

## 2体のロボットによる対話活性度の向上

吉岡 駿† 松元 崇裕† 岨野 太一† 後藤 充裕‡ 今井 倫太†

慶應義塾大学† NTT サービスエボリューション研究所‡

## 1 はじめに

人間と対話を行うロボットは、人の興味を引き始めの内は対話が成立する。しかし、いずれ飽きられて、それ以降、対話相手とみなしてもらえない問題を抱えている。対話が継続しない原因の一つに、対話活性度が低いことが考えられる。

人間同士の対話において、対話活性度は言語情報だけでなく、身体の動作からくる非言語情報にも起因していることが知られている。松山らは対話活性度の指標として発話量、笑顔頻度、頷き頻度を挙げている[3]。また、西村らは音声のオーバーラップを[4]、喜多は動作のオーバーラップを[1]、指標として挙げている。

上記の研究は、対話活性度が高ければ上記の非言語情報が含まれることを示している。対話活性度の判定をすることはできるのに対して、人とロボットの対話活性度を向上させることができない。

本稿では、対話活性度の高い時に見受けられる非言語情報を、ロボットが表出するように制御することで、対話活性度を高くすることを考える。具体的に、非言語情報の中でも動作に注目し、ロボット2体の動作に、真似する動作を2体のロボットが時間的にオーバーラップしながら実行することで、対話活性度を向上させることを提案する。

## 2 ミミクリ動作

人間同士の対話における時間的オーバーラップに着目する。ペントランドは、共感している人同士の対話中にみられるミミクリという現象を、どんなに隠そうとしても無意識に出てしまう非言語情報である、正直シグナルの重要な要素の一つとしている[2]。ミミクリは、対話しながら徐々に互いの動きを真似しあうことで、相手に対する好感や信頼感を強める効果がある。

本稿はミミクリを参考にして、ロボット同士の共感の有無を2体のロボットの動作で表す。動作の同期具合や同じ動作をする確率を変化させることで、違和感のないオーバーラップを実現する。

## 3 Motion Selection using Empathy for Two Robots

本稿では、2体のロボットの対話に動作の時間的オーバーラップを含むことで、人間とロボット間の対話活性度を向上させる Motion Selection using Empathy for Two Robots (MSET) を提案する。視線や指差しなどの意図的な動作と無意識的な動作を分けるため、ロボットの行うことのできる動作を指示動作と非指示動作に分ける。非指示動作について、MSET を適応する。

指示動作とは視線や指差しによって、方向を指示する動作のことである。それ以外の動作は非指示動作である。非指示動作は、あらかじめ一連の動きを定義したもので、動作群として用意されている。例えば、頷き動作、手を振る動作などがある。非指示動作は MSET によって自動で動作群から選ばれる。

非指示動作の選択には、人間同士の対話にみられるミミクリを参考にした。共感の有無  $ee$  は-1か1のいずれかの値を取り、ロボット同士の対話内容からシナリオ作成者がロボットの発話毎に決定する。また、互いの動きを真似する度合いとして、動作開始時間の差  $d$  および同一動作を選択する確率  $p$  を持つ。なお、 $d$  は割合を表し、話し手ロボットの動作の何割が過ぎたときに聞き手ロボットが動作を開始するのかを表す。例えば、 $d=0.0$  なら聞き手ロボットは話し手ロボットと同時に動作を開始し、 $d=0.5$  で、話し手ロボットの動作が4秒なら、聞き手ロボットが動き始めるのは話し手ロボットの2秒後である。 $d$  と  $p$  は話者交代が起こる毎に更新され、ある発話  $i$  においての  $d$  と  $p$  は更新値  $k_d$ ,  $k_p$  を用いて、下記の式によって求められる。すなわち、ロボット同士が共感していることを表現する場合には、2体のロボットの、動作開始時間の差を少なくし、同じ動作をする確率を高くすることで、オーバーラップを実現する。

$$d_i = d_{i-1} - k_d \cdot ee_i$$

$$p_i = p_{i-1} + k_p \cdot ee_i$$

次に、上で求めた値を基にロボットの動作を決定する。これは以下の手順で行われる。

1. 話し手ロボットの動作を動作群よりランダムに選択.
2. 聞き手ロボットの動作は確率  $p_i$  で話し手ロボットと同じ動作, 確率  $1-p_i$  で動作群より話し手ロボットの動作と異なる動作をランダムに選択.
3. 話し手ロボットが発話と同時に動作を開始.
4. 聞き手ロボットは(話し手ロボットの動作にかかる時間)  $\times d_i$  秒だけ待ち, 動作を開始.
5. 次の話者交代まで 1. から 4. を繰り返す.

図 1 に  $d_i=0.4$  の時の同一動作のオーバーラップの例を示す.

#### 4 実験

MSET を評価するために人がロボット 2 体と旅行を計画するシナリオを作成した. シナリオ中に, ロボットが指差す写真を見せるように, 実験参加者に依頼するシナリオを含む. ロボットからの依頼に実験参加者が応えるかどうかを, 実験参加者がロボットとの対話に引き込まれていることを判定する指標として用いる.

MSET の比較対象として話し手ロボットのみが動作するものとロボット同士の動きが完全に同期しているものを用意した.

##### 4.1 展示実験

Keio Techno Mall 2016 において, ロボット 2 体を設置し, 来場者の反応を観察した.

展示中, 計 3 時間の動画を撮影し, 写真をロボットに見せた人数, ロボットに対して何か行動を行った人数, 来場者がロボットに興味を示した時のロボットの動き方などを観察した. なお,  $d_i=p_i=0.5$ ,  $k_i=k_p=0.05$  と設定した.

##### 4.2 結果

記録した 3 時間の内, 60 代女性一名が, ロボットの動作が完全に同期している時に, ロボットに写真を見せた. インタビューを行ったところ, 「かわいいロボットが話していると一緒に話した



図 1. 同一動作のオーバーラップの例

くなる」と言っていた. また, ロボットに手を振る, 話しかけるなど, ロボットに対しての行動を行った来場者は上記の女性を含め 6 人だった.

どの手法の時に来場者がロボットに興味を示すかに関して大きな差はなかった.

##### 4.3 考察

写真をロボットに見せた来場者は「かわいいロボット」とコメントしていることから, ロボットの動作が完全同期していたからではなく, ロボットの容姿に惹かれて, ロボットの依頼に応えたものと考えられる.

今回の展示実験ではオーバーラップの有無による対話の盛り上がり, および引き込みの違いを検証することはできなかった. 一つの要因として, シナリオの初めから終わりまでを聞いていた来場者が少なかったことが挙げられる. 来場者はより多くの展示を見るために, 一つの展示品の前に長く留まらない. 特に, 写真を見せるようお願いする場面は一つのシナリオの最後の方に一度しかなく, その部分を聞いていない, もしくはその部分しか聞いておらず, まだ十分に引き込まれていない来場者が多かったものと考えられる.

#### 5 まとめ

本稿では人間同士の対話を参考にして, ロボット 2 体の動作をオーバーラップさせるシステム MSET について提案した. MSET により, ロボットの対話がより活発になり, ロボット同士の対話に人を引き込むことが可能になる.

今後は実験参加者が発話するシナリオの追加, アンケート調査の実施など, より仔細に観察を行い, 外的要因の少ない環境下での実験を通して, MSET の有用性を明らかにしていく.

#### 参考文献

- [1] 喜多壮太郎:ジェスチャーの認知科学 人はなぜジェスチャーをするのか, 認知科学, Vol. 7, No. 1, pp. 9-21, (2000)
- [2] ペントランド, アレックス(サンディ):正直シグナル 非言語コミュニケーションの科学(柴田裕之訳, 安西祐一郎監訳), みすず書房, (2013)
- [3] 松山洋一, 齋藤彰弘, 秋葉巖, 渡邊萌美, 小林哲則:「置いてけぼり」を救う多人数会話活性化ロボット, HAI シンポジウム, (2012)
- [4] 西村良太, 北岡教英, 中川聖一:対話における音律変化・タイミングのモデル化と音声対話システムへの適用, 人工知能学会言語・音声理解と対話処理研究会, No. 48, pp. 37-42, (2012)