

近距離への呼び込みを目的とした遠距離対話システム

山本 友紀子[†]

慶應義塾大学 理工学部

梅澤 猛[‡]

慶應義塾大学大学院 理工学研究科

今井 倫太[§]

慶應義塾大学 理工学部

{yukiko,umezawa,michita}@ayu.ics.keio.ac.jp

1 はじめに

既存の人間とロボットの音声対話システムでは、遠距離で対話を行うことは想定されていない。これは、内蔵されたマイクやセンサが遠距離では有効に働かないためである。よって、対話を行いたい相手が遠距離にいる場合は、相手を近距離に呼び寄せてから対話を行う必要がある。この際、呼び込む相手の周囲には他の人間がいる可能性があるため、ロボットが誰に呼びかけたのかを周囲の人間が正しく認識することが重要である。

そこで本研究では、遠くにいる特定の人間1人だけを近くに呼び込むシステム COMECAT を提案する。COMECAT は、遠距離にいる人間を画像処理によって発見し、指向性スピーカを用いて話しかける。また、相手の服の色を認識して発話に盛り込むことで、効果的な呼びかけを図る。

2 COMECAT

2.1 コミュニケーションロボット Robovie

本研究では、コミュニケーションロボット Robovie[1] を用いる。また、図1のように、ロボットのスピーカを通常のスピーカから三菱電機エンジニアリング製超指向性スピーカに付け替えた。

以下では、ロボットの頭上に取り付けられた円柱状のカメラを全方位カメラと呼ぶ。また、ロボットの正面、目の部分に取り付けられているカメラを目カメラと呼ぶ。いずれのカメラも、得られる画像の各画素は256階調のRGB表色系である。

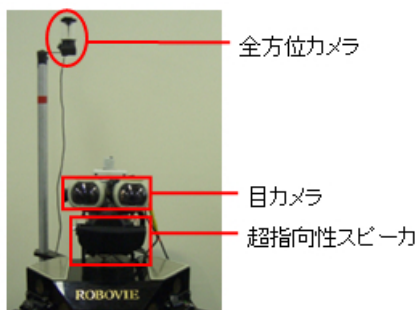


図1 コミュニケーションロボット Robovie

A Distant Calling System Using Directional Speaker

[†]Yukiko YAMAMOTO

Faculty of Science and Technology, Keio University

[‡]Takeshi UMEZAWA

Graduate School of Science and Technology, Keio University

[§]Michita IMAI

Faculty of Science and Technology, Keio University

2.2 システム構成

COMECAT の構成図を図2に示す。COMECAT は、ターンフェーズ 発話フェーズの順に実行する。図2の(1)で示したターンフェーズでは、肌色検出を行い、重心位置を計算し、動作制御フェーズに送る。動作制御フェーズでは重心位置から頭部の回転角度を算出し、アクチュエータで肌色の位置が正面となるように頭部を回転させる。図2の(2)で示した発話フェーズでは、正面の色判定を行い、判定された色を用いて超指向性スピーカにより呼びかける。

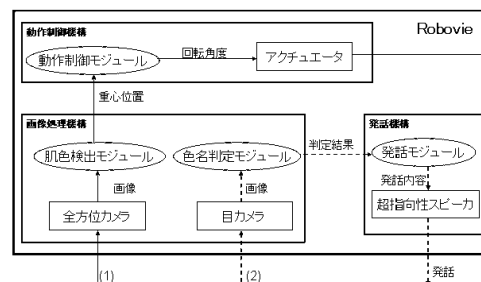


図2 システム構成図

2.3 画像処理機構

画像処理機構は、肌色検出と色名判定の2つのモジュールからなる。画像処理機構における色の判定は全てHSV表色系で行う。よって、入力画像はあらかじめRGB表色系からHSV表色系に変換しておく。

2.3.1 肌色検出による頭部回転角の算出

Robovieの全方位カメラの画像を入力データとし、背景差分による肌色検出を行う。背景はシステム起動直後に得られる100フレームの平均画像とし、

- $H < 30$
- $30 < S < 200$
- $V > 130$

であるものを肌色とする。次に全方位カメラの映像を45度ずつに区切り、各区間で肌色と検出されたピクセル数を数える。そのうち最も肌色の部分が多かった区間に人間がいると仮定し、区間中の肌色部分の重心を求め、その点にRobovieの顔が向くように回転角を計算する。

2.3.2 目カメラによる色判定

色名判定機構では、ロボット正面に見えている画像の色判定を行う。COMECATでは、以下のステップで発話する色を判定する。

1. Vの値が
 - $0 \leq V < 15$: 黒
 - $15 \leq V < 30$: 暗い色
2. Sの値が
 - $S > 15$: 白

3. Hの値が

- $0 \leq H < 5, 340 \leq H < 360$: 赤
- $5 \leq H < 50$: オレンジ
- $50 \leq H < 65$: 黄色
- $65 \leq H < 165$: 緑
- $165 \leq H < 265$: 青
- $265 \leq H < 300$: 紫
- $300 \leq H < 340$: ピンク

3 評価実験

3.1 実験概要

本実験では、ロボットの発話対象になっている場合となっていない場合、服の色の情報を加えなかった場合と加えた場合で、ロボットの呼びかけに対する被験者の反応や印象がどのように異なるかを調べる。

本実験の実験環境を図3に示す。

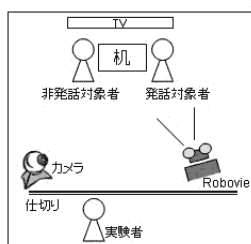


図3 実験環境

被験者には、ビデオを流した実験室に入室してもらう。実験者からは実験の開始を伝えず、立ち位置のみを伝えて自由に過ごしてもらう。被験者の入室後、1分が経過した後で2人1組の一方にRobovieが話しかけを行い、その後30秒おきに話しかける。超指向性スピーカを搭載したRobovieが、2人1組の被験者のうち、一方に話しかける。発話内容は、色情報のない呼びかけ(ねえねえ、こっちを向いて)、色情報を加えた呼びかけ(ねえねえ、その赤っぽい服のあなた、こっちを向いて)の2通りであり、被験者が振り向いたことを確認した場合に「こっちを向いて」から「こっちへ来て」へと変化させた。本実験は大学生の男女32名の被験者を、発話対象か否か、色情報の有無でそれぞれ半分、計4群8人ずつに分けた。

なお、本実験では、色判定ミスによる実験結果への影響を防ぐため、色の判定は行わず、手動で色を入力した。

3.2 実験結果と考察

実験の結果、発話内容の聞き取りについては、色情報の有無は関係なく、発話対象者では”完全に聞き取れた”、”ほとんど聞き取れた”と回答した被験者が60%程度であるのに対し、非発話対象者では20%程度にとどまった。このことから、超指向性スピーカは、発話対象者の絞り込みに有用であることがわかる。次に、Robovieの発話に対して反応(振り向き・接近)した割合を表1に示す。この結果から、Robovieの発話に反応して振り向くことは、色情報の有無とはあまり関係がないが、実際に接近させるためには、色情報を加えた方がよいことがわかる。

また、接近の理由を聞いたアンケート結果を図4、図5に示す。これらの図の選択肢のうち、「ロボットに話

表1 発話に対する反応結果

色情報の有無	発話対象者		非発話対象者	
	振り向き	接近	振り向き	接近
なし	87.5%	75.0%	87.5%	75.0%
あり	100.0%	75.0%	100.0%	50.0%

しかけられた」「ロボットがこっちに来てと言った」はロボットの呼びかけに反応したことを示す。図4、図5からは、色情報がない場合は「ロボットがこっちに来てと言った」という理由が最も多く、非発話対象者でも「ロボットに話しかけられた」と呼びかけに反応した被験者がいたのに対し、色情報を加えると、発話対象者のみが「ロボットに話しかけられた」「ロボットがこっちに来てと言った」と回答し、非発話対象者はパートナーの影響を大きく受けていることが分かる。「その他」の回答者の中には「ロボットがしつこく話しかけた」という理由で呼びかけに反応した被験者が1人いたが、その他はロボット自体に興味を持って反応していた。

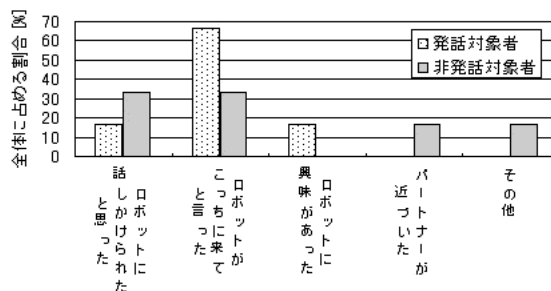


図4 接近の理由 (色情報なしの場合)

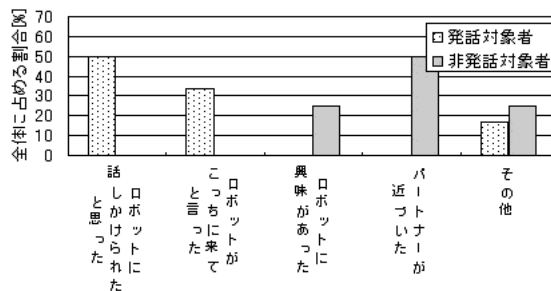


図5 接近の理由 (色情報ありの場合)

4 まとめ

本論文では、超指向性スピーカを用いて遠くにいる人間の呼び込みを行うシステム COMECAT を提案した。COMECATは、対話相手の服の色を認識し、呼びかけることができる。実験の結果から、色情報を含めることによって効果的な呼びかけができることを示した。今後、頭部回転と色判定の精度向上、対話終了後の処理について実装する予定である。

参考文献

- [1] 神田崇行, 石黒浩, 小野哲雄, 今井倫太, 前田武志, 中津良平, “研究用プラットフォームとしての日常活動型ロボット“ Robovie ”の開発”, 信学論 (D-I), vol.J85-D-I, no.4, pp.380-389, Apr. 2002.