

# ヒューマノイドロボットによる擬人化 subtle expression の促進

Promotion of anthropomorphic subtle expression by humanoid robot

桑波田 康太 \*<sup>1</sup> 山田 誠二 \*<sup>2\*1</sup> 小林 一樹 \*<sup>3</sup>  
 Kota KUWABATA Seiji YAMADA Kazuki KOBAYASHI

\*<sup>1</sup>東京工業大学 Tokyo Institute of Technology  
 \*<sup>2</sup>国立情報学研究所/総合研究大学院大学 National Institute of Informatics, SOKENDAI  
 \*<sup>3</sup>信州大学 Shinshu University

Subtle expression(SE) is information to be used in other than the main communication. This study treats SE as an extension of speech command is the main communication. Environment which subtle expression is more expressed is required. Speech command generally use between human and appliance which including microphone. In this case, SE is less expressed. By using a humanoid robot that can be felt more like people by anthropomorphic, it verifies that to encourage SE that can extend the speech command. We discuss the extension of the speech command by SE as application.

## 1. はじめに

Subtle expression(SE) とは人間同士のコミュニケーション時に音声対話などの主コミュニケーション以外で表出される情報のことである。主コミュニケーションは、コミュニケーション時に要となるプロトコルのことであり、音声対話によるコミュニケーションでは言語情報が主コミュニケーションとして使われる。SE は、主コミュニケーションの情報伝達を補足する役割を果たし、典型的な例として話している時に出てくる身振りや顔の表情、驚いた時に発してしまう言葉として意味を成さない発声などがある。コミュニケーションにおいて SE は、お互いに相手の印象を読み取る、または、自分の状態を相手に伝える役割を果たす。例えば、緊張している時に声が震えるなども SE である。SE の表出はユーザーにかかる負荷は低いいため、増加しても負荷にはならない。

本研究では、SE を増やす事で言語情報が主コミュニケーションであるスピーチコマンドを拡張することを考える。そのため、話す際に SE を出しやすい環境が必要になる。一般にスピーチコマンドで話しかける対象は、マイクを内蔵した人工物である。元々 SE は人同士のコミュニケーションで表出されるため、話しかける対象がただの人工物よりも人が話しかけやすい対象を用いれば、より多く有益な SE が表出されることが期待できる。人間が強く擬人化するヒューマノイドロボットを用いる事により、本来人同士で行う SE ような擬人化 SE が促進されると考える。本研究では、スピーチコマンドを使用する状況で、人間の擬人化の度合いによって擬人化 SE がどれだけ促進されるか検証する。この検証のために、マイクを内蔵しており、持ち運び可能な弱く擬人化されるノートパソコンと強く擬人化されるヒューマノイドロボットで比較実験を行い、ヒューマノイドロボットの方が擬人化 SE が多く検出されるか実験する。実験では、SE の中でも人の身振りや顔の表情などを検出する。

## 2. ヒューマノイドロボットによる擬人化 subtle expression の促進

スピーチコマンドは人間が音声によって言葉を発信する事でアプライアンスを操作できる。人間にとって自然にコマンドを発信できる方法であり、その表出にかかる負荷は少ない。スピーチコマンドを使うには、人間が対象の名前と操作を言葉で話せる事が必須である。現在は、スピーチコマンドを使用する際に付随するジェスチャーや言語でない発話などの SE の活用はまだ十分ではない。スピーチコマンドにおける SE には、ジェスチャー、顔の表情、視線、体の姿勢、心拍、発汗などがあり、ユーザーは意識的または無意識的に使用している。例えば、緊張して心拍が上がる、声が震えるなどは無意識的なもので、感情は顔の表情として出やすいが、顔の表情に出ないように演技する事は意識的である。

本研究では、図 1 のように SE を用いて、主コミュニケーションであるスピーチコマンドの拡張を考える。主コミュニケーションと SE を使ってアプライアンスの操作を考えるにあたり、ベクトルの考え方をを用いる。ベクトルの方向は、暖房の温度や音量などを上げるか下げるかどちらかの操作の方向であり、これをスピーチコマンドで決める。また、スカラーは操作方向にどれだけ変化させるかを表し、SE によって決定する。ただし、SE または非言語情報によってユーザーが求めている具体的なスカラーの量はジェスチャーや顔の表情が表す意味は

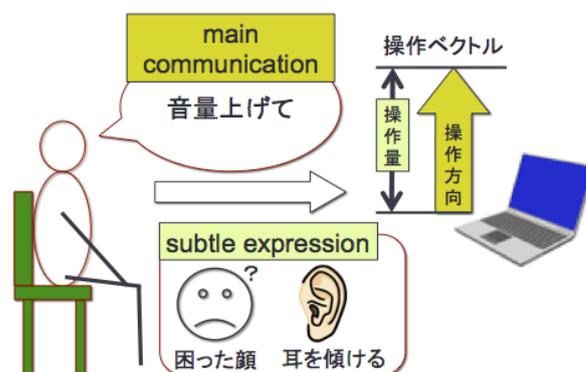


図 1 ユーザーの発信する情報と操作ベクトルの対応

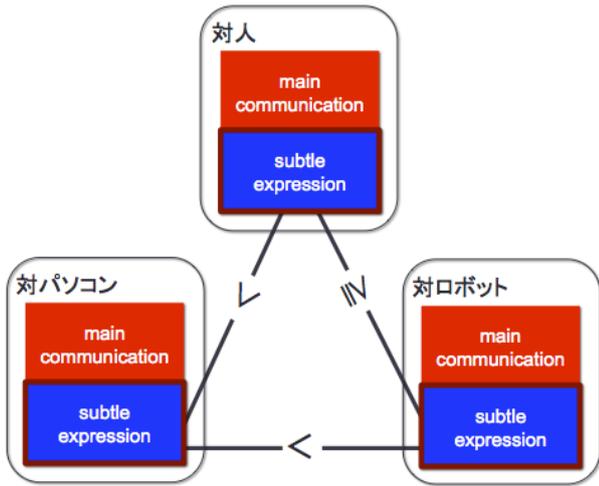


図 2 人間の非言語情報を出す度合い

異なるため一意に決めることは出来ない．今回はスピーチコマンドを行った際の SE の有無でスカラーを決める．SE がある時はスカラーが大きく，SE が無い時はスカラーが小さいといった差異をつける．

SE を利用するには，スピーチコマンドに付随する SE が多く表出されなければならないが，スピーチコマンドでは一般的には人間同士ではなく人間とマイクを内蔵したデバイスであり，SE は人間同士の時よりも少なくなると考える．そこで，話しかける対象が擬人化できるものを用いる事で，擬人化 SE の表出の促進されると考える．特に，ヒューマノイドロボットのような人間によって強く擬人化されるものは，SE が出やすくなると期待できる．人間が行う擬人化は，人間がものと同様の時の不安を軽減する事が知られている [1]．このような背景から，本研究では図 2 のような人間が話しかける対象によって変わる SE の表出について仮説を立てた．SE は人間同士で行う時に一番表出しやすい．パソコンに対して行う擬人化 SE は，人間が弱い擬人化をするため表出が少なくなると考えられる．ヒューマノイドロボットに対する擬人化 SE は人がヒューマノイドロボットを強く擬人化するため，PC よりも多くなり人間の時と同じか少ない表出があると仮定する．本研究では，擬人化 SE の表出は人間が行う擬人化の強さによって変わると仮定し，擬人化の強いヒューマノイドロボットと擬人化の弱いノートパソコンに対し，ユーザーにスピーチコマンドで話しかけてもらい，その時に表出する擬人化 SE を検出し，人間は強く擬人化されるヒューマノイドロボットに対し表出する擬人化 SE は多くなるか検証する．

### 3. 実験・評価方法

図 3 ようにユーザーは話しかける対象が置いてある机の前に座ってもらい，アプライアンスを操作してもらう．操作する時は，ノートパソコンもしくはヒューマノイドロボットに話しかけてもらう．今回の実験では，ヒューマノイドロボットは NAO[2] を使う．ユーザーのスピーチコマンドは机にセットされたマイクで聞き取り，システムからユーザーへのフィードバックは，ユーザーから要求されたアプライアンスの動作を返す．この時，NAO とノートパソコンが動く事はなく，あくまでユーザーに代替してアプライアンスを行うエージェントである．ヒューマノイドロボットなどがジェスチャーを使ってユー



図 3 実験の様子

ザーに反応したり，情報を伝える事は人がタスクを行う時の負荷を軽減する事が分かっており [3]，ヒューマノイドロボットにジェスチャーを追加すると，人間の中で強く擬人化されることが予想される．

本研究では，まず話かける相手としてノートパソコンとヒューマノイドロボットが出てきた時のユーザーの反応を検証する．そして，動かない NAO に対する擬人化 SE が多く表出されるかどうかを確認した後に，擬人化 SE の表出を促すために有効な振る舞いを追加していくことを考えている．スピーチコマンドに付随する擬人化 SE に注目するため，コマンドの内容に直接繋がるようなものを検出しなければならない [4]．検出を想定している擬人化 SE は，冷房の温度を下げるスピーチコマンドに伴って手で扇ぐような動作や音量を上げる際に音が聞こえず，耳を傾ける仕草などである．このような動作の検出には，Microsoft 社の kinect を用いる．

NAO とノートパソコンでどちらが擬人化 SE が多いか比較するために，定量的な評価方法としてジェスチャーの軌跡の総和を用いる．仮説通りであれば，NAO 条件のジェスチャー軌跡がノートパソコン条件より大きくなるはずである．

機械に SE を抽出できるようにするため，ユーザーから検出したジェスチャーから意味のある軌跡を取り出すセグメンテーションを行う必要がある．例えば，暑いから手で扇ぐというジェスチャーであれば，顔に手を近づける段階とその後に顔の近くで手を往復させて実際に扇ぐ動作をする段階でセグメンテーションされる．軌跡は図 4 とすると，始点から中点の近くまでが顔に手を近づける動作であり，中点から終点までが扇ぐ動作になる．よって，扇ぐ動作は中点から終点までの軌跡となる．このようにセグメンテーションを行い，SE のセグメンテーションされたデータを集め，SE を認識するシステムを構築することを考察する．

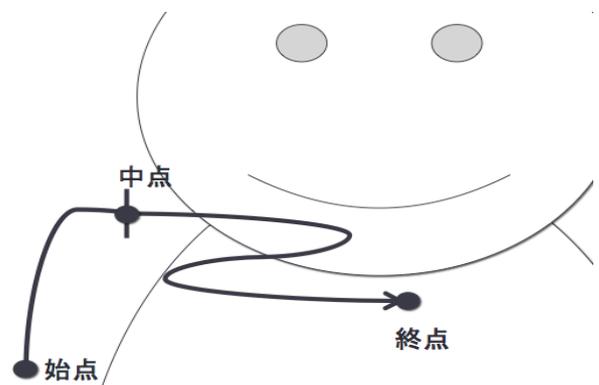


図 4 手で扇ぐジェスチャーの軌跡

## 4. まとめ

本研究では、主コミュニケーションであるスピーチコマンドに付随して表出されるSEを用いて、スピーチコマンドを拡張する事を方針とする。設計要件として、話しかける時にSEが出やすい状況でなければならない。スピーチコマンドで話しかける対象に注目する。人同士であればSEは出ることが分かっているので、人間に近いと認識されるもの、または、擬人化されるものを用いる。これによって、スピーチコマンド中にSEが増えるのか実験により検証する。擬人化の度合いによって、非言語情報の表出の度合いも変化すると予想する。強く擬人化されるヒューマノイドロボットのNAOと弱く擬人化されるノートパソコンで擬人化による擬人化SEの促進が起るのか検証する。

## 参考文献

- [1] Dieta Kuchenbrandt, Friederike Eyssel, Simon Bobinger, Maria Neufeld: Minimal Group - Maximal Effect? Evaluation and Anthropomorphization of the Humanoid Robot NAO, Third International Conference, ICSR ,pp 104-113(2011)
- [2] <https://www.aldebaran.com/ja/xiao-xing-robotutonaotoha>
- [3] Manja Lohse, Reinier Rothuis, Jorge Gallego-Pérez, Daphne E. Karreman, Vanessa Evers: Robot Gestures Make Difficult Tasks Easier: the Impact of Gestures on Perceived Workload and Task Performance, CHI14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp1459-1466 (2014)
- [4] Gang Pan, Jiahui Wu , Daqing Zhang, Zhaohui Wu, Yingchun Yang , Shijian Li: GeeAir: a universal multi-modal remote control device for home appliances, Personal and Ubiquitous Computing Volume 14, Issue 8 , pp723-735(2010)