

教育現場におけるウェアラブル端末を用いた人間行動センシング

酒造 正樹*

概要：社会に出る前段階における大学生は、他者との共存関係を維持・構築するために、グループコミュニケーションの能力を磨く必要がある。例えば、グループディスカッションにおける意見交換やポスター形式における成果発表などが良い例である。大学は、情報発信力を高める教育として適切に学生を指導する必要があるが、未だ手探り状態である。このような教育に対し、従来、映像収録の書き起こしデータをもとに研究が行われてきた。小型モーションセンサなどを活用した人間の行動センシングの手法を取り入れることにより、学生に対する量的な評価と自然な介入を目指す。本稿においては、例としてポスターセッションにおけるデータ収集を紹介し、今後の大規模データ収集について議論したい。

Human Activity Sensing by Wearable Device in Educational Field

MASAKI SHUZO*

1. 加速度センサを用いた行動認識

MEMS 加速度センサは、2000 年代にアナログ・デバイス社が発売して以来、エレメントの小型化、低コスト化により、自動車、ゲーム機、スマートフォンなど様々なデバイスに用いられるようになった。また、小型加速度センサやジャイロセンサを用いて、実用的なアプリケーションを目指し人間の行動センシングを行う研究開発が盛んに行われるようになった[1]。しかし、それぞれの論文の基礎データとなる被験者に関して、少数であること、20 代大学生に偏りを持つこと、研究者毎に実験プロトコルが異なるなどの理由のために、追試を行うことが困難であり、ゆえに実用性に関して課題が残された状態であった。

そこで 2009 年 DICOMO において、名古屋大河口らは人間行動センシングの研究における大規模データベース構築を提案した[2]。人間行動センシングコンソーシアム (HASC) を設立し、分散データ収集アプリ (ロガー) を開発し、2013 までに、4 回の分散型データ収集チャレンジを行い、のべ 4 万ファイルの行動データを取得している。これらの基礎データをもとに機械学習をさせ、人の歩く、走る、静止、階段上り、降り、スキップの分類を行うことができるようになってきた。このような人の行動認識は、ヘルスケア分野、ほか様々な分類に渡り、近年の研究のホットトピックである。

2. グループコミュニケーションにおける情報発信力

大学の教育現場においては、専門課程に進むに連れ、学習者自らの考えを発信し、また他者と議論を重ねることで、コミュニティの中で協調して成長するグループワークを求められることが多くなる。例えば、少人数のグループディスカッションや研究成果発表の場におけるポスタープレゼンテーションなどは代表的なものである。このようなグル

ープコミュニケーションにおいて、話し方マニュアルやプレゼン資料の作り方、プレゼンノウハウ、といった類のマニュアルがあるものの、定量評価を行う試み[3][4][5]や、個人に対する適切なフィードバックを与える試み[6]は少ない。情報発信力を磨こうとするグループコミュニケーションにおける主流となる評価手法は、ビデオカメラによる録画映像を元に、言語・非言語情報の書き起こしが基本である。これには、熟練のアノテータを雇い、実時間の何倍もの時間をかける必要がある。もし、対象となる学習者から小型加速度センサや HASC ロガーによるデータ収集が簡単に行え、手でタグを付与してきた作業が、機械学習の結果により部分的としても、自動化が行えるならば、非常に有用と考える。

3. 研究のアプローチ

本研究においては、教育現場の中でもグループコミュニケーションに着目し、情報発信力を磨く上で人と人とのインタラクションを評価し、適切なフィードバックを与えることを目指す。よって、まず、映像データに個人の振る舞いや発話内容をアノテーションとして付与する。また、「情報発信力」に対する個人のパフォーマンス指標を第三者が与える。これらの情報を総合的に集約した大規模データベースを作成し、機械学習を行うアプローチを取る。

人の内面を評価する研究や、主観評価を伴う研究においては事例研究のスタイルが多い。ここで一旦、少量の事象の観察に多くの時間を割く傾向、2 は学会などの発表形式の一つであるポスターセッションにおけるプレゼンを意味する。次章においては現在進行中の実験の一例について紹介する。

4. ポスターセッションにおけるデータ収録

学習者のポスターセッションにおけるデータ収録に関して現状を述べる。

実験道具として、主に ATR プロモーションズ製小型モーションセンサ TSND151 およびスマートフォン内蔵のモー

* 東京電機大学システムデザイン工学部

ションセンサ (HASC ロガー[8]により加速度, ジャイロ, 地磁気, 音声, WiFi SSID を収集)を用いた。合わせて, 映像記録のためのカメラおよびマイクを用いた。

実験区域に A0 サイズのポスター (3~5 件) を掲示し, 学生発表者が数人の聴講者 (学生) に説明する。セッション全体は 30~40 分とした。1 回の説明の所要時間は約 10 分を目安とし複数回繰り返す。質疑応答の時間は特に制限を設けず, 発表者の裁量による。

蓄積したデータの種類は以下の通りである。

(1) ポスター発表者から見た正面, 後方俯瞰, 上空からの見下ろしの 3 アングルからの映像

(2) 参加者の発話音声

(3) 参加者の行動情報, 空間座標

(4) セッション後のアンケート

上記に加え, プレゼンに対する評価を教員が付与する。

これらのデータセットを用いて, 例えば, 発表の良し悪しに関して定量的な評価を与えることができる。また, 聴講者を含めた場のマネジメントとして, 人の偏りなどを考慮することもできる。

2017 年度に 4 回のポスターセッションを企画し, のべ約 60 人, うち発表者 18 人に参加いただいた[9][10]。2018 年度も不定期に実施する予定である。同様にグループディスカッションについても実施予定である。

なお, 本実験は東京電機大学倫理承認を得ており, 本人の同意取得を行った上で実施した。

5. 映像・センサデータによる行動タグの付与

次に, ポスターセッションの録画映像を分析するにあたり, 発話タグ以外に付与することで有効に働きそうな行動タグを述べる。

(1) 視線

(2) ハンドジェスチャー[11]

(3) 頭部動作 (上下, 左右, 斜め)

(4) 身体動作 (身体ねじり, 方向変更, お辞儀, 移動)

近年の AI 技術の進展に伴い, もし高精細に個人の体全体を, あるいは顔全体を映像記録することが許されるのであれば, 上記の項目は画像の処理により自動的に取得できる可能性がある。しかしながら, プライバシーの観点からたとえ本人が部分的な撮影許可を与えたとしても, 付随して映り込みの可能性がある他者とのインタラクションの場においては映像記録が困難な場合がある。

よって, 視線を除く(2)~(4)の項目のタグ付けにおいては, 加速度などセンサデータの分析手法が有効と思われる。(3) 頭部動作・(4) 身体動作 (ねじり, 方向変更, お辞儀) に関して, 適切な箇所に対してセンサの装着が可能であれば容易に結果を得ることができるであろう。また(2) ハンドジェスチャー, (4) 身体運動 (移動) に関して, 多くの行動認識の研究がなされている。認識の対象とする行動やそ

の認識精度などが明確になれば, 様々な分析モジュールをアドホックに利用すれば良いと考える。

着席形式のグループディスカッションにおいては, 身体の移動を伴わないので一部の分析項目が省略されるのみで, 共通した基本的には分析モジュールが使える見込みである。

6. センシングシステム, HASC ロガーの改良に向けて

これまでに述べてきたように, 小型モーションセンサを装着し, 行動認識を行うことはグループコミュニケーションの評価を行う上で有用である。ツールとしては, 大学生の個人所有率がほぼ 100% (2019 年調べ[12]) となったスマートフォン内蔵のセンサを活用することが第一の選択となる。HASC にて提供されたロガーアプリは, 内蔵センサのログを指定したサーバにネットワーク経由でアップロードすることができ, 他のデバイスに対するデータ収集の迅速正において優位性を持っている。

2017 年度に, HASC ロガーに関する必要最低限の操作説明を参加者に与え実験に参加してもらった。その際に発生した問題点は以下の通りである。

1. 人為的操作の誤り

単に習熟していなことから人為的なミスが発生した。記録開始と終了のボタンを押し間違え, 記録時間が極端に短いケース。また, スリープボタンを押してしまい, 計測もスリープさせてしまったケース。個人の発話内容の記録のために, 指向性ピンマイクを装着してもらったが, 取り付けが不適切な状態となり, 本人の声が録音されていなかったケースなどが起こった。

2. サーバ, ネットワークに起因するもの

操作者側の準備不足が関係しており, 例えばデータ収集サーバを大学の中に立てた場合に, 学内 LAN からのアクセスは可能でも, モバイルキャリアのネットワーク経由ではアクセスができない場合がある。また, 約 30 分の正しく記録されたデータを WiFi 経由でアップロードする場合に, 同程度の時間が必要であった。ネットワークの混雑状況にもよるが, 一度に複数人の同時アクセスが起こりえることを想定しなければならない。なお, ネットワーク混雑や次の予定のためにアップロードを後回しにすることによって, 実験データの回収が遅れ, 最悪の場合にはデータを紛失してしまうケースもあった。

3. 端末性能に起因するもの

最新スマートフォンは高機能であり様々なセンサを内蔵する。しかしながら, 無線のヘッドセットを使うことを前提に外付けマイク端子を持たない端末が存在する。従来型の指向性マイクを配布してプラグ接続する際には注意が必要である。

一方で昨今人気の「格安」スマホにおいては機能を削減されており, 例えばジャイロセンサを内蔵しないものも存

在する。また、旧式のモデルを長年愛用している場合にも、ジャイロセンサを内蔵しないので注意が必要である。

次に、空き容量不足にも注意が必要である。学生の中には、動画のダウンロードやアプリのインストールをしすぎるにより、実験データを途中で記録できないケースが発生した。前項にも関係するがユーザ空きスペースの確保のためにデータを削除したがるため、なるべく早めのデータ回収を完了することが大事である。

今後は、上記の問題点を踏まえて、HASC ロガーの改良を行いたい。

次に、近年の健康志向のブームもあって、脈拍（心拍）計測機能付きのスマートウォッチが普及の兆しを見せており、中でも Apple, Fitbit などの製品が人気を博している。また、JINS は眼電位の計測に基づき集中度を提示する眼鏡型デバイスを発売している。これらのデバイスには、モーションセンサが内蔵されており、実験室における専用デバイスではなく、一般ユーザが実生活で使用に耐えるものと理解される。このようなデバイスは、主にスマートフォンを経由してネットワークサーバに送信される設計であることが多い。

これまでに述べたように、身体中心（衣服のポケット内のスマホ）の行動データのみならず、頭や手からの行動データの取得が望ましい。市販の製品などのモーションデータ・行動データも共通フォーマットにて収容できるようアプリの改良を行いたい。

7. 今後のデータ共有

これまでの HASC においては興味のある研究者が互いにデータを持ち寄り、大規模データベースを構築することに、主たる目標があった。これを基礎データとした行動認識モデルの作成は、いわば副次的なものとして見ることができ、多くの研究者がその恩恵を受けてきている。

著者らのグループでは、HASC 同様に、教育現場に特化したデータ収集のコンソーシアムを再構築することを検討しているところである。しかし、個人を単位とするデータ共有を求めた HASC に対して、多人数の実験参加者を要するグループコミュニケーションのデータ収集は、一定の難しさが存在する。コンソーシアムの形成に向けては、近隣の大学間の連携を行うローカルエリアのアライアンス、学会・研究会レベルでの連携を行うソサイエティアライアンス、大型予算に紐づき国レベルでの連携を行うナショナルアライアンスへと発展させることが理想的である。

センサや映像、また行動タグなどが付与されたデータセットはコーパスとして、他の研究者が利用可能な状態に整備することも肝心である。スイス研究所の IDIAP では、ポスターセッションの映像データを公開している[13]。また、米国 CMU においては、専用のスタジオを作りセンサを装着した人間のマルチアングル映像撮影を可能にしている

[14]。著者らのグループにおいては、実験プロトコル、ツール（センサデバイス、ソフト）、統制環境下の定形行動、データフォーマットなどを議論し、効率よくデータ収集と分析が行えるよう準備を行っているところである。

また、HASC の結成からしばらくの時間が経過し、ヒトを対象とする研究に関して、倫理的ガイドラインに改正があったことも注意を要する。ヒト由来の情報を将来に再利用する場合（既存情報）について、自機関での二次利用・他機関への提供・他機関からの提供などに関して詳細にルール化されるに至った（H29 年 5 月施行 一部改正版「人を対象とする医学系研究に関する倫理指針」（厚労省）[15] および個人情報保護法）。研究を取り巻く法的环境の変化に対応しつつ、著者らの研究内容は本学の倫理委員会にて承認を受けている（H30 年 4 月）。学生教育に関する教育理念を同じとする研究グループの醸成を促しているところである。

参考文献

- [1] L. Bao *et al.*, “Activity Recognition from User-Annotated Acceleration Data,” *Pervasive Computing* pp. 1-17, 2004.
- [2] 河口信夫, “人間行動理解のための装着型センサによる大規模データベース構築にむけて,” *DICOMO 2010*, pp. 579-581, 2010..
- [3] 岡田将吾他, “マルチモーダル情報に基づくグループ会話におけるコミュニケーション能力の推定,” *人工知能学会論文誌*, vol.31, no. 6, AI30-E, 2016.
- [4] 武川直樹他, “グループディスカッションにおける発言者の言語/非言語の表出と評価者評価の関係の分析,” *電子情報通信学会論文誌 D*, vol. J101-D, No. 2, pp. 284-293, 2018.
- [5] H. Hung *et al.*, “Detecting F-formations as Dominant Sets,” *Proc. ACM ICMI*, pp. 231-238, 2011.
- [6] K. Kurihara *et al.*, “Presentation Sensei: A Presentation Training System Using Speech and Image Processing,” *Proc. ACM ICMI*, pp. 358-365, 2007.
- [7] 渡部信一, “教育現場の「コンピテンシー評価」—「見えない能力」の評価を考える”, ナカニシヤ出版, 2017.
- [8] HASC ロガー, <http://hasc.jp/tools/hasclogger.html>
- [9] 森田幸輔他, “ポスターセッションにおける会話場支援者の行動分析,” *HCG シンポジウム 2017*, 2017.
- [10] H. Tokunaga *et al.*, “Preliminary Analyses of Spatial Positions of Poster Session Audience and Their Joining in/Leaving Behaviors,” *Proc. 11th International Conference on Language Resources and Evaluation (LREC 2018)*, 2018
- [11] 片岡邦好, “第 5 章 マルティモーダルの社会言語学 日・英対照による空間ジェスチャー分析の試み,” 井上逸兵編集 *社会言語学*, 朝倉書店, pp. 82-106, 2017.

- [12] <https://saponet.mynavi.jp/release/student/life/2019> 年卒マ
イナビ大学生のライフスタイル調査/
- [13] Idiap Poster Data.
<https://www.idiap.ch/dataset/idiap-poster-data>
- [14] CMU Panoptic Dataset,
<http://domedb.perception.cs.cmu.edu/>
- [15] 厚労省「人を対象とする医学系研究に関する倫理指針」
(平成 29 年 2 月 28 日一部改正)