

E-025

スマートホームにおける音声ニュース提供システムアーキテクチャの構築 An Architecture for the Interactive News Providing System in Smart Homes

東原 智幸† Tomoyuki Higashihara
三吉 達夫† Tatsuo Miyoshi
渥美 雅保† Masayasu Atsumi

1. はじめに

近年、ユーザの興味に合わせた情報提供を行うシステムの研究や、スマートホームのようなユビキタス環境での対話システムについての研究も行われている[1].
本研究では、スマートホーム内での音声対話ニュース提供システムのアーキテクチャについて提案する. 本システムは、ユーザとコミュニケーションや情報提供を行うインターフェース部、ユーザの関心・興味、ニュースデータを管理するユーザ・ニュース管理部で構築される. 本論では特にユーザ管理について詳細に述べる.
本システムの一部を、環境埋め込み型としてウォッシュスタンド、移動ロボット型として移動ロボットRobovie-R ver2[2]に実装し、システムの評価を行った.

2. 興味構造モデル

我々は、ユーザの興味を、ユーザが興味を示したニュースを格構造により重み付けした単語頻度ベクトルとそれに対するユーザの感性的興味の種類の関係で表現した. そして、それらの関係を自己組織化関係ネットワークにより学習しユーザの興味構造モデル(図 1)を表現した[3]. 感性的興味は 8 次元ベクトル(喜び,受容,恐れ,驚き,悲しみ,嫌悪,怒り,期待)で表現する. ニュースに対して“喜び”と“期待”の複合感情である“希望”の感性的興味を感じた場合、感性的興味ベクトル y は(1,0,0,0,0,0,1)で表現される. 本システムでは、単語頻度ベクトルを格情報ではなく、ユーザが興味を示した興味語で重み付けする.
ニュースの選択は、学習された興味構造モデルを用いて行われる. 選択では、ユーザの感性的興味に基づき感性的興味ベクトル \hat{y} を設定し、式(1)により類似度 z_j が最大になるにユニット j を決定する. 類似度 z_j は、 \hat{y} と感情重みベクトル u_j との類似度である. ユニット j の単語重みベクトル w_j をニュース選択に使用する.

$$\arg \max_j z_j = \hat{y} \cdot u_j / (\|\hat{y}\| \|u_j\|) \quad \dots (1)$$

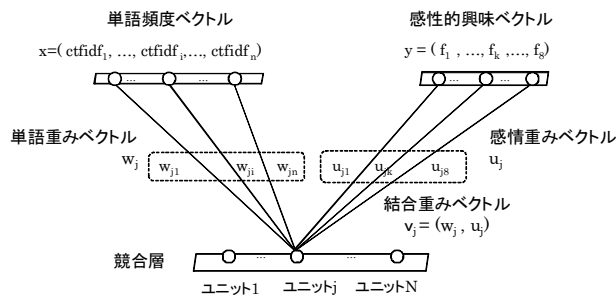


図 1 興味構造モデル

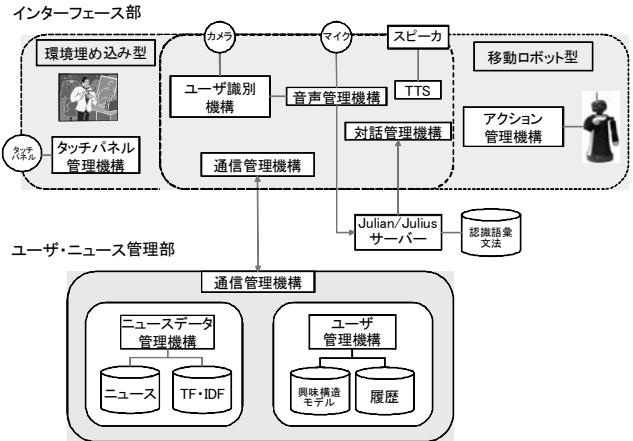


図 2 システム構成

3. システム構成

本システム(図 2)は、ユーザとコミュニケーションや情報提供を行うインターフェース部、ユーザの関心・興味やニュースデータを管理するユーザ・ニュース管理部で構築される. スマートホーム内には、1つのユーザ・ニュース管理部と複数のインターフェース部が存在し、どのインターフェースにおいてもユーザに適応したニュースの提供が行える.

3.1. インターフェース部

ユーザとの対話や情報の提供を行う部分で、環境埋め込み型と移動ロボット型に大別される.
インターフェース部共通の機能としては、ユーザ識別、音声管理、対話管理、通信管理の4つの機能がある. ユーザ識別機構は、カメラ画像や音声からユーザ識別を行う. 音声管理機構は、マイクからの音声取得、認識の管理を行う機構である. 得られた音声認識結果やユーザ情報は対話管理機構の対話で使用される. 対話管理機構は、音声によるニュースの提供、音声認識結果による対話の切り替えを行う機構である. 通信管理機構は、ユーザ・ニュース管理部との通信を管理する機構である.
対話は、ユーザ識別対話、ニュース提供対話、興味確認対話があり、ユーザ識別対話にてユーザを識別後、ニュース提供対話に遷移する. ニュース提供対話は、ユーザにどの感性的興味のニュースを提供するか確認し、通信管理機構を通してユーザ・ニュース管理部にニュース選択を要求し、ニュース集合を取得する. 得られたニュース集合中から順番にニュースを選択し、ニュースのタイトルを提供する. その後、詳しい内容を提供するかどうか確認する. 提供する場合には、興味あるニュース、しない場合には興味なしニュースと判断し、通信管理機構を通してユーザ・ニュース管理部へ送信する.
ユーザが提供を選択した場合には、音声にて内容を提供後興味確認対話に遷移する. 興味確認対話は、提供し

†創価大学大学院工学研究科情報システム工学専攻

た内容に対してどの感性的興味を感じたか確認を行ったあと、ユーザが興味を示した単語(興味語)の取得を行う。興味語の候補は、ニュース集合とともにユーザ・ニュース管理部より受信し、音声認識部[4]にあらかじめ登録する。確認した興味情報は、通信管理機構を通してユーザ・ニュース管理部へ送信し、ニュース提供対話に遷移する。ニュース集合から次のニュースを選択し、同様の処理を行う。選択するニュースがない場合には、対話を終了する。

3.1.1. 環境埋め込み型

環境埋め込み型としてウォッシュスタンドを構築した。ウォッシュスタンドは、共通の機能以外に歯磨きやドライヤを使用しているときなど音声認識が難しい場合に、タッチパネルにてシステムに対する操作を行う機能を備える。

3.1.2. 移動ロボット型

移動ロボット型の機能としては、リビングなどを移動し、ユーザの近くまで移動する機能、情報提供時に顔を見て話す、相槌を打つ、興味にあったニュースを提供する際には大きなジェスチャ、興味に合わなかったときには小さなジェスチャなど状況に応じたジェスチャを行う機能が必要である。車輪移動ロボットとして Robovie-R ver2 を採用する。

3.2. ユーザ・ニュース管理部

ユーザ・ニュース管理部は、通信管理、ユーザ管理、ニュース管理から構成される。

3.2.1. 通信管理

通信管理機構は、インターフェース部とのデータの送受信を管理する機構である。ユーザ情報、ログイン・ログアウト情報、興味情報の受信、選択したニュース集合の送信を行う。

3.2.2. ユーザ管理

ユーザの興味構造モデル、ユーザ履歴、ニュース提供履歴の管理を行う機構である。ユーザ履歴は、(ユーザ識別番号、ログイン日時、ログアウト日時)の組で表現する。インターフェース部にてユーザ識別が完了した時点からユーザのログインとし、ユーザ情報をインターフェース部より受信する。ユーザ情報を解析し、ユーザ識別番号とログイン日時を履歴に登録する。ログアウトは、ユーザがインターフェース部と対話を終了した時点とし、その日時をユーザ履歴に登録する。

ニュース提供履歴は、(ニュース識別番号、興味の有無、感性的興味、興味語₁, ..., 興味語_n, 提供日時)の組で表現する。インターフェース部のニュース提供対話、興味確認対話により得られた興味情報を受信した際に登録される。

ユーザのログアウト後、ニュース提供履歴集合を使用してユーザが興味を示したニュースごとに単語頻度ベクトル、感性的興味ベクトルを作成する。単語頻度ベクトルは、TF・IDF 値をニュース管理部より取得し作成する。ベクトル中の興味語に対応する要素は、その値を γ (≥ 1.0) 倍し重み付けする。作成した単語頻度ベクトル、感性的興味ベクトルを用いて自己組織化関係ネットワーク

の学習により興味構造モデルを更新する。

3.2.3. ニュース管理

ニュース管理機構は、ニュースの管理・検索、ニュースに含まれる単語データの抽出、TF・IDF データの管理・検索を行う機構である。単語データは、(単語表記、読み集合)の組で表現される。単語データは、TF・IDF 作成以外にインターフェース部の興味確認対話にて使用される。本システムでは、ニュースに含まれる単語(名詞、アルファベット語、カタカナ語)を単語データとして登録する。

インターフェース部からニュース選択の要求を受信したときには、ユーザ管理部からユーザのログアウト日時、興味構造モデルを取得し、ニュースの検索と選択を行う。ニュース検索は、ログアウト日時からの現在までに取得されたニュースを検索し、そのニュース集合を作成する。ニュース選択は、まずユーザから要求された感性的興味に基づいて、単語重みベクトル w_j を求める。求めた w_j とニュース集合と類似度を計算し、類似度の高いニュースから順にソートし、上位 n 個を選択する。選択したニュースに含まれる単語データを取得し、ニュースとともにインターフェース部へ送信する。

4. システム評価

本システムの一部を、ウォッシュスタンド環境埋め込み型インターフェース、Robovie-R ver2 移動ロボット型インターフェースに実装し、システムの評価を行った。

準備として goo ニュース 99 件(2005 年 12 月 2 日~13 日)をユーザに提示し、興味構造モデルを作成した。TF・IDF データおよび単語データは、goo ニュース 99 件と毎日新聞 2007 年版[5]から 1 月 1 日、3 日~5 日のニュース 548 件(2 日は 0 件)から作成した。

興味構造モデルは、goo ニュース中のユーザが興味を示したニュース 48 件を使用し、以下の条件で学習した。ユニット数は $6(2 \times 3)$ 個、学習率 0.3、学習回数 10000 回、 $\gamma=5.0$ とした。48 件の感性的興味分布を表 1 に示す。

表 2 は、ユニット間の結合重みベクトルの類似度を示す。 $j(x,y)$ はユニット番号 j と 2 次元ネットワークでのユニット位置座標を示す。遠いユニット間では、類似度が低くなっていることが確認できる。

表 1 感性的興味分布

喜び	受容	怒り	驚き	悲しみ	嫌悪	恐れ	期待
2	4	3.5	13	2	3	5	15.5

表 2 ユニット間の類似度

ユニット $j(x,y)$	1	2	3	4	5	6
1(0,0)	1	0.57	0.21	0.86	0.65	0.10
2(0,1)		1	0.77	0.26	0.52	0.52
3(0,2)			1	0.05	0.54	0.88
4(1,0)				1	0.69	0.09
5(1,1)					1	0.63
6(1,2)						1

表 3 ユニット毎の上位 10 位の単語

アニストン,牛肉,恐竜, 耐震,米国,選挙, 劇場,議員,脂肪,心臓, 1(0,0)	劇場,心臓,脂肪,充電, 記憶,解体,テレビ, アニストン,小泉,アジア, 4(1,0)
牛肉,電話,選挙,米国, アニストン,フジモリ, 西村,議員,イラク,撤退, 2(0,1)	ウイルス,牛肉,カンガルー, トロイ,木馬,エイズ, 米国,レーン,配備,空母,選挙, 5(1,1)
電話,牛肉,カンガルー, 木馬,トロイ,撤退, 選挙,カズ,イラク,米国, 3(0,2)	電話,カンガルー,トロイ, 木馬,ウイルス,レーン,空母, 配備,応答,スモーカー, 6(1,2)

表 4 感性的興味によるニュース選択

	1 位	2 位	3 位
喜び	70101016	70101037	70101052
受容	70101052	70101107	70101031
恐れ	70101029	70101016	70101027
驚き	70101052	70101107	70101031
悲しみ	70101103	70101029	70101027
嫌悪	70101103	70101029	70101027
怒り	70101103	70101027	70101029
期待	70101037	70101016	70101029

表 3中の単語は、作成した興味構造モデルの各ユニットの単語重みベクトル w_i 中から値の大きい上位 10 個の単語を示す。各ユニットで毎に分類されている単語を大まかに知ることができる。

次に、作成した興味構造モデルを使用してニュースの選択を行った。ニュースは、毎日新聞の 1 月 1 日のニュース 88 件を使用した。表 4は、感性的興味毎にニュース選択を行い、上位 3 位までのニュース識別番号を示した表である。

ユーザの感性的興味に適合しているニュースは、70101037, 70101103, 70101052 である。70101037 は新しい携帯決済サービスに関するニュースで“期待”と適合している。また、70101103 はある議員が速度違反についてのニュースで“嫌悪”または“怒り”，70101052 はある選手の競技会辞退についてのニュースで“驚き”と適合している。“喜び”で 1 位である 70101016 は、原子力委員会の人事のニュースで一致は確認できなかった。興味構造を作成する際に、使用したニュースに“喜び”に対応するニュースが少なかったため，“喜び”の興味構造がモデルに反映されなかったと考えられる。

最後に、表 4 に示されたニュース 8 件を使用し、ニュース提供対話にてユーザに提供し、興味確認対話にて興味の取得を行った。

興味語の取得では、ユーザが興味語を応答する際に手間がかかるケースがあった。70101037 のニュースで、ユーザが“携帯決済サービス”に興味がある場合、まず“携帯”，“決済”，“サービス”と単語に分割し、「“携帯”と“サービス”に興味があります」や「“決済”と“サービス”に興味があります」などのようにシステムに回答する必要がある。複合語の認識機能を導入することで、ユーザの手間を低減することが可能となる。

5. まとめ

スマートホームでのニュース提供システムアーキテクチャについて述べ、音声によるニュース提供・対話による興味語の取得、ニュースの検索・選択、興味構造モデルの更新について評価を行い動作について確認を行った。今後の課題として、

- (1) 形態素情報だけでなく、単語の位置関係、係り受け情報を利用した複合語認識機能の追加による興味語の取得
- (2) ニュースを一度に読み上げるのではなく、ユーザの状態に適した段落単位で提供[6]、提供の速度の調整、適切な読み上げへの対応
- (3) 興味モデルを更新する際に、新規のネットワークから学習するのではなく、過去のモデルにインクリメンタルに更新する手法の構築およびユニット数、学習率、学習回数の自動設定
- (4) 固定されたニュース集合ではなく、日数が経過するごとに増加するニュースへの対応がある。

本研究では、ウォッシュスタンド環境埋め込み型インターフェース、Robovie-R ver2 移動ロボット型インターフェースの個々の機能について評価を行っていない。ウォッシュスタンドでは、タッチパネルにニュース分野や興味語候補を表示することで、ニュース検索、興味の取得をタッチパネル操作で行える。また、移動ロボット型では、能動的にユーザのいる場所まで移動し、ニュースの提供を行う機能やジェスチャを交えながらのニュース提供機能の実装を目指す。また、これらの機能についてユーザとのインタラクションの評価を行う。

システム全体の機能としては、複数ユーザの同じインターフェースに対しての同時対話(話者分離)を実現する機能が必要である。

謝辞

本研究の一部は、私立大学社会連携研究推進事業「測位/光神経複合センサノードによるユビキタス・モニタリング・ネットワークの開発とその産業応用への展開」の支援のもとに実施された。

【参考文献】

- [1] 上田:ユビキタス環境における対話型ロボットインターフェースのための対話戦略の構築, 情報処理学会誌, Vol147, No. 1, 87-97, 2006
- [2] 普及版日常活動型ロボット Robovie-R ver2, <http://www.atr-robo.com/product/r2/robo-r2.html>
- [3] 東原, 三吉, 渥美:ウェブニュース提供のための自己組織化ネットワークと格重み付き単語頻度ベクトルを用いたユーザの興味構造表現, 情報処理学会第 70 回全国大会, 2D-1, p. 2-63~p. 2-64, 2008
- [4] 河原, 李:連続音声認識ソフトウェア Julius, 人工知能学会誌, vol. 20, No. 1, pp. 41-49, 2005
- [5] CD-毎日新聞(データ集)2007 年版
- [6] 三吉, 東原, 渥美:ユーザの興味プロファイルに基づく音声ニュースシステムにおける複数関連ニュース提供の対話的制御, 情報処理学会第 70 回全国大会, 2D-2, p. 2-65~p. 2-66, 2008