

質問応答型検索の音声認識誤りに対するロバスト性向上

筒井 秀樹[†] 真鍋 俊彦[†] 福井 美佳[‡] 藤井 寛子[†] 浦田 耕二[†] 酒井 哲也[†]

†(株)東芝 研究開発センター 知識メディアラボラトリー

‡(株)東芝 研究開発センター 研究企画室

我々はこれまで、質問応答型マルチモーダルヘルプシステムの開発を行ってきた。これはユーザからの質問に対し、映像・音声・取扱説明(テキスト)などで構成される表現力豊かなマルチモーダルコンテンツの検索技術、および、質問内容を理解し、ユーザが必要としている情報に対して的確に回答する質問応答技術を融合することにより、よりわかりやすい情報提供を実現したシステムである。この中で、音声入力による質問を処理する際、音声認識誤りが起きると、その後の処理がうまく行かず、適切な回答ができない場合があった。失敗原因を検討した結果、具体的な時間や量をきく Factoid 型の質問に対する音声認識誤りの影響が大きいことがわかった。これは、音声認識誤りによって疑問詞についての情報が失われることにより、具体的に何を回答すべきかを判定する回答タイプ判定に失敗することが原因であった。そこで今回、音声認識誤りに対するロバスト性向上を目的とし、回答タイプが正しく判定されるように、音声認識結果を補完して検索する手法を開発した。その結果、上位 3 位までの MRR (Mean Reciprocal Rank)による検索精度で、従来手法が 0.429 であったのに対し、今回の手法では 0.597 に向上した。

Improving the Robustness to Recognition Errors in Speech Input Question Answering

Hideki Tsutsui[†] Toshihiko Manabe[†] Mika Fukui[‡] Hiroko Fujii[†] Koji Urata[†] Tetsuya Sakai[†]

† Knowledge Media Laboratory, Corporate R&D Center, TOSHIBA Corp.

‡ Research Planning Office, Corporate R&D Center, TOSHIBA Corp.

We have been developing a multimodal question answering system that combines the search technology for multimodal contents with high expressive power such as video, speech and text, and the factoid question answering technology for understanding the user's information need and extracting exact answers from text. Failure analyses of our system showed that speech recognition errors were fatal for answer type recognition and therefore for the final Mean Reciprocal Rank (MRR) performance, especially with numerical factoid questions. We therefore propose a new method which is robust to speech recognition errors. This method improves our MRR based on top 3 answers from 0.429 to 0.597.

1. はじめに

近年、家電や AV 機器の高機能化やネットワーク化が進み、操作が複雑になってきている。また、多機能化が進み、ユーザが全ての機能を把握して使いこなすのが難しくなっている。製品には必要十分な取扱説明書が付属しているが、コールセンターへの問合わせ事例をみると、取扱説明書に記述されている内容に関する問合わせも少なくない。今後さらに取扱説明書のデータ量が増えるにつれ、知りたい情

報を探せない、操作が複雑で取扱説明書を読んでもわからない、といった問題が増えると考えられる。

そこで、我々は、映像・音声・取扱説明(テキスト)などで構成されるマルチモーダルナレッジ(MM ナレッジ)による表現力豊かなコンテンツの蓄積、検索技術と、質問内容を理解し、ユーザが必要としている情報に対して的確に回答する質問応答(QA)技術を融合し、ユーザに対して、よりわかりやすい情報提供を可能とする質問応答型マルチモーダルヘルプ

(家電ヘルプシステム)を開発し、バイロットシステムをタブレットPC上で試作した[1]。

本家電ヘルプシステムは、ユーザの状況に依存しない検索システムを実現するため、音声による質問入力インターフェースを備えている。また知識源として家電の取扱説明書のテキストと、手段や方法を説明するためのアノテーション付き映像を別途用意し、予め登録しておく。システムは入力された質問の種類によって回答の適切な形式／メディアを選択し、映像、音声、テキストを切り替え、あるいは、組み合わせて回答を提示する。例えば、手段や方法に関する質問、すなわち、How型の質問に対しては、映像やテキストで回答し、数量などを尋ねたFactoid型の質問に対しては、音声(音声合成)とテキストで回答する(図1)。質問の解釈、および、検索処理の部分には、我々が開発している質問応答型検索の評価システムASKMiを利用した[2]。

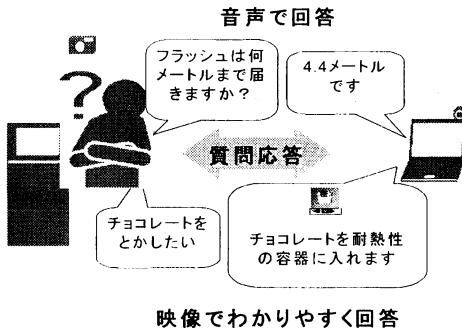


図1 質問応答型マルチモーダルヘルプ

しかし、ユーザからの音声による質問文に音声認識誤りが発生した場合、検索精度が著しく劣化してしまう。特にFactoid型の質問は、「何グラム」や「何メートル」など、回答の種類を判別するための手がかり表現が誤認識されると正しい回答を検索できないという性質から、検索結果が音声認識誤りに大きな影響を受けることがわかった。

関連研究としては、質問応答技術がNTCIRワークショップのQACタスク[3]において扱われているなど、注目を集めている。また、テキスト以外の簡単なインターフェースを持つシステムとして、音声による検索システムが盛んに検討されている。堀ら[4]は、新聞記事などの大規模コーパスを対象とした質問応答システムの研究を行っている。この中で、

質問文を構文解析し、応答対話文を生成することで、質問文の曖昧性を解消している。入力に音声を用いる場合、これまでのキーボード入力と違い、音声認識誤りによる検索精度低下が問題になる。西崎ら[5]は音声キーワードによりニュース音声データベースを検索するシステムを開発している。このなかで、音声認識の次候補まで考慮しキーワード候補間の関連性を用いて音声認識誤りに対する耐性を向上させている。また音声入力を適用した質問応答システムについても研究がなされており、秋葉ら[6]は、単語間のネットワーク文法とN-gramを用いた音声認識用の言語モデルについて検討している。また製品の使い方をナビゲーションする取り組みとして、オンラインヘルプを対象とし、自然言語による質問への回答や、曖昧な質問に対する問い合わせなどの研究が行われている[7]。この中で、N-best候補を用いた検索を行い、候補間での検索結果の違いから検索重要度を求め、検索結果に大きな影響を与える候補の相違箇所を提示してユーザに確認している。

本研究では、音声認識誤りによる検索精度低下を最小限にとどめるため、質問文の音声認識結果を補完する手法について検討し、その効果の評価結果について報告する。具体的には、音声認識誤りに大きな影響を受けるFactoid型の質問において、回答の種類を判別するための手がかり表現を補完する。この補完は、音声認識誤りを含む質問文中の部分文字列と、あらかじめリストアップしておいた手がかり表現の候補リストとの類似性に基づいて行われる。また、文書中の数量、操作部の名前、時間表現などを抽出した回答候補と、それから一定の距離にある単語との共起関係を基に、その単語が質問中にある場合、その回答候補を尋ねる手がかり表現に絞り込んで補完する。

以下では2節で家電ヘルプのバイロットシステムの構成について、3節で音声認識誤りの影響について、4節でFactoid型質問に対するロバスト性向上策、5節で評価実験の結果について説明する。最後に考察とまとめを行う。

2. 家電ヘルプシステムの構成

家電ヘルプのバイロットシステムは図2に示す通りにクライアント・サーバで構成され、クライアント側のヘルプインターフェース、サーバ側のヘルプマネージャ、MMナレッジ検索エンジン、質問応答検索エンジン、メディアサーバから成る。クライアント上のヘルプインターフェースは質問の音声認

表1 質問文のタイプ

質問文のタイプ	検索エンジン	例
Factoid型	質問応答検索	数字を聞く(時間、量など)操作部の名前、場所 いつ どのくらい 時間は 量は どこを どの なにを 例：オートメニューを選ぶキーは
How型	MMナレッジ検索	一連の操作方法 ～の方法 やり方 して欲しい するには どうやる できない 例：夜景を撮りたいんだけど

識、および、音声合成などを用いた検索結果の提示を行う(図3)。より自然な質問応答ができるようになるため、Factoid型の質問とHow型の質問の両方を受け付ける。サーバ上のヘルプマネージャは質問文のタイプ(表1)判定の結果によってMMナレッジ検索エンジンと質問応答検索エンジンのどちらかを呼び出す。

MMナレッジ検索エンジンは、複数の映像、音声、画像のファイル(リアルストリーム)からMPEG7により必要部分を繋ぎ合わせた仮想ストリームを検索対象とし、その時間軸上に関連づけられたテキスト情報(アノテーション)を用いて検索処理を行う[8]。この検索はベクトル空間モデルの文書検索モデルをストリームデータ向けにカスタマイズした手法を採用している[9][10]。検索結果の3位までがヘルプインターフェース上に提示されるが、2位以下はタイトル情報だけが表示される。1位のデータについては、映像であれば、即座に再生できる画面(図4)が、また、取扱説明書のページ画像であれば、それがそのまま表示されるようにヘルプインターフェースの画面が構成される。

質問応答検索エンジンは、予め取扱説明書から数量、操作部の名前、時間表現などの情報を抽出しておく。これら回答候補のデータは抽出元のページと関連づけておく[2][11]。質問に対する回答は3位までをヘルプインターフェース上に提示し、1位の回答については音声合成による読み上げと、根拠ページの提示が行われる(図5)。

本家電ヘルプシステムでの、音声による質問文入力と、音声合成による検索結果の読み上げについては、東芝LaLaVoice2001[12]の連続音声認識、音声合成機能を利用した。

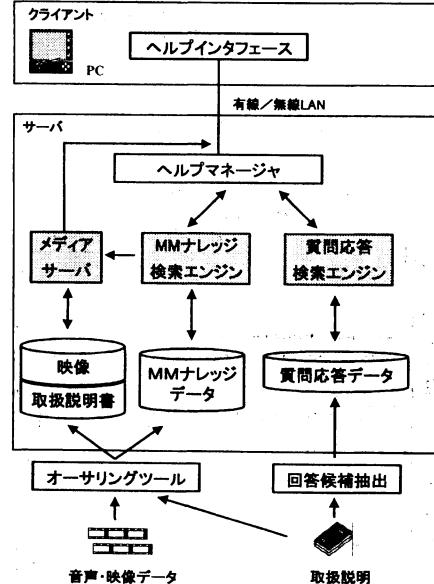


図2 全体構成

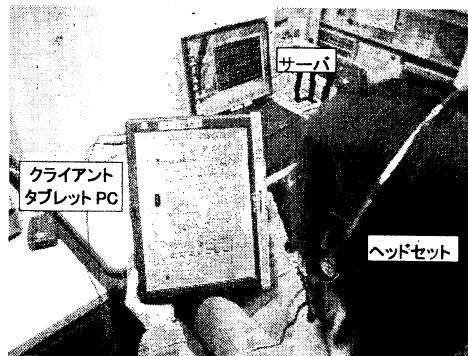


図3 パイロットシステム全景

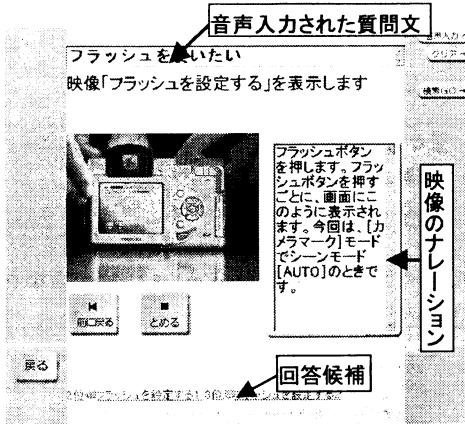


図 4 映像の提示画面

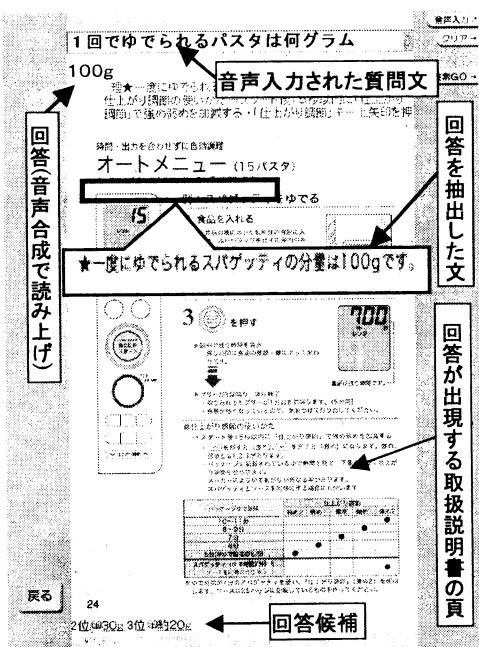


図 5 Factoid 型質問に対する回答画面

3. 従来システムの問題点

前節のパイロットシステムを用いて、男女 10 人の被験者で 96 件の質問に基づいた評価実験を行った。この質問文は、実際にコールセンターへ問い合わせのあったデータから Factoid 型と How 型の質問文に注目し、それらの質問の割合から Factoid 型が 40 件、How 型が 56 件の全 96 件からなる質問セットを構成した。この時、音声認識誤りによる検索精

度劣化を見るために、後で述べる表 3 におけるキーボードによる質問文入力の検索精度が 0.7 付近になるように質問セットを選択している。

検索対象はオープンレンジとデジタルカメラの取扱説明書で、事前に音声認識の辞書について検索対象の文書を用いた語彙学習を行った。学習用の文書には検索対象の取扱説明書を利用し、テキストを LaLaVoice2001 の機能(単語の自動登録&学習)を利用して登録した。この機能は専門用語(未知語)、言い回しについて自動的に行うことができる。この結果、音声認識結果と正解テキストとの比較により以下の式で計算した実験環境における文字単位の平均認識精度は 0.788 であった。

$$1 - D/N$$

ここで

D: 認識結果と正解テキストの編集距離(文字単位)

N: 正解テキストの文字数

質問とその音声認識誤りの例を表 2 に示す。

表 2 音声認識誤りの例

Factoid 型	(正) 動画を撮影するときの画像サイズは何ピクセル
	(誤) 動画を撮影するときの画像サイズ阿南ピクセル
Factoid 型	(正) シャッター速度は何秒
	(誤) シャッター速度は難病
How 型	(正) 予熱の設定はどうやるの
	(誤) 余熱の設定はどうやるの

Factoid 型と How 型の質問について、質問文をキーボードで直接入力した場合と、音声により入力した場合の検索結果を、上位 3 位までの MRR[†]で比較すると表 3 のようになった。

表 3 音声認識誤りによる検索精度劣化

	(1)	(2)	(3)=(2)/(1)
Factoid 型	0.742	0.429	0.578
How 型	0.732	0.625	0.854

(1)キーボード入力、(2)音声入力、(3)検索精度劣化の割合

[†] MRR(Mean Reciprocal Rank)正解が最初に出現した順位の逆数を得点としたもので、全質問にわたって平均したもの[12]

パイロットシステムでは Factoid 型の質問に対しては上位 3 件までに回答そのものが、また、How 型の質問に対しては正解となる映像、あるいは、取扱説明書のページが上位 3 位までに含まれているかを基準に測定した。さらに、それぞれの質問の型が正しく認識され、適切な回答メディアの選択が行われていることも条件とした。例えば、Factoid 型の質問に対して、How 型と誤認識され、映像や取扱説明書のページが出力された場合、たとえその中に正解情報が含まれていても不正解とした。

上記の結果では、How 型の質問での音声入力による精度劣化が音声認識精度と比較して小さく抑えられているのに対し、Factoid 型の質問での精度劣化は認識精度より大きくなっている(表 3 の(3))。How 型の質問で呼び出される MM ナレッジ検索エンジンは、ベクトル空間モデルをベースとした文書検索を行っており、質問文中に検索タームが複数の含まれているため、音声認識誤りが起きたとしても、その影響を受けにくい。

一方、Factoid 型の質問を処理する ASKMi は、質問タイプ/回答タイプ判定、根拠文書の検索、回答候補スコアリングの三段階で処理している。質問タイプとは Factoid 型や How 型など質問の種類を指す。回答タイプ判定とは Factoid 型の質問に対してはその回答の種類(重さ、長さ等のこと)である。質問タイプと回答タイプは質問の言い回しやそこにある語を用いて正規表現で記述されたルールベースで判定する。根拠文書の検索は質問中の検索タームに基づきを行い、所定の上位検索結果の文書について、その中の回答候補を検索タームとの距離などに基づいてスコアリングし、結果を出力する。キーボード入力と音声入力のそれぞれについて各処理フェーズの精度を測定したところ表 4 の結果が得られた。

表 4 各処理フェーズの精度

	(1)	(2)	(3)
キーボード入力	1.000	1.000	0.953
音声入力	0.700	0.964	0.671

(1) 質問タイプ判定：Factoid 型の質問が正しく Factoid 型と判定された割合

(2) 回答タイプ判定：Factoid 型と正しく判定されたうち、回答タイプも正しく判定された割合

(3) 回答と文書の MRR 比：最終結果の MRR(回答単位)/根拠文書の検索結果の MRR(文書単位)

なお、根拠文書の検索結果の MRR(文書単位)では、Factoid 型質問の正解が含まれている文書を正解文書とした。

上記の結果では、(1)の質問タイプ判定の精度劣化が全体の精度劣化に繋がっていることが示されている。すなわち、質問タイプ判定に用いる情報(以後、手がかり表現と呼ぶ)が音声入力の際に失われて(=誤認識されて)いることが精度劣化の主要因となっている。(2)ではキーボード入力と音声入力の両方ともあまり精度が劣化していない。これは質問タイプ判定に成功した質問文では、手がかり表現が正しく音声認識されており、その結果、手がかり表現に基づく回答タイプも正しく判定されていることを示している。(3)では根拠文書の検索精度があまり劣化していないにもかかわらず、根拠文書から Factoid 型の回答を抽出する段階で精度が劣化していることを示している。これは、手がかり表現以外の検索タームを基にしている根拠文書の検索は、音声認識誤りの影響を受けにくいため、根拠文書から Factoid 型の回答を抽出する段階では、手がかり表現を基にするため、音声認識誤りの影響を大きく受けることがある。

4. 提案手法

前節の音声認識誤りの影響を検討した結果、音声入力では手がかり表現の誤認識が最終検索結果の精度劣化に繋がっていることがわかった。ただし、誤認識されても、音声認識という処理の性質上、類似の発話表現になっている可能性が高いという仮説の下に、以下の方針で失われた手がかり表現の復元を試みることにした。

- (1) あらかじめ手がかり表現の候補をリストアップしておく
- (2) Factoid 型に判定されなかった質問の認識結果について、(1)の手がかり表現の候補と発話上での編集距離を測定し、所定の閾値以内のものがあれば、編集距離の最短のものを質問に追加する。ただし、
- (3) (1)の手がかり表現の候補を、質問文中の語との文書中での共起関係を基に絞り込み、(2)の照合処理の効率向上、および、精度向上を図る

この処理の例を図 7 に示す。

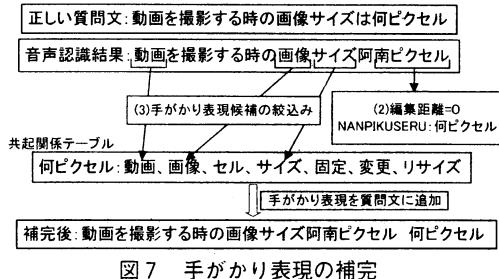


図 7 手がかり表現の補完

この例では正しい質問文の手がかり表現「何ピクセル」が、音声認識誤りによって「南ピクセル」と変化している。そのため、手がかり表現を基に判定している質問タイプ判定が、これを Factoid 型と正しく判定できない。この時、音声認識結果の質問文中の語と、リストアップされた手がかり表現の候補との編集距離を測定し、手がかり表現「何ピクセル」を編集距離が 0 で見つける。ただし、この時、リスト中の全ての手がかり表現との編集距離を求めるのではなく、共起関係テーブルにおいて、質問文中の語「動画」「画像」「サイズ」との共起関係のある手がかり表現に絞り込み、編集距離を求める。

(1)の手がかり表現の候補の例を表 5 に示す。これは、ASKMi による質問タイプ判定と回答タイプ判定用いたルール(表 1)から Factoid 型と判定される表現を抽出した。今回の実験ではこのような手がかり表現を 113 語リストアップした。

表 5 手がかり表現の例

何回	何倍	何枚	何本
何秒	何分	何時間	何週間
何年	何ミリ	何センチ	何ピクセル

(2)の発話上の編集距離はローマ字表現を用いて計算した。また、手がかり表現の候補と照合するのは質問全体ではなく、質問中で、照合中の候補と編集距離が最短の部分文字列を求めている。さらに、編集距離の値をそのまま用いると文字数の多い手がかり表現が不利になるので、以下の式で正規化した類似度を求めた。

$1 - D/N$

ここで

D:手がかり表現の候補と質問中の部分文字列の編集距離(ローマ字表記)

N:手がかり表現の候補のローマ字表記上での文字数(アルファベット単位)

(3)の絞り込は以下の手順で作成した手がかり表現の候補と検索タームの共起関係テーブルを用いて行う。

- (3)-1 Factoid 型質問の回答候補となる表現を検索対象の文書で走査する
- (3)-2 回答候補となる表現を文書中で見つけたら、その度に(3)-3 以下の処理を行う
- (3)-3 (3)-2 の回答候補を回答とする手がかり表現を推定する
- (3)-4 (3)-2 で見つけた表現の前後一定文字数だけ文字列を切り出す
- (3)-5 (3)-4 で切り出した文字列中で検索タームとなる語(例えば、形態素解析を施し、名詞を探す)があれば、その検索タームと(3)-3 の手がかり表現を共起関係テーブルに登録する。

この処理の例を図 8 に示す。

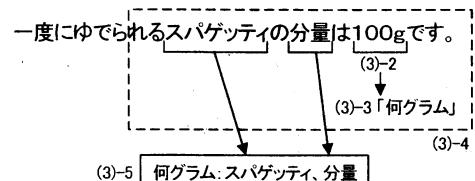


図 8 共起関係テーブル作成処理の例

この例では(3)-1 において対象の文書で走査した結果、(3)-2 の処理で「100g」という回答候補となる表現を見つけている。次に(3)-3 の処理で「100g」を回答とする「何グラム」という手がかり表現を推定している。次に(3)-4 の処理で、「100g」の前後一定文字数区間「れるスパゲッティの分量は 100g です。」を切り出す。(3)-5 の処理で、この部分文字列中で検索タームとなる語「スパゲッティ」「分量」を抽出し、手がかり表現「何グラム」と共起関係テーブルに登録する。

(3)-1 の回答候補は「XXcm」のような単位のついた数量表現について行い、また、(3)-3 での推定する手がかり表現はその単位の数量を求める疑問表現、例えば、「XXcm」であれば「何 cm」にした。上記の処理により、「何 cm」などの手がかり表現と検索タームとの対から成る共起関係テーブルが作成される。このようにして作成した共起関係テーブルの例

を表6に示す。今回の実験では113語の手がかり表現とそれぞれ最大49語の共起する単語からなる共起関係テーブルを作成した。

表6 共起関係テーブルの例

何回	連續,動き,被写体,赤目,軽減,フラッシュ,発光,強制,...
何倍	ズーム,被写体,光学,デジタル,距離,最大,拡大,画素,...
何度	リモコン,受光,左右,タン,矢印,ボタン,回転,表示,...
何枚	ソフトウェア,ケーブル,撮影,間隔,枚数,再生,モード,...
何本	アルカリ,乾電池,メモリー,カード,リチウム,電池,...
何杯	一度,分量,解凍,キー,中華,バター,ロール,調理,パン,...
何秒	撮影,間隔,枚数,ビュー,画像,表示,プレビュー,発光,...

(3)の手がかり表現の候補の絞り込みでは、この共起関係テーブルを用いて、質問中にある検索タームと対で登録されている手がかり表現があれば、それらを対象に(2)の照合処理を行うようにしておらず、リストアップされた全ての手がかり表現の候補で(2)の照合を行なうわけではない。これは、手がかり表現の候補の中には、例えば「何度(NANDO)」のような短い語もある。このように短い語で(2)の照合を行うと多くの語にマッチしてしまう。また「何度」「何℃」のように発音が同じで意味の違う手がかり表現について、その曖昧性の解消のためにも、この共起関係テーブルを用いた絞り込みは有効である。なお、「長さは?」や「どのくらい」などの特定の単位に依存しない手がかり表現は、質問中の検索タームにかかわらず、すべての質問に対して(2)の照合処理を行っている。

5. 実験

3節の音声認識誤りの影響を検討した実験データに基づき、補完前と補完後のMRR値を比較した。Factoid型質問の音声認識結果について4節の方式で補完したものに入力してMRR値を測定し、4節に示した結果と併記すると以下の表7のようになる。

表7 補完後の検索精度

	(1)	(2)	(3)
Factoid型	0.742	0.429	0.597

(1)キーボード入力、(2)音声入力、(3)補完後

これらの結果を表3と表7を合わせて図9に示す。Factoid型の検索精度が補完前では0.429であったが、補完後は0.597にまで精度を向上させることができた。

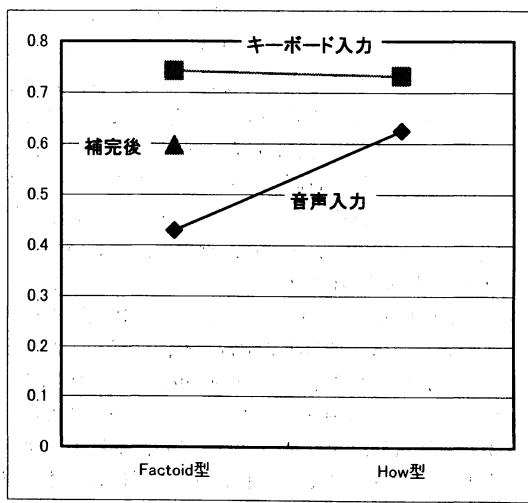


図9 実験結果

質問応答型検索の各処理フェーズの精度についても調査すると、表8の結果が得られた。

表8 各処理フェーズにおける精度

	(1)	(2)	(3)
キーボード入力	1.000	1.000	0.953
音声入力	0.700	0.964	0.671
補完後	0.923	0.959	0.926

(1)質問タイプ判定：Factoid型の質問が正しくFactoid型と判定された割合

(2)回答タイプ判定：Factoid型と正しく判定されたうち、回答タイプも正しく判定された割合

(3)回答と文書のMRR比：最終結果のMRR(回答単位)／根拠文書の検索結果のMRR(文書単位)

この結果からも音声入力の際のネックとなっていた質問タイプの判定が補完処理によって改善されていることがわかる。

6. 考察

音声認識誤りに大きな影響を受ける Factoid 型質問について、How 型の質問と同程度の検索精度まで向上させることができた。Factoid 型質問に対しては、質問タイプ／回答タイプ判定、根拠文書の検索、回答候補スコアリングの三段階で処理しており、音声認識誤りの影響を大きく受ける。特に音声認識誤りにより手がかり表現が正しく認識されない場合、質問タイプ判定と回答タイプ判定が正しく行われず、Factoid 型の質問を正しく処理することができなかつた。今回の実験では、その手がかり表現を補完することにより、Factoid 型の質問においても、How 型と同様の検索精度にすることができた。

How 型の質問では、事前に取扱説明書を使った音声認識辞書の語彙学習でも、検索タームが取扱説明書に現れやすい。一方 Factoid 型の質問では、「何グラム」や「何メートル」などの手がかり表現は取扱説明書には現れないため、音声認識辞書の語彙として学習されにくく、手がかり表現に音声認識誤りが起きる 1 つの要因となっていると考えられる。

表 7 におけるキーボード入力に対する補完後の精度劣化は、

$$(3)/(1) = 0.597 / 0.742 = 0.805$$

となり、実験環境での音声認識精度 0.788 とほぼ同じ値にまで精度劣化を抑えることに成功した。これは今後音声認識精度が向上すれば、それに伴い全体の検索精度も向上することを表している。

7. おわりに

音声入力に基づく家電ヘルプシステムにおいて、音声認識誤りに対するロバスト性向上のための質問補完処理を実装し、その効果を測定した。音声入力による精度劣化の大きかった Factoid 型質問に対して、文書中の回答候補と検索タームの共起関係に基づく補完処理を実装した。

Factoid 型質問において、質問タイプ判定の処理が音声入力の際にキーボード入力に比べて 0.700 の精度に劣化し、検索精度全体のネックになっていたが、上記の補完処理を施すことでキーボード入力に比べて 0.923 の精度まで復元することができた。

今回、より簡単な操作性を実現するために、質問応答技術に音声入力インターフェースを持たせたが、単に音声認識と情報検索を組み合わせるだけでは、それぞれの掛け算の精度に到達せず、それらを連携させるための工夫が必要なことがわかった。今

回は音声認識結果を情報検索側で補完する手法をとったが、今後、補完結果を音声認識側にフィードバックして認識精度を向上させたり、また対話処理を組み込んでユーザに確認するなど、これらの技術をより密に連携させることで全体の精度を向上させる工夫を検討したい。

8. 参考文献

- [1]浦田他、質問応答技術に基づくマルチモーダルヘルプシステム、情報処理学会研究報告、2004-FI-74, pp.23-29(2004)
- [2]Sakai, T. et al., ASKMi: A Japanese Question Answering System based on Semantic Role Analysis, RIAO 2004, pp.215-231(2004)
- [3]Fukumoto, J. et al., Question Answering Challenge(QAC-1): An Evaluation of Question Answering Tasks at the NTCIR Workshop 3, AAAI Spring Symposium: New Directions in Question Answering, pp.122-33(2003)
- [4]堀他、音声インラクティブ Q A のための応答対話文生成法、日本音響学会秋季講演論文集 Vol1, pp.81-82(2002)
- [5]西崎他、音声キーワードによるニュース音声データベース検索手法、情報処理学会論文誌、Vol.42, No.12, pp.3173-3184(2001)
- [6]秋葉他、音声入力による質問応答システムのための音声認識用言語モデルの検討、言語処理学会第 8 回年次大会発表論文集, pp.244-247(2002)
- [7]翠他、音声対話による大規模知識ベース検索システム - 音声版ダイアログナビ -、情報処理学会研究報告 2004-HI-109/2004-SLP-52, pp.21-26(2004)
- [8]MPEG7 Japan 情報規格調査会 SC29/WG11/MPEG-7 小委員会編
<http://www.itscj.ipst.or.jp/mpeg7/>
- [9]中山他、知識情報共有システム(KIDS)の開発と実践 -組織におけるノウハウ共有の促進-、人工知能学会誌、Vol.16, No.1, pp.64-68, Jan.2001
- [10]鈴木他、マルチモーダルナレッジ技術の展示案内システムへの適用、人工知能学会誌、Vol.18, No.2, Mar.2003
- [11]市村他、固有表現抽出と回答タイプ体系が質問応答システムの性能に与える影響、電子情報通信学会論文誌、D-II, 2005, [掲載予定]
- [12]LaLaVoice2001,(株)東芝
http://www3.toshiba.co.jp/pc/lalavoice/index_j.htm