

携帯電話の操作方法を対象とした中国語話し言葉による音声検索

石川 開 赤峯 享 花沢 健

NEC メディア情報研究所

{k-ishikawa@dq, s-akamine@ak, k-hanazawa@cq}.jp.nec.com

1. はじめに

近年、携帯電話などの情報機器の多機能化が急速に進み、利用者が機器の使い方を全て把握することが困難になっている。そのため、利用者は、機器の使い方が分からなかったり、トラブルが生じた際に、その都度、マニュアルや FAQ を調べたり、友人やコンタクトセンターに問い合わせることで、使い方を理解している。我々はこれまで、このようなプロセスを自動化し、何時でも誰でも手軽に使える音声入力によるテキスト検索方式を提案してきた[1]。

今回、携帯電話の操作方法に関する要求文として、約1万文のバリエーションを収集し、これを同一の要求ごとにまとめ、要求 ID に対する要求文のバリエーションと、正解マニュアル文書の組を作成し、検索のモデル化を行った。本稿では、利用者の要求を識別することで、利用者が入力する特定分野の要求に対して適合率の高い検索方式を提案する。本方式では、(1) 利用者の情報要求を識別する構文解析ベースの識別器を構築し、(2) 文字組ベースの類似文書検索と組み合わせる。我々は、本方式を用いて、携帯電話の操作方法に関する質問や操作要求を中国語で音声検索できるプロトタイプシステムを開発し、PDA 上で動作を確認した。また、中国語の携帯電話マニュアルを対象とする評価実験を行い、提案手法の有効性を示した。

2. 操作方法を対象とした検索のモデル化

2.1. 類似文書検索を適用する問題

中国語の携帯電話の操作方法に関する要求文の例を表 1 に示す。これらはいずれも「天気予報を知る方法」という同一の要求 (ID=37) に関する要求文であり、ユーザーマニュアルの同じ箇所を回答として利用者に提示すれば問題が解決できる。このような情報要求は、利用者の最終目的が特定の問題解決であるという点で、製品ポータルサイトやコンタクトセンターで利用される FAQ の検索と本質的に似ている。いずれも、情報提供側は特定の製品やサービスに特化して、利用者の質問やトラブルごとに、対応する回答や解決方法に関する情報を管理・提示する。利用者は、特定の機能を使うことや、直面するトラブルや疑問の解決を

望んでいるため、問題解決に直結する情報が提示されなければ満足しない。このような検索要求に対して、利用者参照用にテキスト化されたマニュアルや FAQ を検索対象とし、類似文書検索のモデルを適用すると、次のような問題を生じる。

(1) 特定ドメインにおける内容語の重複

特定機器の操作方はドメインが限定されるため、要求内容の異なる文であっても、特徴語は重複することが多い。類似文書検索は、要求文中の特徴語を多く含む文書が検索上位候補となるため、正解文書が要求内容の異なる候補中に埋もれやすい。

(2) 同じ要求内容に対する表現の多様性

要求内容が同じでも、利用者によって、要求文には様々な表現が使われる。類似文書検索は、要求文と検索対象の表現が一致しないと、類似性を正しく評価することが難しいため、正解の検索漏れが生じる。

2.2. 要求を識別するアプローチ

携帯電話の操作方法に関する要求文のバリエーションに基づいて、利用者の要求を的確に識別するモデルを検討した。ここでは表 2 に示すような識別対象文を例に、要求の識別問題について説明する。この識別対象文は、要求 C「待ち受け画面の時刻合わせの方法」に関する文であり、特徴語 (□ で囲まれた語) を 4 つ含んでいる。一方、バリエーション中には、要求 C 以外にも同じ特徴語を持つ要求 A、B の文が存在する。また、同じ C の文でも、例のように、特徴語は同義語や類義語などの多様性があり、同じ特徴語を持つとは限らない。このような要求の識別問題に対しては、

表 1: 「天気予報を知る方法」の要求文の例

要求 ID	要求文 (□ は要求の特徴語)
37	/手机/如何/实现/天气/查询/的/功能/
37	/手机/能否/查询/天气预报/
37	/天气/查询/功能/该/如何/实现/
37	/天气预报/查询/步骤/
37	/天气预报/自动/查询/功能/
37	/天气预报/是/如何/查询/的/
37	/天气预报/的/查询/方法/是/什么/

表 2: 要求の識別問題

識別対象文

/怎样/更改/待机/画面/的/时间/
(要求 C: 待ち受け画面の時刻合わせの方法)

同じ特徴語を持つバリエーション中の文

/如何/更改/待机/画面/的/文字/
(要求 A: 待ち受け画面の挨拶文の変更方法)

/怎么/设置/待机/画面/
(要求 B: 待ち受け画面の設定方法)

/怎么/设置/显示/画面/的/时间/
(要求 C: 待ち受け画面の時刻合わせの方法)

識別誤りが少なく、かつ、表現の多様性に対応できることが重要である。

そこで我々は、構文構造に基づくアプローチを考えた。図 1 にその概要を模式的に示す。まず、個々の情報要求に依存しない共通の辞書と文法を用いて、要求文に対する構文構造(図の上段)を得る。この構文構造の特徴語を網羅する句構造 VP に着目し、VP を親ノードとする次のような要求固有の句構造ルールと辞書ルールを作成する。

要求 C 固有の句構造ルール $S_c \rightarrow r VP_c$ $VP_c \rightarrow s_c \text{ 设置 } s_c \text{ 显示 } s_c \text{ 画面 } p s_c \text{ 时间}$ 要求 C 固有の辞書ルール $s_c \text{ 设置} \rightarrow \text{更改}$ $s_c \text{ 显示} \rightarrow \text{待机}$ $s_c \text{ 画面} \rightarrow \text{画面}$ $s_c \text{ 时间} \rightarrow \text{时间}$
--

これらのルールを共通ルールと共に用いて構文解析を行うと、下段のような構文構造が最尤構造として得られる。この構造の初期記号 S_c から、識別対象文が要求 C に関する文であると識別することができる。

3. 検索方式の提案

前章で述べた要求識別方式は、高い適合率が期待できる一方、未知の要求文に対する頑健性に課題がある。そこで、構文解析ベースの要求識別と、多様な検索要求文に対する頑健性の高い文字組ベースの類似文書検索を組み合わせた検索方式を提案する。

(1) 構文解析ベースの要求識別器

前章で述べたアプローチに基づき、本提案方式では、文脈自由文法において、入力文の文法構造を得るための共通の辞書ルール、句構造ルールと共に、非終端記号が要求 ID を素性に持つ要求固有ルールを用いる。終端記号 V_T を単語の集合、 V_C を品詞の集合、 V_X を句記号の集合、 V_{C_i} (i は要求 ID、 $i=1, \dots, N$) を要求固有の同義語の集合、 V_{X_i} を要求固有の句記号の集合とすると、本方式で用いる辞書ルールと句構

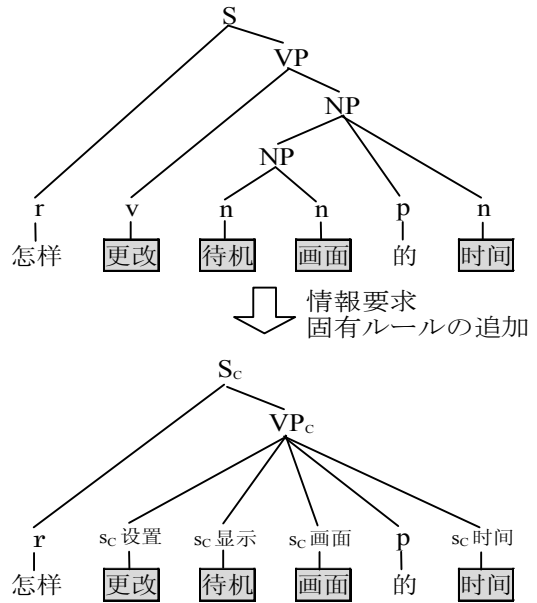


図 1: 構文構造に基づいた識別方法

造ルールは次のように表現される。

共通の辞書ルール $A \rightarrow B \quad A \in V_C, B \in V_T$ 共通の句構造ルール $A \rightarrow \alpha \quad A \in V_X, \alpha \in (V_C \cup V_X)^+$ 要求情報 ID=i 固有の辞書ルール $A \rightarrow B \quad A \in V_{C_i}, B \in V_T$ 要求情報 ID=i 固有の句構造ルール $A \rightarrow \alpha \quad A \in V_{X_i}, \alpha \in (V_T \cup V_C \cup V_{C_i} \cup V_X \cup V_{X_i})^+$

このようなルールを用いて入力文の文法構造を求め、最尤構造の初期記号 $\sigma (\in V_X, V_{X_1}, \dots, V_{X_N})$ の素性 i から入力文の要求 ID を識別する。本稿では最尤構造を求めるために各ルールの出現頻度が等しいと仮定した確率的文脈自由文法を用いる。

(2) 文字組ベースの類似文書検索

本稿では、類似文書検索として、文字組(文字 **uni-gram** および文字 **bi-gram**)ベースの検索を用いる。要求文 Q に対する文書 D の検索スコア $Score(D, Q)$ を次の Okapi BM25 式を用いて計算する[2]。

$$Score(D, Q) = \sum_{T \in Q} w^{(1)} \cdot \frac{(k_1 + 1) \cdot tf}{K + tf} \cdot \frac{(k_3 + 1) \cdot qtf}{k_3 + qtf}$$

$$w^{(1)} = \log \frac{N - n + 0.5}{n + 0.5}$$

$$K = k_1 \cdot \left((1 - b) + b \cdot \frac{dl}{avdl} \right)$$

ここで、 T は要求文 Q 中の文字列、 N は全文書数、 n は T を含む文書数、 tf は文書 D における T の出現頻度、 qtf は Q に含まれる T の出現頻度、 dl は D に含ま

れる文字列数、 $avdl$ は全文書に含まれる文字列数の平均を表す。 k_1, k_3 、および、 b は定数で、本稿ではそれぞれ $k_1=1, k_3=1000$ 、および、 $b=1$ を用いる。

(3) 検索結果のマージ

あらかじめ、各要求 ID と検索対象中の正解文書との対応関係を作成・保持しておく。要求識別器が識別した要求 ID に対応する正解文書を、検索結果の最上位に並べ、それに続く候補として、類似文書検索の結果をスコア順に並べる。なお、要求識別器の構文解析が失敗した場合は類似文書検索の結果のみを用いる。

4. プロトタイプシステム

提案した検索方式を音声認識と組み合わせて、市販の PDA (OS: PocketPC2003, CPU: Intel PXA255 (400MHz), メモリ: 64MB) 上で動作する音声検索システムを開発した。

検索対象は、中国の携帯電話(N840)のマニュアルから 94 文書(表 3)を抽出した。また、これとは別に、一般的な携帯電話の操作方法に関する55個の要求を定義し、その回答文書を 94 文書に求めることで、要求と回答との対応関係を作成した。

図 2 にシステム構成の概要を示す。利用者は携帯電話の操作に関する要求文を中国語で音声入力する。音声認識モジュールは認識結果を単語列で出力する。この認識結果から、要求識別モジュールは要求 ID (全 55 個)の識別を行い、類似文書検索モジュールは類似文書検索を行う。検索結果マージモジュールでは、要求識別と類似文書検索の結果を用いて最終的な検索結果を生成する (図 3)。

表 3: 検索対象文書のタイトルの例

検索対象文書タイトル	日本語訳
待机时间与通话时间	待受時間と通話時間
检查电池	電池残量の確認
插入和取出 miniSD 卡	ミニ SD カードの挿入取出し
开机	電源を入れる
接听一个来电	電話の受けかた

4.1. 音声認識モジュール

音声認識モジュールは、利用者が入力する中国語音声に対して、大語彙連続音声認識を行う。音響モ

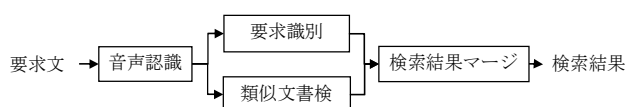


図 2: システム構成

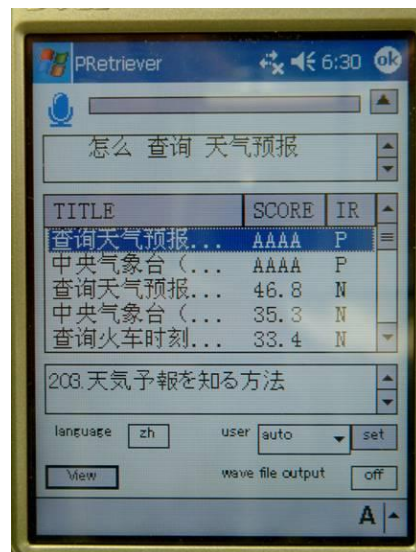


図 3: プロトタイプシステムの検索画面

デルは、中国語の標準語である普通話の話し言葉を対象に開発したものをを用いた[3]。言語モデルは、携帯電話の操作方法に関する要求文のバリエーション約 10,000 文を学習データとして、単語 N-gram モデルを構築した。

4.2. 要求識別器のルール作成

言語モデルで使用した要求文約 10,000 文のうち、55 個の要求に関する要求文 4846 文を要求識別用の訓練セットとし、共通の句構造・辞書ルールおよび要求固有の句構造・辞書ルールを手で構築した。

5. システムの評価

5.1. 評価セット

55 個の要求に対応し、訓練セットに含まれない要求文 101 文を評価セットとして作成した。さらに、音声検索の性能を評価するために、中国語ネイティブ 5 名の収録音声を用いた。検索対象文書は、中国の携帯電話のマニュアルから作成した 94 文書である。

5.2. 要求の同定性能の評価

要求識別用の訓練セットと検索評価用の評価セットにおける要求の識別性能を表 4 に示す。ここで、適合率は、識別結果の出力文数に対する正解文数の割合を表し、再現率は、全データセットの文数に対する識別結果の出力文数の割合を表す。また、評価セットの

表 4: 要求の識別性能

データセット	訓練セット (4846 文)	評価セット(101 文)	
		書き起こし	音声
適合率(%)	95.9	79.4	81.47
再現率(%)	50.4	33.7	36.4

表 5: 類似文書検索のテキスト検索性能

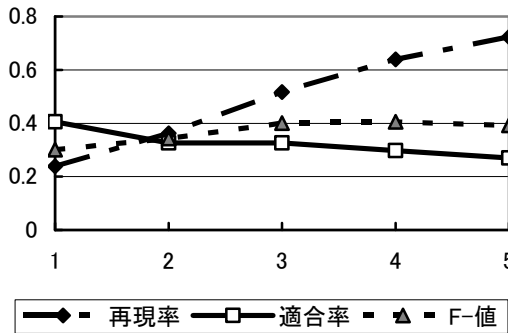
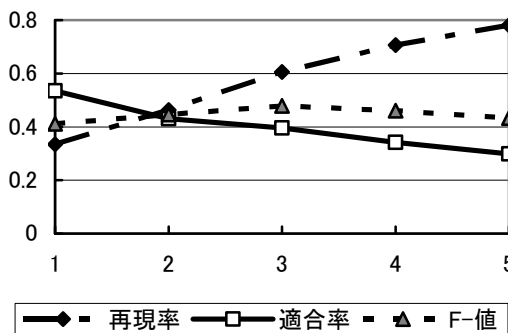


表 6: 提案手法のテキスト検索性能



うち、「書き起こし」は音声の書き起こしテキストに対する評価結果を示し、「音声」は 5 名の収録音声に対する音声認識処理を含んだ評価結果の平均値を示す。

5.3. 検索性能の比較評価

提案手法の有効性を確認するため、評価セットを対象としたテキスト検索と音声検索の性能を、適合率、再現率に関して類似文書検索と比較評価した。横軸は検索上位の候補数を示す。類似文書検索によるテキスト検索の評価結果を表 5 に、提案手法によるテキスト検索の評価結果を表 6 に示す。さらに、音声検索における検索性能についても、同様に比較評価を行った。類似文書検索による音声検索の評価結果を表 7 に、提案手法による音声検索の評価結果を表 8 に示す。音声検索の評価結果は、評価用に収録した話者 5 名の音声に対する評価結果の平均値を示す。

6. 評価結果の考察

要求識別の適合率は、評価セットで 77.1%であり、類似文書検索の1位の適合率 40.6%を大きく上回ることから、提案方式が適合率の高い識別方法であるといえる。また、音声認識でもテキストと同等な識別性能が得られた。一方、再現率は評価セットで 33.7%と低い値であるが、訓練セットに対しても 50.4%と低い値であることから、今回作成したルールが十分でなかったことによる影響であると考えられる。一方、検索性能に関

表 7: 類似文書検索の音声検索性能

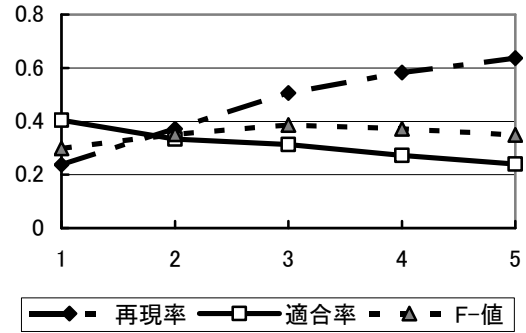
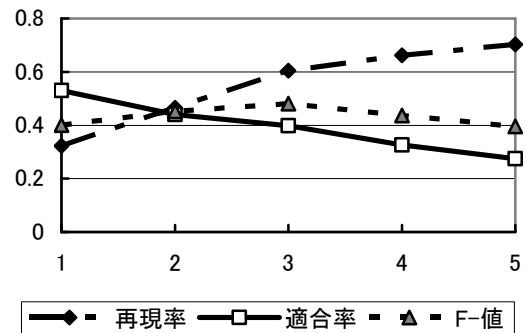


表 8: 提案手法の音声検索性能



する提案手法と類似文書検索の比較評価では、検索上位で再現率、適合率共に提案手法が 10 ポイント以上上回る結果を得た。これにより、適合率の高い要求識別と、再現率の高い類似文書検索の組み合わせが、検索上位の適合率を向上する効果が示された。要求識別の再現率が改善されれば、提案手法の効果はさらに明確になるとと思われる。

7. おわりに

本稿では、要求の識別に基づく適合率の高い検索方式の提案を行った。本方式を用いて、中国の携帯電話の操作方法に関する音声検索システムを開発し、PDA 上で動作を確認した。さらに、要求識別と類似文書検索を組み合わせた本方式が適合率が高いことを実験で確認した。今後、ルールの強化を行い、さらに認識誤りと検索精度の関係について評価を行う予定である。

参考文献

- [1] 池田崇博 他:自由発話音声入力による携帯電話向けテキスト検索システム, 言語処理学会第 10 回年次大会, pp.564-567, 2004.
- [2] S. E. Robertson et al.: Okapi at TREC-3, TREC-3, pp.109-126, 1995.
- [3] 花沢健 他:携帯端末上で動作する日中通訳システムの開発, 情報処理学会第 68 回全国大会, 2006.