results.

	音響モデル	
	オープン	クローズド
単語正解精度	29.8%	42.1%

版 [12] のうち，男性話者 200 名の講演音声を用いて作成した．音響分析条件と HMM の仕様を表 2 に示す．これらの条件で音響モデルを作成し，さらに，MLLR+MAP [13] により音響モデル適応を行った．音響モデル適応は，テストセットを含めた適応をクローズド，含めない適応をオープンとした．ただし，どちらの場合も適応データの話者は，テストセットと同一のものを用いた．適応データの分量は，クローズドの場合で約 10 分，オープンの場合で約 5 分であった．

言語モデルは，実験で用いた発話を書き起こしたテキストから作成した．ただし，テストセットに対しオープンとなるように，話者 B の発話を用いて話者 A の認識用言語モデルを作成した．

音声認識による単語正解精度を表 3 に示す．音響モデルをオープンにした場合は単語正解精度が 30%を下回ってしまう．将来的に，適応に用いるデータを増やした際，オープンでもクローズド程度の性能に近づくことを期待して，音響モデルはクローズドの方を用いることとした．

4.2　システム要求識別モデルの学習

前節の条件で音声認識を行い，得られた 1-best の認識結果を用いてブースティングの学習を行った．素性には，unigram 及び bigram を用いた．このときに選択された素性後の例を表 4 に示す．システム要求に投票を行う素性には bigram が多く選択されている．これは，システム要求発話がある程度決まったフレーズによって行われているためと考えられる．また，システム要求の素性には <s> や </s> との組み合わせが多数存在することから，システム要求発話は，ある程度会話を区切った上で為されているものと考えられる．

ここで得られたブースティングモデル，及び 1-best の認識結果から sigmoid 関数のパラメータを推定した．パラメータの推定は勾配法を用いて繰り返し計算で行う．

表 4　AdaBoost によって選択された素性語
Table 4　Selected features by AdaBoost.

システム要求	雑談
<s>+ここ <s>+向こう <s>+写真 ください+</s> ここ+は て+</s> て+ください 場所 来+て 来い+</s> etc.	あー これ たら って とか ない なんか まあ やつ 言っ etc.

ただし，繰り返しすぎると，sigmoid 関数の識別境界付近の傾きが急峻になりすぎてしまい，出力が 0 か 1 かの二値に近くなってしまう．そこで，パラメータの更新が緩やかになって来た時点，本研究では，パラメータの更新の比率が 0.01 以下となった時点で学習を止めるようにした．

4.3　システム要求検出

4.1 の条件で音声認識を行い，出力されたワードグラフからシステム要求検出実験を行った．

システム要求か否かに依存した言語モデルを用いた手法を “multi N-gram”，ブースティングによる識別結果を sigmoid により疑似確率化した手法を “sig-Boosting”，従来の認識結果を用いたブースティングによる手法を “Boosting”，trigram の信頼度を用いる手法を “Confidence” とする．結果として，それぞれの手法において F 値が最も高かったケースを図 5 に示す．

実験の結果，提案手法のうち，ブースティングの結果を sigmoid により疑似確率化して用いる手法が最も良い性能を示した．次いで，認識結果を用いたブースティングが良い性能を示した．元々ブースティングによる識別が高性能な上に，提案手法では，音声認識誤りを原因とする識別誤りが，ワードグラフ中から正しい単語を拾ってくることで解決し，性能が向上したものと考えられる．一方，システム要求か否かに依存する言語モデルを用いる手法や信頼度を用いる手法では，性能が低下した．特に，「こっちに来て，とか」のように「とか」という 1 単語の有無によって結果が左右されるような発話の識別を多く誤った．これは，文の大半がシステム要求発話と同一であるため，システム要求に依存した言語モデルの方が高いスコアを出力してしまうためであると考えられる．同様に，雑談の中にシステム要求が混ざるような場合も雑談と識別してしまう傾向が見られた．信頼度を用いた手法では，multi N-gram と同様に，「こっちに来て，とか」の様な発話の識別を誤ってしまう傾向がみられた．

5.　ま と め

本稿では，システム要求検出を音声認識の定式化の中に組み込むことで，認識仮説を利用して要求検出を行う手法について述べた．ベイズの定理により，音響モデル・言語モデル，及び単語仮説から識別的にシステム要求検出を行うモデルを導いた．識別的にシステム要求検出を

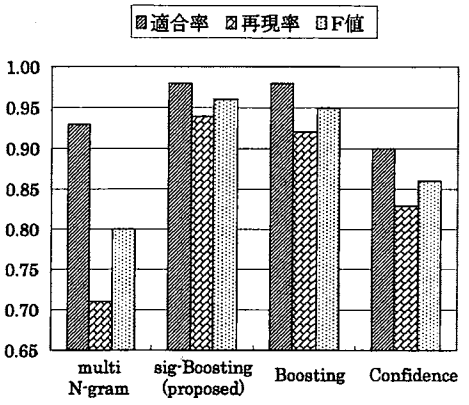

図 5 システム要求判別結果

Fig. 5 Result of system request discrimination.

行うモデルには，ブースティングの出力スコアを sigmoid 関数を用いて疑似確率化したものを用いることを提案した．これにより，音響モデル・言語モデル・要求検出モデルの3つのスコアを統一的に扱うことが可能となり，適切な認識仮説，及びその要求検出結果を選択することができた．実験の結果，提案手法が最も良い性能を示した．1-best の認識結果が誤っている場合に，ワードグラフ中から正しい単語を拾うことで識別性能が向上したものと考えられる．他方で，要求に依存した言語モデルを用いる手法では，本タスクでは，ひとつの単語の有無により結果が変わってしまうことがあるため，N-gram はそのような識別には不向きであると考えられる．

今後の課題として，大規模なコーパスを構築し実験を行うこと，多様な表現によりシステム要求が可能なタスクにおいて実験を行うことがあげられる．

文　献

[1] 堀内靖雄，庵原彩子，西田昌史，市川熹，“自然対話における聞き手の反応と話し手のうなずき・言語情報・韻律情報との関係に関する予備的検討，” 情処学研報，SLP-52, pp.93-98, July 2004.

[2] 伊藤 敏彦，山田 真也，荒木健治，“音声対話認識率や状況の違いによる音声対話の言語的・音響的特徴の比較，” 情報処理学会研究会報告 2005-SLP-56, pp.101-106(2005年5月).

[3] 杉本夏樹，北岡教英，中川聖一，“音響特徴を用いた対システム発話と対人間発話の識別，” 電子情報通信学会，総合大会，D-14-9, pp.133 (2006.3)

[4] 石塚健太郎，河原達也，堂下修司，“発話検証用モデルを用いた音声操作プロジェクタ，” 信学技報，SP98-5, pp.33–38, Apr. 1998.

[5] 佐古淳，滝口哲也，有木康雄: “AdaBoost を用いたシステムへの問い合わせと雑談の判別，” 第8回音声言語シンポジウム，SIG-SLP64, pp.19-24, 2006-12.

[6] R.Schapire, Y.Freund, P.Bartlett, and W.Lee, “Boosting the margin: A new explanation for the effectiveness of voting mathods,” Annals of Statistics, vol.26, no.5, pp.1651-1686, Oct. 1998.

[7] Y.Freund and R.Schapire, “Experiments with a new Boosting algorithm,” Proc. 13th International Conference on Machine Learning Bari, Italy Morgan Kaufmann, pp.148-146, July 1996.

[8] H.Schwenk and Y.Bengio, “Adaboosting neural networks,” Proc. ICANN'97, vol.1327 of LNCS Berlin Springer, pp.967-972, Oct. 1997.

[9] G.Ratsch, T.Onoda, and K.-R. Muller, “Soft Margin for AdaBoost,” Machine Learning, vol.42, no.3, pp.287-320, March 2001.

[10] 小野田崇，“Boosting の過学習とその回避，” 電子情報通信学会論文誌，Vol.J85-D2, No.5, pp. 776-784, 2002 年 5 月.

[11] 工藤拓／松本裕治，“半構造化テキストの分類のためのブースティングアルゴリズム，” 情報処理学会論文誌，Vol.45, NO.9, 2004 年 9 月.

[12] 古井貞熙，前川喜久雄，伊佐原均，“『話し言葉工学』プロジェクトのこれまでの成果と展望，” 第2回話し言葉の科学と工学ワークショップ，pp.1-6, 2002.

[13] 緒方淳，有木康雄，“音素事後確率に基づく信頼度を用いた音響モデルの教師なし適応，” 信学技報，SP2001-105, 2001.