社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

多人数インタラクション評価のための姿勢推定

小松　和朗† 　嶋田　和孝† 　遠藤　　勉†

† 九州工業大学情報工学部，福岡県飯塚市川津 680-4
E-mail: †{k_komatsu,shimada,endo}@pluto.ai.kyutech.ac.jp

あらまし　多人数インタラクションの円滑化において対話者の内面を推定することは重要な課題である．対話者の内面推定を実現するには言葉などの言語情報だけでなく，話者の表情や姿勢，ジェスチャといった非言語情報も重要な役割を果たす．我々は非言語情報の 1 つである姿勢に着目し，話者の姿勢から内面を推定する手法を検討している．本論文では，その前処理として深度センサーを用いた姿勢推定の手法を提案する．提案手法では頭上方向から撮影した人物画像より得られた深度情報を基に抽出した (1) 深度ヒストグラム，(2) 人物の面積，(3) 人物の体型，(4) 人物領域の中心座標，(5) 頭部領域の中心座標，(6) 身体の傾き度合い，(7) 身体の向き の計 7 種類の特徴量を用いる．抽出した特徴量を学習アルゴリズムの 1 つである Adaboost に適用することで姿勢推定に最適な分類器を作成する．実験の結果，提案手法が姿勢推定において有効であることを示す．

キーワード　非言語情報，姿勢推定，頭上カメラ，深度情報，深度センサー

Posture identification for evaluation of multi-party interaction

Kazuaki KOMATSU†, Kazutaka SHIMADA†, and Tsutomu ENDO†

† Kyushu Institute of Technology，680-4 Kawazu，Iizuka-shi，Fukuoka，820-8502，JAPAN
E-mail: †{k_komatsu,shimada,endo}@pluto.ai.kyutech.ac.jp

Abstract　The goal of our study is to understand a user's state in multi-party conversation. To realize the task, we need to estimate the state of user's mind. There are many approaches to estimate the state: use of verbal information and non-varbal information such as posture and gesture. In this paper, we propose a method for posture estimation. Characteristics of our method are (1) use of a top-view image and (2) use of depth information. Using top-view images solves a problem of occlusion. Using depth information reduces mistakes for detection of a person area and is effective for extracting features. We apply seven features, such as depth histogram and body information to the Adaboost algorithm. Experimental results show the effectiveness of our method.

Key words　Non-verbal information，Posture estimation，Top-view camera，Depth information，Depth sencor

1. ま え が き

対話は，私たち人間が生活する上で，他人とコミュニケーションを図るための手段である．コミュニケーションは，自らの意思を相手に伝え，相手の意思を受け入れることで成立し，他人との情報共有，双方の感情理解および全体の意思決定を目的として行われる．雑談や会議といった複数人で行うコミュニケーションの形態は多人数インタラクションと呼ばれ，現在対話分析を中心に研究が進められている [1].

円滑な多人数インタラクションを実現するには，各話者が自らの意図を的確に伝えるとともに相手の意図を正確に理解する必要がある．各話者は自他の心理状態，話題に対する前提知識，現時点での話題の論点といった様々な状況を考慮した上で適切な行動をとらなければならず，対話者全員が同様の行動をとるのは対話人数に比例して困難になると考えられる．加えて，対話では言葉のやりとりだけでなく，アイコンタクトやボディランゲージなどの非言語コミュニケーションも同時に行われている．したがって，言葉などの言語情報だけでなく，話者の表情，顔の向きや姿勢，口調や声の抑揚，ジェスチャといった非言語情報も話者の内面推定に有効であると考えられる．熊野ら [2] は複数人対話において各話者の表情と話者間の視線から共感状態を求めることで各話者の対人感情推定を行っており，円滑な多人数インタラクションの実現に向けた新しい研究の枠組みを提案している．Mahmoud ら [3] は自発的に生じたジェスチャとその時の精神状態との関係性を考慮した感情推定手法を提案しており，人物の感情認識においてジェスチャの有効性を示している．

我々は複数人対話において言語情報および非言語情報を効果

的に組み合わせて使用することで話者の内面を推定し，推定結果を基に各話者の意図伝達，意図理解を促進するインタラクション支援システムの構築を目指している．これまでに，単語の出現頻度や発話長，声の抑揚や発話時間を用いた複数人対話における盛り上がり判定 [4] や，対話中に発生する笑いに着目し，笑いの発生要因や笑いの大きさに基づく盛り上がり推定手法 [5] について研究を進めてきた．これらの研究においてより精度を向上させるには，画像情報の利用が必要である．対話中の画像情報から取得できる特徴としては話者の表情，顔の向きや姿勢，ジェスチャなどが挙げられるが，我々は対話中に話者がみせる姿勢に着目する．

本論文では姿勢推定手法の提案およびその有効性について検証する．姿勢は話者の内面状態を無意識のうちに表現する特徴の 1 つであり，話者の内面推定に有効である．例えば，話している相手に対して前傾姿勢の状態でタイミングよく相槌をうてば，話題に興味や関心を持っている状態であるといえる．姿勢推定には頭上方向から深度センサーを用いて撮影した人物画像を用いる．頭上方向から撮影することで人物の重なりによって発生するオクルージョンを回避することができ，かつ，被験者に与える心理的負担の軽減が期待できる．また，撮影に深度センサーを利用することで物体との距離情報を含んだ深度情報を取得することができ，撮影時に照明変動の影響を受けないというメリットがある．撮影したデータに含まれる深度情報から 7 つの特徴量を抽出，機械学習に適用することで姿勢推定を行う．

2. 関 連 研 究

本節では話者の内面推定における姿勢の重要性および頭上方向から撮影するメリットについて述べた関連研究を紹介する．Vargas [6] は対話中に現れる姿勢を言語調整動作における主要素であると述べている．言語調整動作とは聞き手が対話内容を理解し受け入れているかどうかを無意識の内に話し手に知らせる動作のことを指しており，話者の内面推定を行う際に非常に有効な要素であるとされている．また，姿勢は教育支援の分野においても学習者の内面推定を行う際にその有効性が述べられている．Mota ら [7] は「学習者の内面が姿勢に現れる」という仮定に基づき，姿勢推定を行うことで学習者の内面推定を図っている．なお，姿勢推定に用いる特徴量は圧力センサーを搭載した椅子から物理的に取得した圧力値から算出している．椅子と接触した部位の情報を詳細に取得できる反面，腕や手の動きなど接触していない部位の情報を取得することはできない．加えて，複数人対応を実現するためには人数分の装置を用意する必要があるなど制約が多い．

Grafsgaard ら [8] は深度センサーを学習者の正面に設置し，学習者の各部位の情報を間接的に取得することで姿勢推定を行い，姿勢と心理状態との相関関係を算出している．しかし，人物や身体の部位の重なりによって身体の一部にオクルージョンが生じた場合，取得できる情報量が減少するため識別精度が下がることが予想される．我々は人物識別のタスクでこの問題を解消するために頭上画像を利用する手法を提案している [9]．頭上画像を利用することで人物の重なりによるオクルージョンの問題は生じなくなる．さらに，被験者の視野に入らない位置にカメラを設置して撮影を行うことで，被験者の心理的負担を軽減することができる．以上の関連研究を踏まえて，本研究では深度センサーを備えたカメラを用いて頭上方向から人物を撮影することで姿勢推定を行う．

図 1　各姿勢の例.

3. 姿勢と内面状態

本節では我々が目指す多人数インタラクション支援のシステムを構築する上で重要な要素である話者の心理状態について議論する．冒頭でも述べたように，円滑な多人数インタラクションを実現するには, 各話者が自他の心理状態を自分自身で考慮した上で適切な行動をとらなければならない．話者の心理状態は対話中に非言語情報として現れることが多く，非言語情報の分析によって話者の心理状態，すなわち話者の内面を推定することは，多人数インタラクションの円滑化を実現するためには必要不可欠であるといえる．本論文では対話中にみられる姿勢に着目しているが，多人数インタラクションにおいて現れる姿勢には様々な種類が存在すると考えられ，すべての姿勢を判別することは容易ではない．そこで，対話環境を小規模な会議に限定して試験的にデータ収集を行うことで推定する姿勢の種類を制限している．なお，本論文で推定する姿勢は以下の 7 種類とする．図 1 に各姿勢の例を示す．

1. 前傾：興味・関心がある状態を表す
2. 後傾：退屈な状態を表す
3. 背伸び：疲労している状態を表す
4. 腕組み (前)：思考中である状態を表す
5. 腕組み (後)：退屈・疲労している状態を表す
6. 頬杖：思考中・退屈している状態を表す
7. ノート記入：興味・関心がある状態を表す

4. 提 案 手 法

本節では，提案手法について述べる．図 2 は提案手法の大まかな流れを表しており，深度情報を利用した人物領域の抽出，特徴量抽出，姿勢推定の大きく 3 つによって構成される．また，本研究では ASUS 社が提供している Xtion PRO LIVE(注1) の深度センサーを使用することで深度情報を取得している．

(注1)：ASUS Xtion PRO LIVE, http://www.asus.co.jp/Multimedia/Motion_Sensor/Xtion_PRO_LIVE/

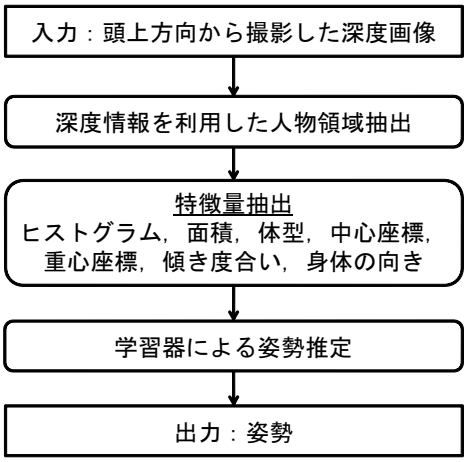

図 2　提案手法の概略図.

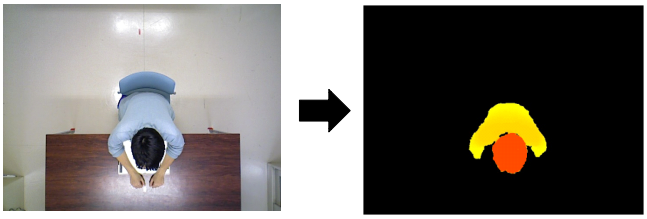

図 3　人物領域抽出の例.

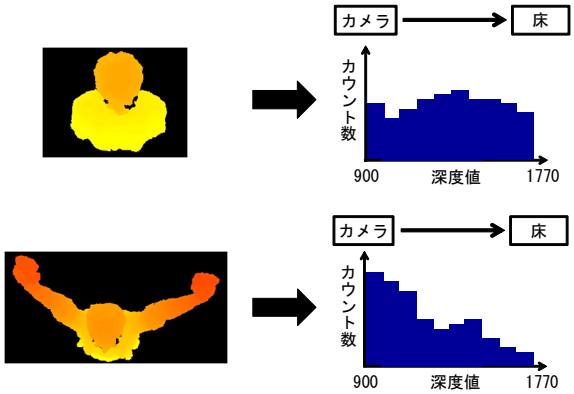

図 4　深度ヒストグラムの例.

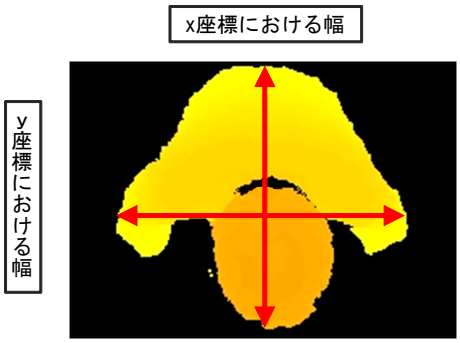

図 5　人物の体型特徴の例.

4.1　人物領域抽出

姿勢推定を行うためには，まず初めに撮影した画像から人物領域を抽出する必要がある．通常，人物抽出を行う際には背景差分を利用したものが多いが，照明変動など環境による影響を強く受けてしまうことが問題として挙げられる．この問題を解決するために深度情報を利用して人物と背景を切り離す．本論文で想定している状況では，人物・机・床が写るような画像が撮影される．そこで，人物領域の深度値と机領域の深度値の境界となる値を閾値として用いることで人物領域を抽出する．図 3 は深度情報に基づいた人物領域抽出の例である．なお，人物領域は深度値に応じて色の濃淡を変えて表現しており，濃淡が濃いほどセンサーとの距離は近く，濃淡が薄いほどセンサーとの距離は遠い．

4.2　特徴量抽出

この節では人物識別に用いる特徴量について述べる．本手法では，深度センサーによって深度情報を取得しており，取得した深度情報を基に以下の計 7 種類の特徴量を抽出している．

- 深度ヒストグラム
  深度センサーから取得した距離情報を基に深度ヒストグラムを生成する．姿勢の変化に伴って身体の各部位と深度センサーとの距離は異なるため，深度センサーとの距離情報を含んでいる深度情報のヒストグラムを作成することで姿勢推定に有効な特徴量が抽出できると考えられる．図 4 に後傾と背伸びの姿勢における深度ヒストグラムの例を示す．後傾は頭から胴体までの部位が深度センサーから一定の間隔で位置するため，全体的に緩やかなヒストグラムが作成される．一方，背伸びの場合は手や腕といった部位が深度センサーと近いため，ある範囲において突出したヒストグラムが作成される．
  深度センサーで取得できる距離情報の範囲は 500mm～5500mm であるが，この範囲では人物領域以外にも机や床領域など姿勢推定に不要な情報を含んだ深度ヒストグラムが生成されてしまう．今回は人物領域のみを表す範囲のヒストグラムを生成することでより正確な特徴量を抽出する．基礎実験の結果，人物領域のみを取得するのに最適な距離情報の範囲が 900mm～1770mm であると判明したため，この範囲における深度ヒストグラムを作成している．
- 人物の面積
  人物の面積および体型は姿勢推定において直感的で特有の特徴量となると考える．そこで本手法では人物領域における面積および体型特徴を利用する．面積特徴は抽出された人物領域のピクセル数の総和を算出することにより求める．
- 人物の体型
  体型特徴は人物領域の x 座標と y 座標の幅の値を使用することにより求める．これらの値は人物領域画像の周辺分布から抽出しており，人物領域の x 座標と y 座標において幅が最大となる値を使用する．図 5 は体型特徴の取得例を表している．
- 人物領域の中心座標
  深度センサーを用いて取得した人物領域は各姿勢で形状が異なる．このことから，人物領域の中心座標も姿勢毎

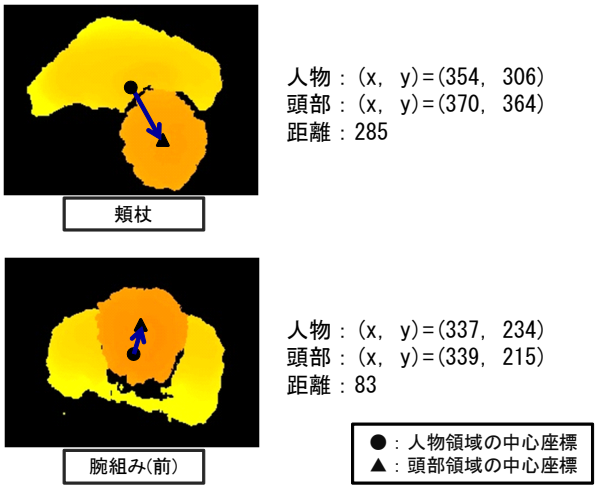

図 6　人物領域および頭部領域の中心座標と身体の傾き度合い特徴の例.

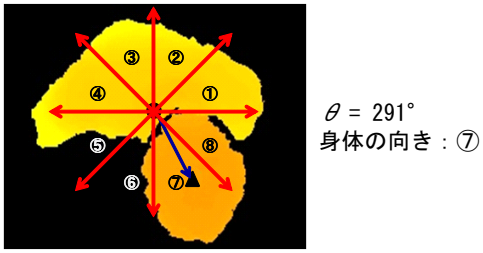

図 7　身体の向き特徴の例.

に変化すると考えられるため，姿勢推定を行う特徴量として利用する．図 6 は類杖の姿勢と腕組み (前) の姿勢より取得した人物領域画像を示している．なお，図 6 中の●は人物領域の中心座標点を表しており(注2)，各姿勢において●の座標位置 (x, y) が異なることがわかる．

本論文では，画像モーメントを利用することで人物領域より取得した重心を中心座標として扱う．人物の面積および体型特徴を取得した際に用いた人物領域を入力画像とし，以下の式において人物領域の 3 次の空間モーメント M_{pq} を求めることで領域の重心座標 $(\bar{x}, \bar{y})$ を算出する．式 (1) の p, q はそれぞれ x 方向，y 方向の次数を示している．

$$\begin{cases} m_{pq} = \sum_{x,y} (x^p \cdot y^q \cdot f(x,y)) \\ \bar{x} = \dfrac{m_{10}}{m_{00}}, \bar{y} = \dfrac{m_{01}}{m_{00}} \end{cases} \tag{1}$$

- 頭部領域の中心座標

  深度ヒストグラムの項目において，姿勢の変化によって身体の部位の位置に差分が出てくることを述べたが，特に頭部の位置の変化が大きいことがわかった．頭部の位置を示す頭部領域は図 6 において人物領域内の濃淡が濃い部分である．なお，図 6 中の▲は頭部領域の中心座標点を表している．類杖の姿勢では▲の座標位置が画像の下側に位置している一方，腕組み (前) の姿勢ではやや上部に位置していることがわかる．そこで，本研究では人物領域の部分領域である頭部領域より取得した重心座標を頭部領域の中心座標として扱う．人物領域の中心座標算出と同様に，頭部領域を入力画像として式 (1) を用いることで頭部領域の重心座標を算出する．

- 身体の傾き度合い

  取得した画像データを分析した結果，各姿勢の身体の傾き方には一定の傾向が現れていることが分かった．図 6 の矢印がその傾きを意味している．類杖の姿勢は矢印が画像の中心から下側に向かって伸びている一方，腕組み (前) の姿勢では矢印が画像の中心から上側に向かって伸びていることがわかる．これより，身体の傾き度合いによって話者がどのような姿勢をとっているのか推定可能であると考えられる．算出には前項で求めた人物領域の中心座標と頭部領域の中心座標を用いる．入力画像より人物領域の中心座標からみて頭部領域の中心座標がどこに位置するか求めた後，座標間の距離を取得する．この座標間の距離の値を身体の傾き度合いとする．身体の傾き度合いを表す座標間の距離 D は以下の式で算出する．

$$D = \sqrt{(x_c - x_g)^2 + (y_c - y_g)^2} \tag{2}$$

  式 (2) の x_c と y_c はそれぞれ人物領域の中心座標 (x_c, y_c) を，x_g と y_g はそれぞれ頭部領域の中心座標の (x_g, y_g) を表現している．図 6 に身体の傾き度合い特徴の取得例を示す．図中の矢印の長さが身体の傾き度合いである座標間の距離 D を表している．D の値が大きければ大きいほど矢印の長くなり，身体が大きく傾いているころを表す．

- 身体の向き

  身体の向きは人物領域の中心座標 (x_c, y_c) および頭部領域の中心座標 (x_g, y_g) を用いて角度を算出することで求める．身体の向き θ は以下の式で算出する．

$$\theta = -\arctan((y_g - y_c), (x_g - x_c)) * (180/\pi) \tag{3}$$

  身体の向きは θ の値と 45 ℃ずつで区切った範囲のどこに属しているかで表現する．図 7 は身体の向きの取得例を表している．

4.3　学習器による姿勢推定

抽出した特徴量に機械学習アルゴリズムを適用することで姿勢推定に最適な学習器を作成し，作成した学習器を用いて姿勢推定を行う．本論文では，機械学習アルゴリズムの Adaboost [11] を用いる．Adaboost は統計的学習手法 boosting の 1 つで，Freund らが提案した機械学習アルゴリズムである．Boosting とは，単純な予測が可能な弱分類器を組み合わせて，より高精度な分類器を作成する手法の 1 つである．本研究では弱識別器として C4.5 アルゴリズム [12] を用いる．C4.5 は，Quinlan らが考案した決定木学習アルゴリズムであり，属性とクラスで構成されたデータを与えることで，判別ノードと葉 (クラス) から成る決定木形式で分類器を作成する．また，機械学習の実装お

(注2)：本論文の座標位置は図 3 の画像において，その左上を (0,0) とした絶対座標である．図 6 の人物領域に関する画像だけが対象でないことに注意.

表 1　実験結果．

特徴量	精度 [%]
深度ヒストグラム	74.9 %
人物の面積	24.9 %
人物の体型	45.7 %
人物領域の中心座標	43.2 %
頭部領域の中心座標	39.2 %
身体の傾き度合い	24.1 %
身体の向き特徴	32.5 %
深度ヒストグラム+人物の面積	72.0 %
深度ヒストグラム+人物の体型	76.3 %
深度ヒストグラム+人物領域の中心座標	81.4 %
深度ヒストグラム+頭部領域の中心座標	82.7 %
深度ヒストグラム+身体の傾き度合い	75.4 %
深度ヒストグラム+身体の向き特徴	**83.0 %**
all	82.1 %

よび姿勢推定にはオープンソースソフトウェアである Weka [13] を使用する．

5.　実　　験

本章では，前節で述べた特徴抽出を行い，抽出した特徴量を用いて姿勢推定を行うことで，提案した手法の有効性を検証する．

5.1　実験データ

本実験に用いたデータを以下に示す．

- 訓練データ：1400 枚
- テストデータ：630 枚
- 被験者：13 人
- フレームレート：30fps
- 入力画像サイズ：640 × 480

なお，訓練データとデストデータは別の日に取得したものであり，各姿勢データは収集した動画データから 10 フレーム間隔で 10 枚ずつ抽出している．

5.2　実験結果と考察

本実験では 5 つの特徴量の組み合わせすべてについて評価実験を行った．表 1 は実験結果の一部である．なお，“＋”は特徴量の組み合わせを表しており，all は 7 つの特徴量をすべて用いたときの結果である．

まず，各特徴量についての考察を行う．今回取得した特徴量の中では最も有効な特徴量は深度ヒストグラムであった．この特徴量はカメラから人物の頭部や肩といった各部位までの距離情報を利用している．領域内における様々な距離情報を統合することで多くの情報を含んだ特徴量となり，その結果，他の特徴量に比べて高い精度を得ることができたと考えられる．一方で深度ヒストグラム以外の特徴量に関しては，今回の実験では精度が 24.1 %～45.7 %と深度ヒストグラムの精度 74.9 %と比較すると大幅に低くなった．精度が低い主な原因は特徴毎に取得できるデータのばらつき具合にあると考えられる．例えば，同じ種類の姿勢である図 8 において，人物の面積特徴はそれぞれ異なる値が取得されていた．また，人物領域の中心座標や重心座標，身体の向きなどは，異なる種類の姿勢であってもほぼ同一の特徴量が取得されたことによって誤識別が発生したことが考えられる．

図 8　同一姿勢の誤識別例．

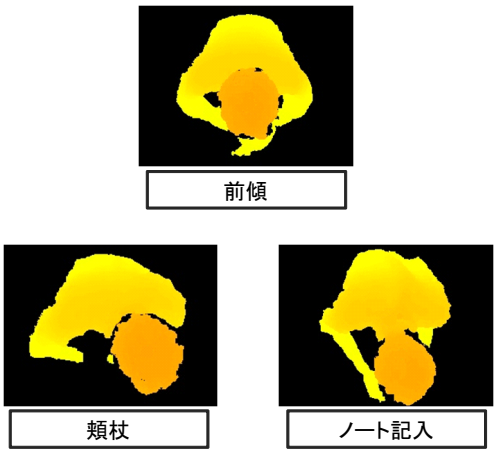

図 9　前傾・頬杖・ノート記入時の画像例．

続いて，推定精度の最も良かった組み合わせである「人物の深度ヒストグラム+身体の向き」について分析を行う．この 2 つの特徴量を組み合わせた推定結果の詳細を表 2 に示す．表 2 において図 9 のような前傾と頬杖 (14 事例)，ノート記入 (12 事例) の姿勢を識別する際に誤識別が多くみられた．また，腕組み (後) と後傾 (6 事例)，背伸び (12 事例) の姿勢を識別する際にも誤識別が多くみられた．これは，各姿勢より取得した深度ヒストグラムの傾向および身体の向きがほぼ同一であることから誤識別が発生したと考えられる．

このような誤識別を解消する方法の 1 つとして，時系列データの利用が考えられる．本論文では動画データから一定間隔で抽出した画像データを用いて姿勢推定を行っている．しかし，この手法では異なる種類の姿勢から類似した姿勢データを抽出する可能性があり，誤識別を引き起こす要因になっている．1 枚の画像のみで判断するのではなく，連続する複数枚の画像を用いることで身体の各部位がどのように遷移しているかを取得でき，この問題を解決することができると考えられる．例えば前傾と頬杖，ノート記入は身体全体が前に傾く動作までは同じであるが，頬杖は腕が顔の付近に移動し，ノート記入は手や腕の位置が小刻みに移動する．このように時系列データより状態

表 2　hist+direct の推定結果の詳細.

姿勢の種類	前傾	後傾	背伸び	腕組み (前)	腕組み (後)	頬杖	ノート記入
前傾	58	0	0	6	0	14	12
後傾	0	82	0	8	0	0	0
背伸び	0	0	82	0	8	0	0
腕組み (前)	1	3	0	86	0	0	0
腕組み (後)	0	6	12	6	65	0	1
頬杖	4	0	0	3	0	81	2
ノート記入	13	0	0	3	0	5	69

遷移の差分を取得することで姿勢推定の誤識別を解消し，精度向上につながると考えられる．

6. ま と め

本論文では対話者の内面推定の前処理である姿勢推定の手法として，深度センサー付きカメラで頭上方向から撮影した画像と深度情報を利用した姿勢推定手法を提案した．頭上方向から撮影することでカメラが視界に入ることによる心理的拘束性の問題を解消するとともに，頭上方向画像を用いた姿勢推定の手法において深度情報の有効性を確認した．姿勢推定に用いる特徴量として，(1) 深度ヒストグラム，(2) 人物の面積，(3) 人物の体型，(4) 人物領域の中心座標，(5) 頭部領域の中心座標，(6) 身体の傾き度合い，(7) 身体の向き の計 7 種類使用した．学習アルゴリズムには Adaboost と C4.5 を適用した．本手法において特に有効である特徴量は深度ヒストグラムと身体の傾き度合いを組み合わせたものであり，実験の結果，83.0 %の識別精度を得ることができた．このことから，深度情報を姿勢推定に用いることの有用性を確認することができた．

今後の課題としては，複数人対話時における姿勢推定の実現と興味推定にむけた環境の整理およびアノテーション手法の確立が挙げられる．本研究の最終目標は複数人対話における話者の内面推定および推定結果に基づいたインタラクション支援システムの構築である．複数人の姿勢状態を同時に推定する必要があるが，現段階では単体での姿勢推定のみ構築しているため，複数での姿勢推定へと拡張する予定である．

また，興味推定を行う環境として小規模な会議の場における複数人対話を想定しているが，人数や性別，対話内容等の異なる対話環境を評価することで本研究の枠組みを検証する．その際，推定された姿勢から内面推定を行うために姿勢状態へのアノテーションを進めていく．中村ら [14] はアノテーションを行う際には単純に姿勢状態と感情を一対一対応させるのではなく，対話のトピックや話者間の関係を考慮する必要があると述べている．同じ姿勢対象に対するアノテーションが評価者間で大きく異なるようだとアノテーションの信頼性に欠けるため，今後も実験を重ね，評価基準を考案していくことでアノテーションの信頼性を確保する．

文　献

[1] 坊農真弓, 高梨克也, “多人数インタラクションの分析手法,” オーム社, 2009.

[2] 熊野史朗, 大塚和弘, 三上 弾, 大和淳司, “複数人対話を対象とした顔表情の共起パターンの確率モデルに基づく共感状態の推定,” 画像の認識・理解シンポジウム (MIRU2011), IS2-35, 2011.

[3] M. Mahmoud, T. Baltrusaitis, P. Robinson and L. Riek, “3D corpus of spontaneous complex mental states,” In Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII 2011), Lecture Notes in Computer Science, Oct, 2011.

[4] 横山貴彦, 嶋田和孝, 遠藤 勉, “複数人談話における言語情報と非言語情報を利用した盛り上がり判定,” 言語処理学会第 18 回年次大会 (NLP2012), P1-26, 2012.

[5] 嶋田和孝, 楠本章裕, 横山貴彦, 遠藤 勉, “複数人談話における笑いの情報を考慮した盛り上がり判定,” 電子情報通信学会, 言語理解とコミュニケーション研究会 (NLC), NLC2012-7, pp.25-30, 2012.

[6] Marjorie Fink Vargas, 石丸 正, “非言語コミュニケーション,” 新潮社, 1987.

[7] S. Mota and R. W. Picard, “Automated Posture Analysis for detecting Learner’s Interest Level,” In Proceedings of Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction, CVPR HCI, June, 2003.

[8] J. F. Grafsgaard, K. E. Boyer, E. N. Wiebe and J. C. Lester, “Analyzing Posture and Affect in Task-Oriented Tutoring,” In Proceedings of the 25th Florida Artificial Intelligence Research Society Conference, 2012.

[9] R. Nakatani, D. Kouno, K. Shimada and T. Endo, “A Person Identification Method Using a Top-view Head Image from an Overhead Camera,” In Proceedings of 2nd International Workshop on Advanced Computational Intelligence and Intelligent Informatics(IWACIII2011), SS4-1, pp.739-746, 2011.

[10] V. Manohar, M. Shreve, D. Goldgof, S. Sarkar, “Modeling Facial Skin Motion Properties in Video and Its Application to Matching Faces across Expressions,” In Proceedings of IEEE International Conference on Pattern Recognition(ICPR), August, 2010

[11] Y. Freund and R. E. Schapier, “Experiments with a new boosting algorithm,” In Proceedings of ICML pp.148-156, 1996.

[12] J. R. Quinlan, “C4.5 Programs for Machine Learning,” Morgan Kaufmann Publishers, 1993.

[13] S. R. Garner, “WEKA: the Waikato environment for knowledge analysis,” In Proceedings of the New Zealand Computer Science Research Students Conference, pp.57-64, 1995.

[14] 中村和晃, 角所 考, 正司哲朗, 美濃導彦, 澤木美奈子, 南 泰浩, 前田英作, “擬人化エージェントとの音声対話時におけるユーザの非言語動作からの難/易及び興味/退屈の推定,” 電子情報通信学会和文論文誌 A, Vol.J95-A No.1 PP.85-96, 2012-1.