ロボット食事介助における声がけの応答検出

松田 輝[1,a)]　三枝 亮[1,b)]

概要：食事介助における声掛けは, 食事者に次の動作を予測させるだけでなく, 反応を見ることで食事を摂る準備ができているかを確認する役割がある. 前研究では, ロボットアームの接近動作に先行して動作タイミングを表す音響を提示し, 接近動作に対する心理的抵抗と馴化への影響を検証した. 音響提示によりアームの動作予測が可能になることが示唆されたことから, 音響は介助者の意識をアームへ向かわせることが可能であると考えられる. 本研究では, 食事前から口に入れる直前までの声がけに対する食事者の反応をシステムで検出することにより, 食事者の意識が食事に向いているのか判断する手法を提案する.

キーワード：食事介助, ロボット, 音響

Response Detection of Voice Calls in Robotic Meal Assistance

Abstract: In meal assistance, voice calls not only help the person anticipate the next action, but also serve to confirm whether the person is ready to eat by observing his or her reaction. In the previous study, we presented an acoustic sound that indicates the timing of the approaching motion of the robotic arm prior to the approaching motion, and examined the psychological resistance to the approaching motion and its effect on habituation. Since it was suggested that the acoustic presentation enabled prediction of the arm's movement, it is thought that the sound can direct the caregiver's attention toward the arm. In this study, we propose a method to determine whether the caregiver's awareness is directed toward the meal by detecting the caregiver's response to the call in the interval from before the meal to just before the meal is placed in the mouth using the system.

Keywords: feeding, Robot, Sound

1. はじめに

近年, 障害者の自立支援において介助ロボットへの期待が高まっている. 脳や頸椎の損傷による重度な麻痺を有する患者は単独で食事を行うことが難しく，介護従事者や家族による介助が不可欠である. また, 食事介助時には食事者とコミュニケーションをとることが重要である. 食事介助における声掛けは, 食事者に次の動作を予測させるだけでなく, 反応を見ることで食事を摂る準備ができているかを確認する役割がある. これまでの研究では，利用者がジョイスティックでアームを操作して食物を口元へ運ぶ方法 [1] や, ロボットが自動で食物と利用者を認識して専用フォークで食物を口元に運ぶ方法 [2] が提案されている. また, 前研究 [3] では, ロボットアームの接近動作に先行して動作タイミングを表す音響を提示し, 接近動作に対する心理的抵抗と馴化への影響を, 身体動作を計測することで検証した. このとき, 音響の提示により動作を予測できることで身体動作を低減する可能性があることから, 音響は介助者の意識をアームへ向かわせることが可能であると考えられる. 本研究では, 食事前から口に入れる直前までの声がけに対する食事者の反応をシステムで検出することにより, 食事者の意識が食事に向いているのか判断する手法を提案する. また, 本研究で提案するシステムは将来的に食事介助ロボットに載せることを想定している. 本手法でロボットが物体への注意を検出することで, 食事中に話しかける際の話題づくりのための人行動認識に利用可能なほか, 食事へ向かう姿勢が整っているか検出することにより, 無理なタイミングで食物を食べさせないように判定できると考えられる.

[1] 神奈川工科大学
Shimoogino 1030, Atsugi, Kanagawa 243-0203, Japan
[a)] akira.matsuda@syblab.org
[b)] ryo.saegusa@syblab.org

© 1959 Information Processing Society of Japan

2. 方法

2.1 注意対象の計測システム

本計測システムでは, 顔の姿勢から注意方向を推定し, 食事者の注意対象を検出する. また, 合成音声を出力することで, 注意対象の変化を計測する. 注視点を検出する方法として, グラス型の視線センサや, 設置型の視線センサがある. しかし, 本システムは将来的に食事介助ロボットに搭載することを想定しており, グラス型の場合は食事者に装着してもらう手間がある. また, 設置型の視線センサでは, 頭の位置を固定しなければならない制約がある. そこで本システムでは, 装着の手間が無く, 頭部の姿勢を固定させない注意検出方法として, 顔の方向を利用したシステムを提案する. 本システムの構成を図 1 に示す. また計測のフローについて図 2 に示す. また, 計測装置は図 3 であり, 模式図を図 4 に示す。計測には深度カメラ (astra stereo) を用い, 計測装置上部と中央に 2 台のカメラが設定されている. 顔の姿勢推定では図 5 のように目, 眉, 鼻, 口, 顎, について, 12 か所を特徴点とし, 特徴点の抽出は MediaPipe[6] の顔ランドマーク検出機能を用いる. まず, 顔を正面カメラで撮影し, 各特徴点に対し, 鼻の頂点の 3 次元座標を原点とする顔モデルを作成する. OpenCV を用い DLT 法で計算することにより, 顔の姿勢ベクトルを推定する. 姿勢ベクトルを顔座標系の回転行列に変換し, 鼻の頂点から顔の正面方向に向かって延びる直線から 1 点の 3 次元座標を取得し注意座標とする. 顔の座標変換は以下の計算を行う. 推定された顔の姿勢を表す回転行列 M_f は Z 軸の負方向が顔の方向であるため, Z 軸の向きを逆にする行列 M_{flip} が必要である. 注意座標の鼻からの距離を $z$mm としたとき、顔座標系における位置ベクトルは V_{face} となる。M_f の位置ベクトル (鼻のカメラ座標) を V_f とすると, カメラ座標における注視座標 a_f の式は以下のようになる.

$$M_{flip} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$V_{face} = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$$

$$a_f = M_{flip} M_f V_{face} + V_f$$

図 6〜図 9 は推定時の様子である. 鼻から 400mm 先に注意点を設定しており, 画像中心が X, Y 軸の原点とする 3 次元座標系で, 注意点が X: 正, Y: 正のとき赤色で表示されており, X: 正, Y: 負で緑色, X: 負 Y: 正で青色, X: 負, Y: 負で白色で表示されている. また, 人差し指の先端の座標, 口の中心の座標も取得する. 次に正面カメラから得た座標情報を上カメラの座標系に一致させ, 上カメラ座標を本システムの基準座標系とする. このとき上カメラと正面カメラの位置関係は手動で設定した. 上カメラからみた正面カメラの回転行列 M_t と位置ベクトル (正面カメラの位置)V_t, 正面カメラにおける各座標 p_f から, 上カメラにおける各座標 p_t を求める式は以下のようになる.

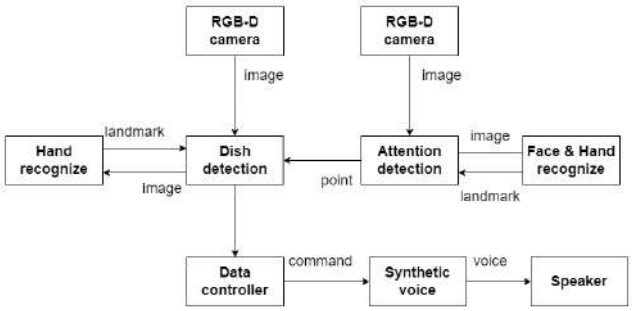


図 1 システム構成.
Fig. 1 System configuration.

$$p_t = M_f p_f + V_f$$

注意対象の判定を説明する. まず対象物の位置を記録し, 対象ごとに円形の判定領域を作成する. 次に, 鼻と各対象物の距離 x をそれぞれ計算し, 注意座標を作成する. これは, 対象物が n 個存在すれば, n 個の注意座標が作られ, 対象物と注意座標でそれぞれ対応している. 各対象物と対応する注意座標の距離を比較し, 最も距離が短い注意座標を判定に使用する. 判定に使用する注意座標が判定領域の中にあればその対象物を見ていると判定する. 上カメラからは指の座標も取得する. 正面カメラからも指の座標を取得していたが, これは, 1 つのカメラでは画角をカバーしきれないためである. 正面カメラから取得した指座標は口の座標と合わせて, 口に物を入れたかを判定するのに用いる. 上カメラの指座標はどの皿に手を伸ばしているか検出するのに用いる. 声がけには合成音声を用い, テキスト読み上げソフトウェアの VOICEVOX で音声を作成する.

2.2 システムの精度検証

注意対象の推定精度を検証する実験を行った. 実験風景を図 10 に示す. 被験者は椅子に座り, 机に乗ったマーカの方に顔を向ける. その時の注意座標とマーカの深度カメラによる座標の誤差を計測した. 被験者には注意対象を見る際, なるべく眼よりも顔を動かすことを指定し, 20 秒間注視したときの注意座標を計測した. マーカと顔の距離 400mm から 1m, マーカ, 顔, カメラの角度 0 度の条件と, マーカと顔の距離 400mm, マーカ, 顔, カメラの角度 0 度から 75 度の条件で実験した. 距離を変化させる実験では図 11 のようにカメラの位置, 椅子, 机は一直線になるように配置した. 角度を変化させる条件では, 図 12 のような配置で実験した. 実験結果を以下に示す. 図 13 は角度 0 度で距離のみ変化させたときの誤差である. 距離が離れるにつ

© 1959 Information Processing Society of Japan

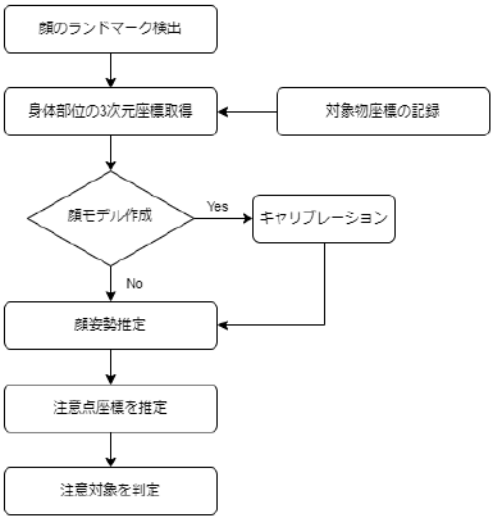


図 2 システムフロー.
Fig. 2 System Flow.

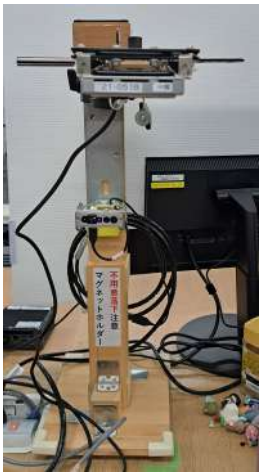


図 3 計測装置.
Fig. 3 Measuring device.

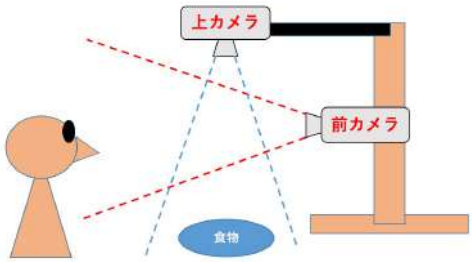


図 4 計測装置の模式図.
Fig. 4 Schematic diagram of the measurement device.

れて推定結果のばらつきが大きくなっていることがわかる. これは, 距離が長いほど微小な角度変化で発生する注意点の移動が大きくなるためである. また, 400mm のときばらつきが最小になっているものの, 誤差は大きくなっているが, これについては後述する. 図 14 は距離 400mm で角度のみ変えたとき誤差である. この図では, 角度によるばらつきの変化は少ないが, 角度が大きくなるほど誤差が大きくなっていることがわかる. 先ほどの 400mm, 0 度のデータで誤差が大きくなっていたのも, マーカと顔の距離が近かったために, ピッチ方向の角度が大きくなり, 誤差が大きくなったと考えられる. 図 15 は距離による誤差が X, Y, Z 軸のどの方向に発生したかを示す図である. この図からは, Y 軸方向の誤差が特に大きいことがわかる. 図 16 は角度による誤差が X, Y, Z 軸のどの方向に発生したかを示す図である. こちらは, Y, Z 軸方向が同程度に誤差が大きいが, X 軸方向は誤差が少ないことがわかる. このことから, X 軸方向の注意の推定は精度が高いと考えられる.

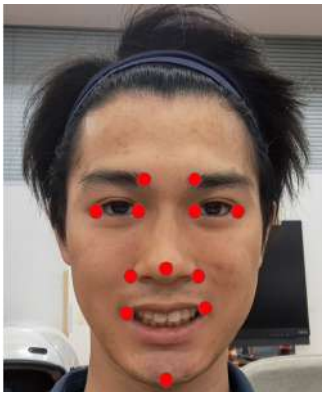


図 5 顔計測箇所.
Fig. 5 Facial measurement.



図 6 X=正,Y=正.
Fig. 6 X=positive,Y=positive



図 7 X=正,Y=負.
Fig. 7 X=positive, Y=negative.



図 8 X=負,Y=正.
Fig. 8 X=negative,Y=positive

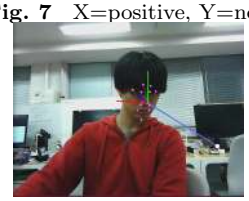


図 9 X=負,Y=負.
Fig. 9 X=negative,Y=negative.

3. 実験

3.1 概要

本実験では, システムを用いて食事の様子を計測し, 声がけに対する食事者の反応をシステムが検出できるか検証する. 実験風景を図 17 に示す. 被験者はシステムの前に座り, 食物を順番に手で取って食べる. 机上の配置は図 18 のように計測装置, 皿に乗った菓子とスピーカーが配置されており, 皿とスピーカーは計測装置の認識範囲内に設置した. 菓子はチョコレートを用意し, 皿ごとに種類が異な

© 1959 Information Processing Society of Japan



図 10 実験風景.
Fig. 10 Experimental Scene.

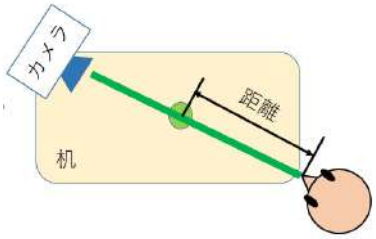


図 11 距離による精度検証.
Fig. 11 Accuracy verification by distance.

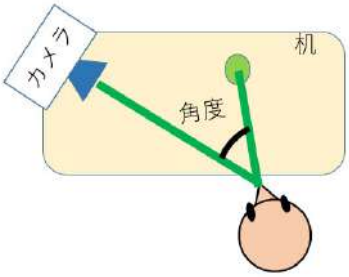


図 12 角度による精度検証.
Fig. 12 Accuracy verification by angle.

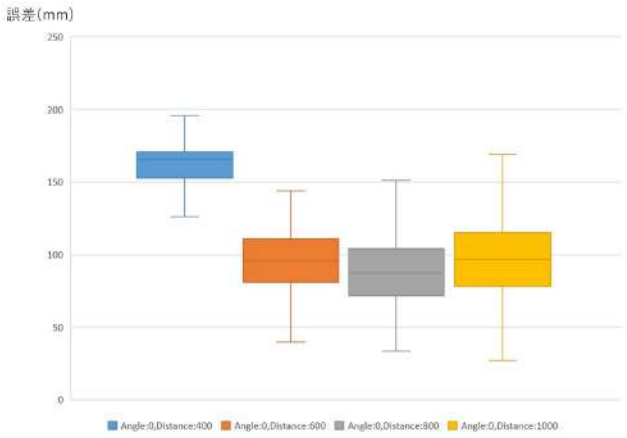


図 13 距離による誤差.
Fig. 13 Error due to distance.

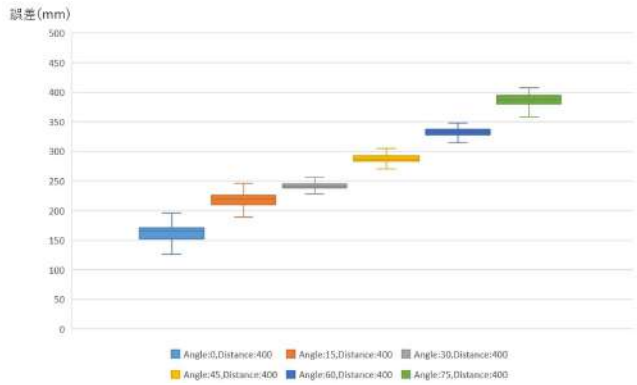


図 14 角度による誤差.
Fig. 14 Error due to angle.

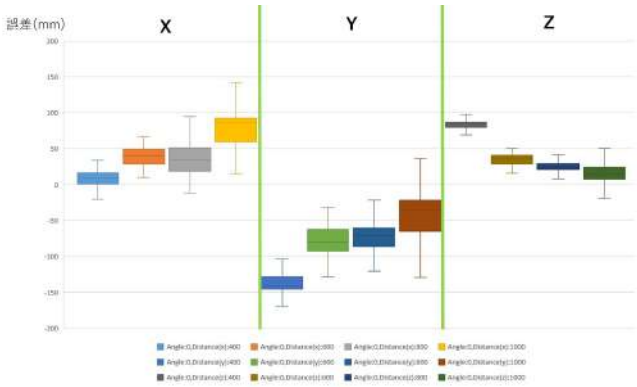


図 15 軸方向ごとの距離による誤差.
Fig. 15 Error due to distance per axis direction.

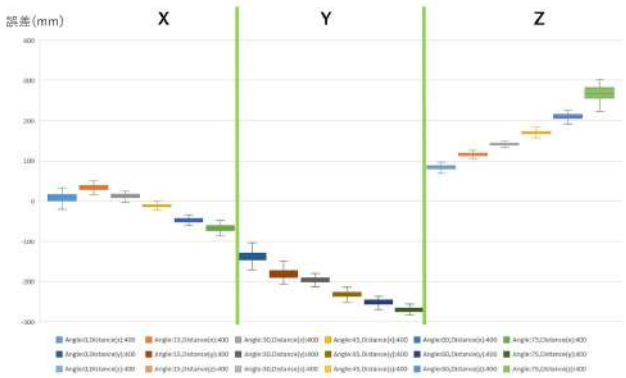


図 16 軸方向ごとの角度による誤差.
Fig. 16 Error due to angle per axis direction.

る. 食べる順番は左, 奥, 右, の順番である. 最初に音声でメニューの説明をし, その後, 食べる前に一言ずつ音声を提供する. 食べているものが口から完全に無くなったとき, 被験者が「食べ終わりました」と発声することで, 次の食物に移る. 音声の内容は図 19 である. 被験者には事前に, 片手のみを使って食事をすること, 使わない手は机の下へ下げること, 眼だけを動かして物を見ないことを指定した. また, 実験中に被験者がどこを向くかは指定していない.

© 1959 Information Processing Society of Japan

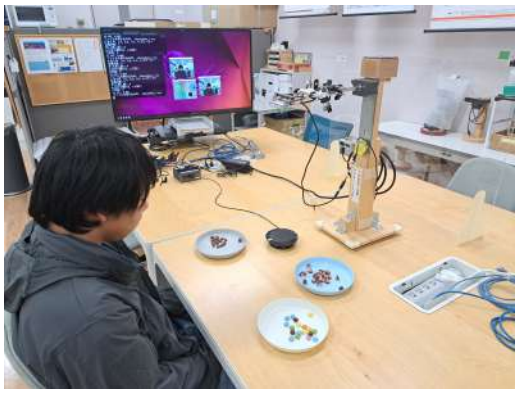


図 17 実験風景.
Fig. 17 Experimental Scene.

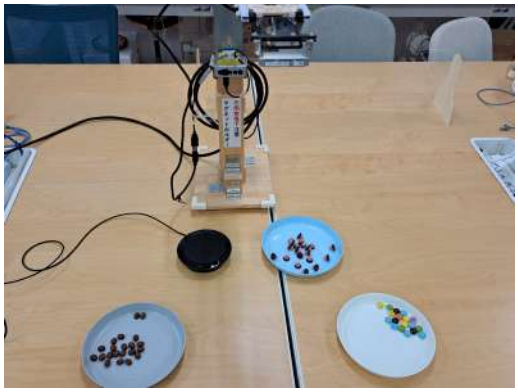


図 18 机上の配置.
Fig. 18 Desk Arrangement.

タイミング	内容
メニュー説明	おやつの時間です。これからメニューを説明します。左から順に、コーヒービート、アポロチョコレート、マーブルチョコレートです。
1つ目の皿	コーヒービートはおいしいですよ。
2つ目の皿	アポロチョコレートはどうですか。
3つ目の皿	マーブルチョコレートがまだ残っていますよ。
終了時	美味しかったですか？

図 19 音声の内容.
Fig. 19 Audio Contents.

3.2 結果

計測結果を図 20〜27 に示す．小さい円形の点は注意座標であり, 対象物の判定結果に応じた色でプロットされている．大きい四角形の点は, 皿及びスピーカーの座標である．また, 赤い線は指先の座標である．計測結果の図は判定結果を見やすくするため 2 次元で描かれているが, 実際に取得している座標は 3 次元の座標である．メニュー説明のセクション (図 20) ではスピーカーの近くに注意が多く集まっていることがわかる．やや注意がばらけているのは, 説明された食事の皿を見ようとしたためだと思われる．図 21 は一つ目の皿について声がけをしているときのデータであるが, スピーカーに向かって注意が集中していることがわかる．図 22 では, 手を食物へ伸ばすときにときその場所に

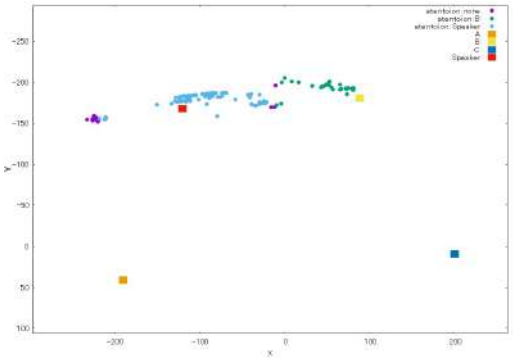


図 20 メニュー説明.
Fig. 20 Menu Description

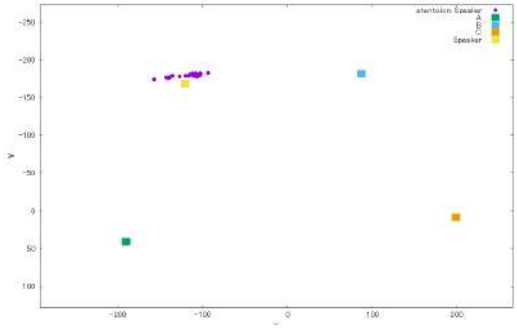


図 21 1 つ目の皿：声掛け中.
Fig. 21 The first plate: in the process of calling out to the client.

注意が集まっている様子が見て取れる．図 23,24 では 2 つ目の皿についての声がけをしているが, スピーカーと皿の間に注意が集まっている．これは対象が遠くにあることと, 対象物とスピーカー間の距離が近いことにより, 注意が分散されたと考えられる．図 25,26 では声を聴いているときは正面を向いていたが, 3 つ目の皿に手を伸ばしたときには皿に注意が集まった．また, 終了時 (図 27) の注意は正面を向いている．メニューの説明時はスピーカーに向いていたが, 音声を聴くことを繰り返すうちに食事に集中し始め, スピーカーに注意を払わなくなったのではないかと考えられる．これらより, 離れた場所にある対象物や, 対象物間の距離が近いときには判別が難しくなるが, 近くにある対象に対しては注意を検出し判定できていることが見て取れる．また, 本実験では, 注視対象の指定はしていない．よって, 声がけのタイミングと食物を手に取るときの注意が大きく変化していることを検出できたことにより, 本システムでは声がけによる反応が検出できていると考えられる．

4. まとめ

本研究では, 口に食物が入っていない区間の声がけに対する応答を検出する手法を提案し, 顔方向による注意の検出システムを提案した．声掛けから食物を食べようとするまでの注意の変化を検出できたことから, 本システムで声がけによって生じた注意の応答を検出可能であると言える．

© 1959 Information Processing Society of Japan

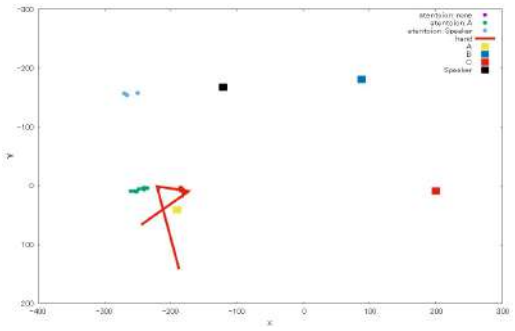


図 22 1 つ目の皿：声掛け後.
Fig. 22 First plate: after calling out.

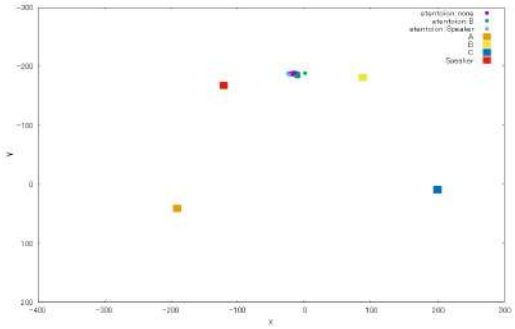


図 23 2 つ目の皿：声掛け中.
Fig. 23 Left shoulder measurement results.

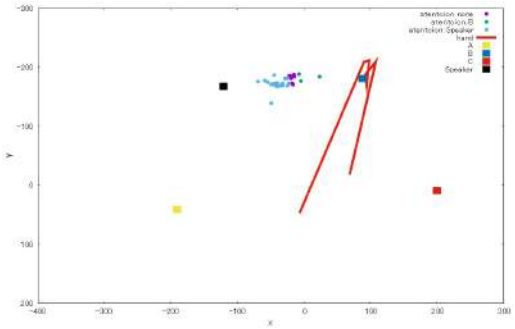


図 24 2 つ目の皿：声掛け後.
Fig. 24 Right shoulder measurement results.

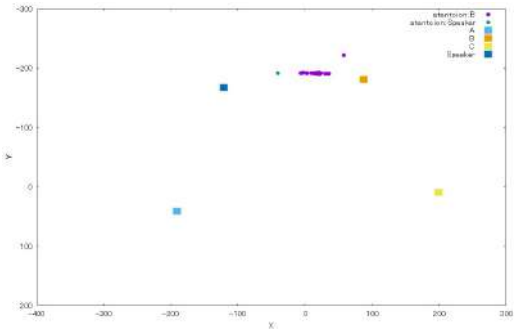


図 25 3 つ目の皿：声掛け中.
Fig. 25 Left elbow measurement results.

注意対象のおおまかな判定ができたことから, ロボットに搭載する際, 食事者が食事から十分離れた対象に注目している場合において, 食事への意識が向いておらず, 食べる準備ができていないと判定できると思われる. また, 対話コミュニケーション機能における話題提供のための, 人の行動認識手法として利用可能だと考えられる. 本システムでは 3 次元座標情報を使っているが, 注意判定には 2 次元座標で判定しているため, 縦に物が並んでいる場合や, 上にある対象への注意判定ができない. 今後は立体的に注意判定ができることにより, 机の上だけでなく周囲の環境への注意も行えるようにしたい. また, ロボットに本システムを載せた場合に実用可能か検証していきたい.

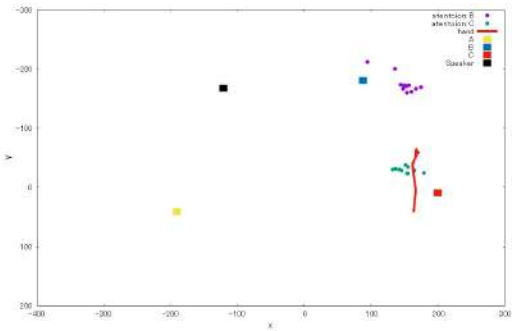


図 26 3 つ目の皿：声掛け後.
Fig. 26 Right elbow measurement results.

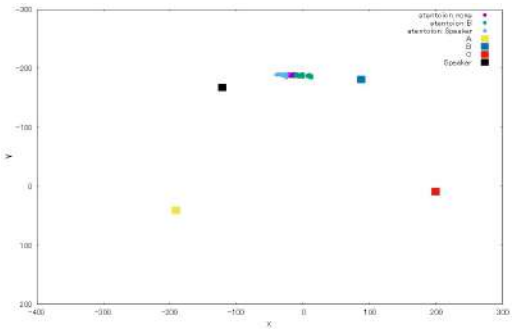


図 27 終了時.
Fig. 27 Left hand measurement results.

謝辞 本研究にご協力いただきました小田原福祉会の時田佳代子理事長と関係各位に御礼を申し上げます.

参考文献

[1] 石井 純夫, 食事支援ロボット「マイスプーン」, 日本ロボット学会誌, 21(4), pp.378-381, 2003.
[2] Feng, Ryan, et al., Robot-assisted feeding: Generalizing skewering strategies across food items on a plate, Inter. Symp. of Rob. Res., Springer, Cham, 2019.
[3] 松田 輝, 三枝 亮, ロボット食事介助における動作タイミング提示と心理的抵抗への効果, 2023-AAC-22(2), pp.1-6, 7 月 21 日, 2023
[4] 小山 秀紀, 可部 明克, 対話型ロボットの動作と音声が感情状態に与える影響：予備調査, 人間工学, 2020, 56 巻, 4 号, p. 123-129
[5] 松田 輝, 村田 圭 , 三枝 亮, 身体拡張認識によるロボットアーム制御と食事介助, 2023 年情報処理学会第 85 回全国大会講演論文集, 調布, 3 月 2 日-4 日, 2023
[6] MediaPipe https://github.com/google/mediapipe 2023 年

© 1959 Information Processing Society of Japan