National University Corporation
Tsukuba University of Technology

Shiraishi Lab

IARIA

One-Handed Signs: Standardization for Vehicle Interfaces and Groundwork for Automated Sign Language Recognition

○Akihisa Shitara[†*], Taiga Kasama[*], Fumio Yoneyama[*], Yuhki Shiraishi[*]

[†]Graduate School of Library, Information, and Media Studies, University of Tsukuba, Japan
[*]Faculty of Industrial Technology, Tsukuba University of Technology, Japan

Email: theta-akihisa@digitalnature.slis.tsukuba.ac.jp

DigitalWorld 2024 Congress
The Seventeenth International Conference on Advances in Computer-Human Interactions
ACHI 2024
May 26, 2024 to May 30, 2024 - Barcelona, Spain

Presenter: Akihisa Shitara

- d/Deaf
- Ph.D Student
  - Graduate School of Library, Information, and Media Studies, University of Tsukuba
- Part-time Researcher
  - Faculty of Industrial Technology, Tsukuba University of Technology
- Research Topics
  - Accessibility
  - Human Computer Interaction
  - Deaf Physicality

Contents

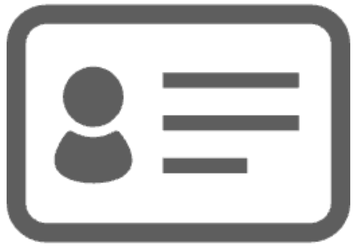
1. Introduction
2. Related Work
3. Needs Survey
4. Proposal of Vehicle Interfaces Standard
5. Data Collection and Annotation Rules
6. Discussion
7. Conclusion and Future Work

Contents

1. Introduction
2. Related Work
3. Needs Survey
4. Proposal of Vehicle Interfaces Standard
5. Data Collection and Annotation Rules
6. Discussion
7. Conclusion and Future Work

Introduction

d/Dhh

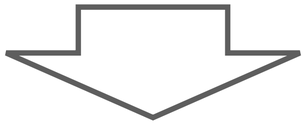


driver’s license

+

hearing aids / cochlear implants
or
magnifying glasses and hearing impairment markers

- Act for Eliminating Discrimination against Persons with Disabilities
- Disqualification provisions stipulated in Article 88 of the Road Traffic Act



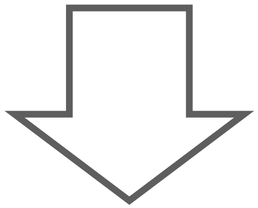
environmental sounds (emergency information etc)

Communication issues

Introduction

In-vehicle sign language recognition standard

- Installation location of drive recorders
- One-handed sign that occurs during sign language conversations



This research topics:

1. Proposing limited sign language by selecting one-handed signs
2. Evaluation of the one-handed signs by the parties concerned

Contents

1. Introduction
2. Related Work
3. Needs Survey
4. Proposal of Vehicle Interfaces Standard
5. Data Collection and Annotation Rules
6. Discussion
7. Conclusion and Future Work

Related Work

Image recognition

Mukai et al. [4]: 41 static fingerspelling in JSL
average recognition accuracy of 86%

Hosoe et al. [5]: only static fingerspelling in JSL
recognition rate of 93%

Jalal et al. [6]: static fingerspelling in ASL
recognition rate of 99%

Kondo et al. [7]: static and dynamic fingerspelling in JSL
93.8% identification rate

[4] N. Mukai, N. Harada, and Y. Chang, “Japanese fingerspelling recognition based on classification tree and machine learning,” in 2017 Nicograph International (NicoInt). New York, NY, USA: IEEE (Institute of Electrical and Electronics Engineers), June 2017, pp. 19–24.
[5] H. Hosoe, S. Sako, and B. Kwolek, “Recognition of jsl finger spelling using convolutional neural networks,” in 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA), May 2017, pp. 85–88.
[6] M. A. Jalal, R. Chen, R. K. Moore, and L. Mihaylova, “American sign language posture understanding with deep neural networks,” in 2018 21st International Conference on Information Fusion (FUSION). New York, NY, USA: IEEE (Institute of Electrical and Electronics Engineers), July 2018, pp. 573–579.
[7] M. Kondo, N. Kato, K. Fukui, and A. Okazaki, “Development and evaluation of an interactive training system for both static and dynamic fingerspelling using depth image,” IEICE technical report, vol. 114, no. 512, 2015, pp. 23–28, (in Japanese).

Related Work

Sensor glove recognition

Cabrera et al. [8]: static fingerspelling in ASL
recognition rate of 94.07%

Mummadi et al. [10]: fingerspelling in FSL
average recognition rate of 92%
with an F1-score of 91%.

Kakoty et al. [11]: fingerspelling (C, I, J, L, O, U, Y, W) in ISL
fingerspelling (A to Z) in ASL
signed numbers (0 to 9)
average recognition rate of 96.7%

Chong et al. [12]: 28 sentences in ASL
accuracy of up to 99.89%

[8] M. E. Cabrera, J. M. Bogado, L. Fermin, R. Acuna, and D. Ralev, “Glove-based gesture recognition system,” in Adaptive Mobile Robotics. World Scientific, 2012, pp. 747–753.
[10] C. K. Mummadi, F. P. P. Leo, K. D. Verma, S. Kasireddy, P. M. Scholl, and K. Van Laerhoven, “Real-time embedded recognition of sign language alphabet fingerspelling in an imu-based glove,” in Proceedings of the 4th International Workshop on Sensor-Based Activity Recognition and Interaction, ser. iWOAR ’17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 1–6. [Online]. Available: https://doi.org/10.1145/3134230.3134236
[11] N. M. Kakoty and M. D. Sharma, “Recognition of sign language alphabets and numbers based on hand kinematics using a data glove,” Procedia Computer Science, vol. 133, 2018, pp. 55–62.
[12] T.-W. Chong and B.-J. Kim, “American sign language recognition system using wearable sensors with deep learning approach,” The Journal of the Korea Institute of Electronic Communication Sciences, vol. 15, no. 2, 2020, pp. 291–298.

Related Work

Data collection

Image recognition

- PHOENIX-2014, PHOENIX-2014-T (GSL)[16, 17]
- Chinese Sign Language[16]
- ChicagoFSwild, ChicagoFSwild+ (Chicago Sign Language)[16-19]

[16] H. Zhou, W. Zhou, Y. Zhou, and H. Li, “Spatial-temporal multi-cue network for continuous sign language recognition,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 07, 2020, pp. 13 009–13 016
[17] “PWTH-PHOENIX-Weather 2014-T),” 2019, URL: https://wwwi6. informatik.rwth-aachen.de/~koller/RWTH-PHOENIX-2014-T/ [retrieved: April, 2024].
[18] B. Shi and et al., “Fingerspelling recognition in the wild with iterative visual attention,” in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 5399–5408.
[19] “Chicago Fingerspelling in the Wild Data Sets (ChicagoFSWild, ChicagoFSWild+),” 2019, URL: https://home.ttic.edu/~klivescu/ChicagoFSWild [retrieved: April, 2024].

Sensor glove recognition

Not require pictures of the face; thus privacy concerns

Contents

1. Introduction
2. Related Work
3. **Needs Survey**
4. Proposal of Vehicle Interfaces Standard
5. Data Collection and Annotation Rules
6. Discussion
7. Conclusion and Future Work

Needs Survey

Method: Google Forms

Respondents: 143 (88: male, 53: female, 2: Unpecided)
Mean age: 32.01 (SD: 13.87)
Identity: Deaf 68, deaf 20, hard of hearing 35, hearing 4, unsure etc 11, others 3, invalid 2

Findings:

When calling out to the driver while driving

- hard of hearing more likely to use voice than Deaf
  (with a significant difference)

Communicate with the driver while the vehicle is in motion

- more than 75% of d/Dhh individuals wanted to use sign language

Over 75% of the respondents

- difficulties communicating with a person who can use sign language
  while sitting in the passenger seat
- want to communicate with
  the person sitting in the passenger seat while driving

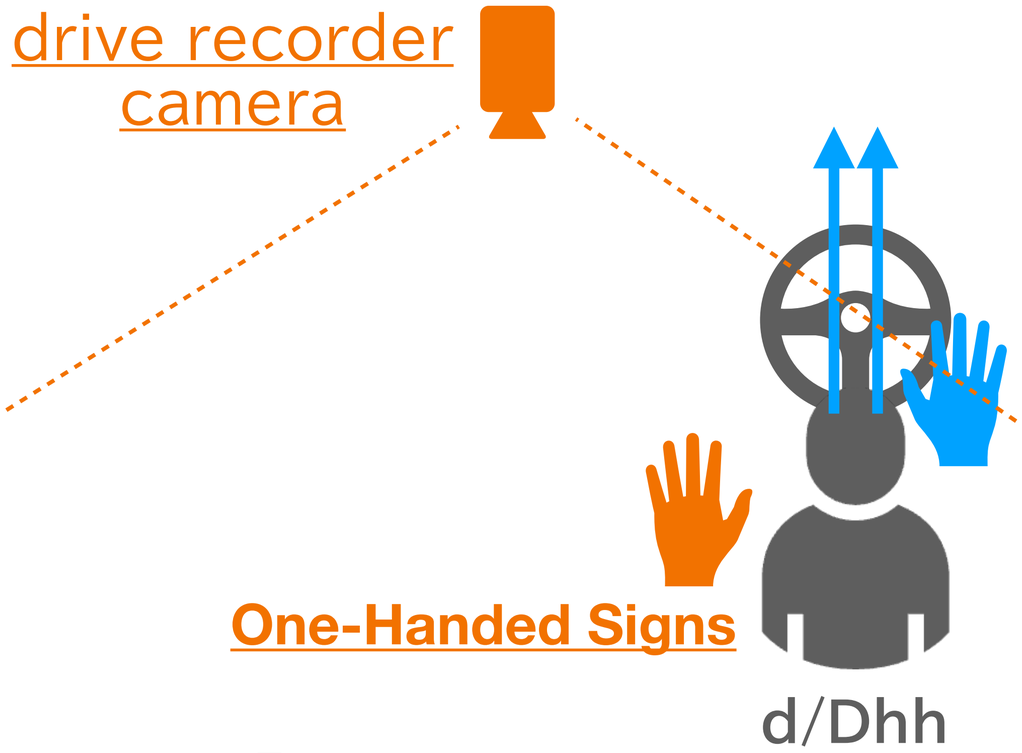
Contents

1. Introduction
2. Related Work
3. Needs Survey
4. **Proposal of Vehicle Interfaces Standard**
5. Data Collection and Annotation Rules
6. Discussion
7. Conclusion and Future Work

Proposal of Vehicle Interfaces Standard

In-vehicle sign language recognition standard

Limitation:
- must always face forward
- needs to grip the steering wheel

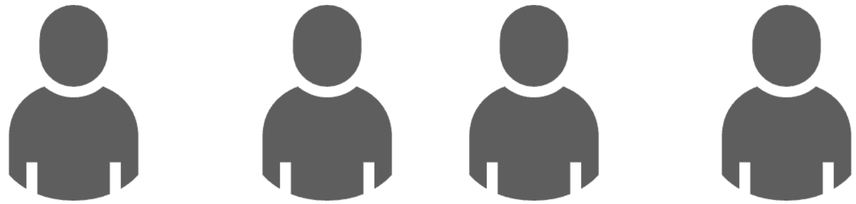


Proposal of Vehicle Interfaces Standard

One-Handed Signs

Rule:

- Select one-handed sign words and short sentences
- Not using grammar that:
  employs facial expressions, other elements



d/Deaf deaf deaf hearing

- driver’s license
- driving experience

Words: 58 Short sentences: 40

“danger”, “watch out”, “emergency vehicle”, some directions, small talk

Proposal of Vehicle Interfaces Standard

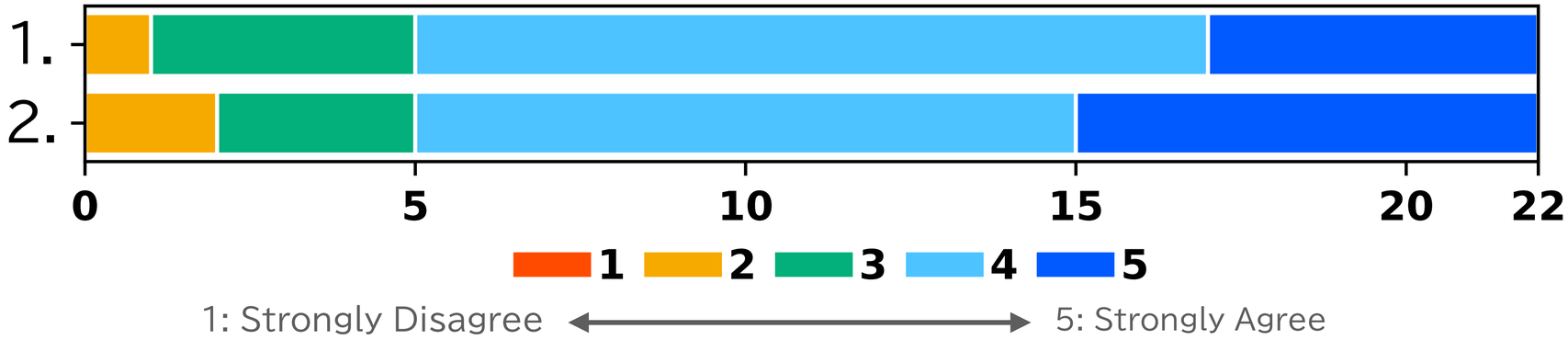
After completing the data collection experiment:

Participants: 22

Sign language experience: 11.3 years (SD: 7.1years)

Evaluation of Necessity:

1. In-vehicle sign language recognition
2. Use of one-handed signs in side



Binomial test's two-sided significance level of 0.05:

1. and 2. is p=0.017, confirming a significant difference

Proposal of Vehicle Interfaces Standard

After completing the data collection experiment:

Participants: 10

Sign language experience: 11.6 years (SD: 5.76years)

One-Handed Signs

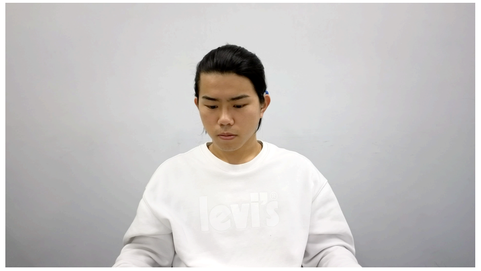
Expression for each word

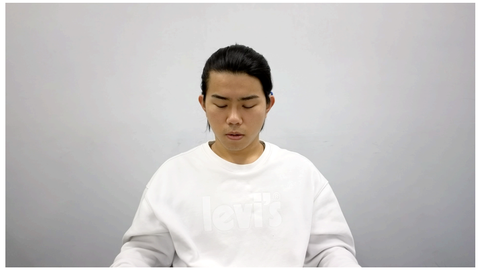
1. Unnaturalness with all words
2. Selected word that negative responses was 3 or more
   - Binomial test’s two-sided significance level of 0.05
   - No significant difference was confirmed (p=0.344)
3. Acceptability with 7 words
4. Negative responses was 1 or less
   - Binomial test’s two-sided significance level of 0.05
   - No significant difference was confirmed

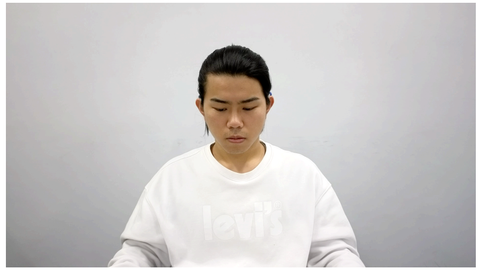
Proposal of Vehicle Interfaces Standard

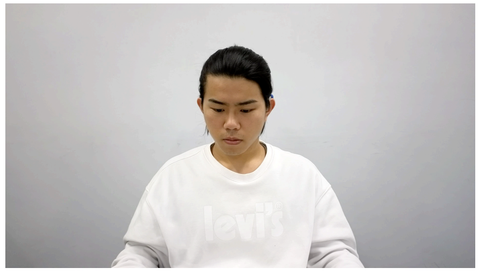
One-Handed Signs

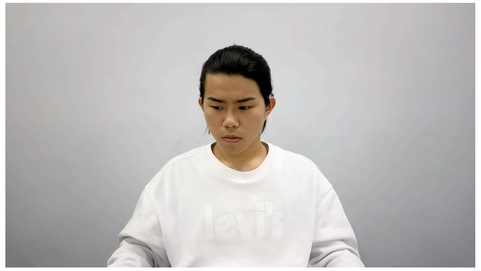
Expression for each word that Uncomfortable but acceptable


w2: 休憩(rest)

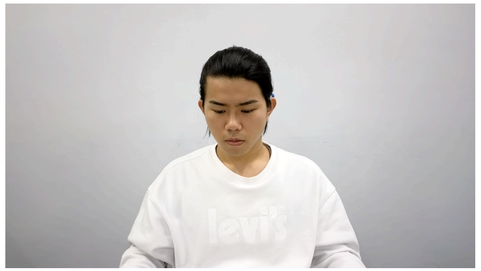

w4: 落ち着く(calm)

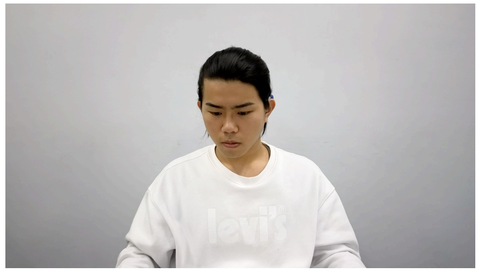

w6: コンビニ(convenience)


w7: 事故(accident)


w10: 信号(traffic light)


w16: 次(next)


w22: 道(way)


w27: 別(another)

Proposal of Vehicle Interfaces Standard

After completing the data collection experiment:

Participants: 10

Sign language experience: 11.6 years (SD: 5.76years)

One-Handed Signs

Expression for each short sentence

1. Unnaturalness with all short sentences
2. Selected short sentences that negative responses was 3 or more
   - Binomial test’s two-sided significance level of 0.05
   - No significant difference was confirmed (p=0.344)
3. Acceptability with 12 short sentences
4. Negative responses was 1 or less
   - Binomial test’s two-sided significance level of 0.05
   - No significant difference was confirmed

Proposal of Vehicle Interfaces Standard

One-Handed Signs

Expression for each short sentence that Uncomfortable but acceptable

s3: 危険, 落ち着いて見て
[Danger, watch calmly]

s4: コンビニ, 行きますか？
[Are we going to
the convenience store]

s9: 次の信号で右折
[Turn right at the next signal]

s12: 次はどこ行く？
[Where are we going next?]

s13: 次に左折するのはどこ？
[Where do we turn left next?]

s18: 別の道があるか調べて
[Check if there is another route]

Proposal of Vehicle Interfaces Standard

One-Handed Signs

Expression for each short sentence that Uncomfortable but acceptable

s21: 緊急車両が来るから右寄せて
[Pull over to the right because an emergency vehicle is coming]

s24: ガソリンスタンドはない
[There is no gas station]

s27: 左に止めても構わない
[You can stop on the left]

s32: おーい, ウィンカー消して
[Hey, turn off the turn signal]

s34: 次の交差点で左折だから, 左寄せして
[It’ s a left turn at the next intersection, so stay to the left.]

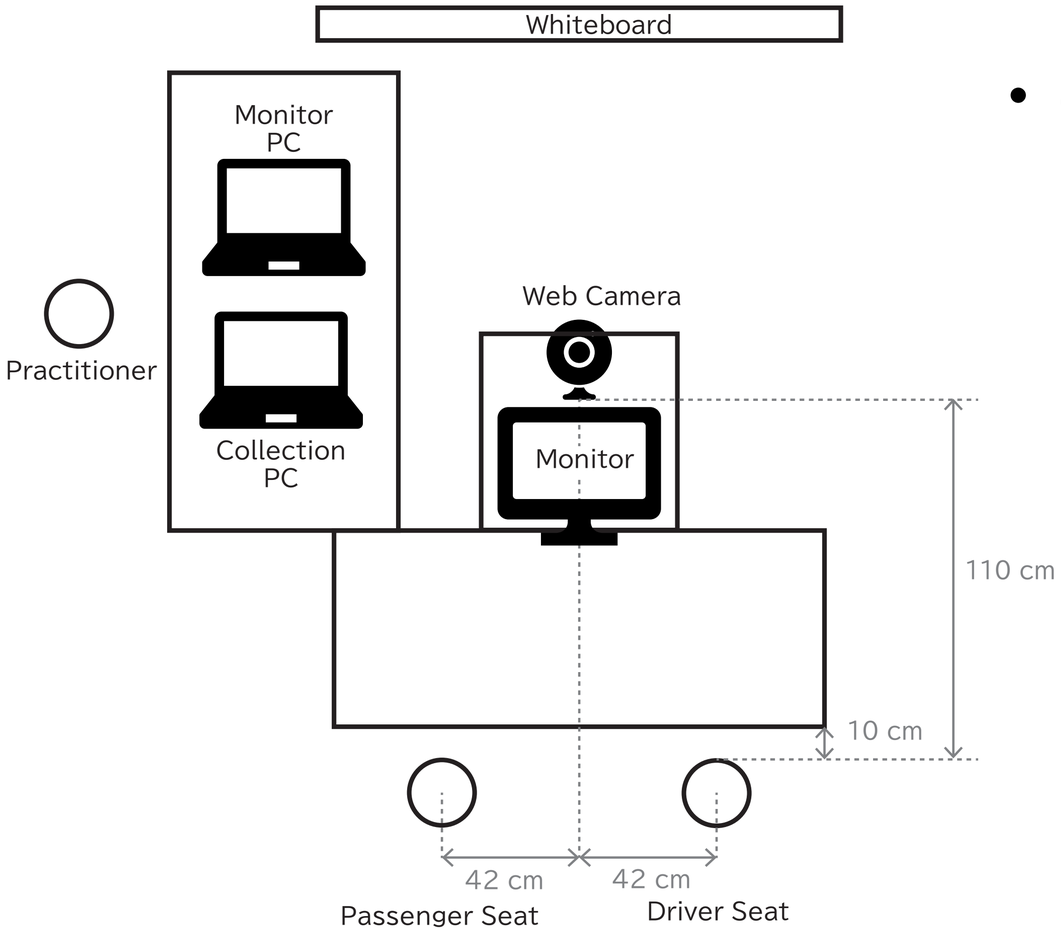
s37: 3つ目の交差点で右折だから, 右寄せして
[It’ s a right turn at the third intersection, so stay to the right.]

Contents

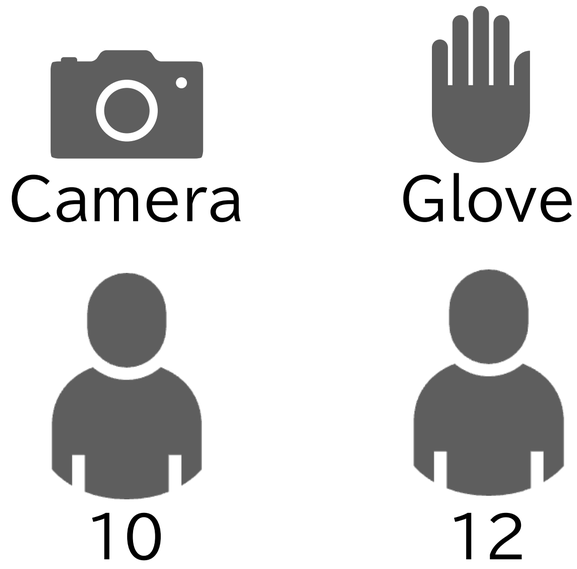
1. Introduction
2. Related Work
3. Needs Survey
4. Proposal of Vehicle Interfaces Standard
5. **Data Collection and Annotation Rules**
6. Discussion
7. Conclusion and Future Work

Data Collection and Annotation Rules

Data Collection



- 4 time each one short sentence
- driver seat and passenger seat



National University Corporation Tsukuba University of Technology
Shiraishi Lab

Data Collection and Annotation Rules

Annotation Rules

- Annotation labels: “word” and “transition”
- Dynamic sign language expressions
  - Start point: handshape is determined
  - End point: handshape collapses
    or hand moves away from the ending position
- Static sign language expressions
  - Start point: moment just before
    reaching the fixed position
  - End point: moment when the hand moves away
    from the fixed position

Contents

1. Introduction
2. Related Work
3. Needs Survey
4. Proposal of Vehicle Interfaces Standard
5. Data Collection and Annotation Rules
6. **Discussion**
7. Conclusion and Future Work

Discussion

Validity of One-Handed Sign Expressions

Necessary to increase the number of participants and consider participant attributes

- Sign language experience, JSL or SJ
- Identity, Involvement in deaf school communities

Annotation Rules

- “Here” and “There” use “number 1” that handshape

  “Here”: points to one’s own feet

  “There”: depends on which place is being pointed
  —> but since it is a high-context expression in Japanese

- “Why” and “Go” use “number 1” that handshape and twist
  the wrist downward from one’s chest to the front

  “Why”: aiming for 45 to 90degrees, twice

  “Go”: aiming for more than 80 degrees, once

- “(Number)th”

  Each “th” as a single “th”

Discussion

Differences between Camera and Gloves

Not yet been examined:
- Differences between cameras and gloves
- As well as which one is more suitable

Necessary:
- Evaluate the accuracy of sign language recognition
- User studies to determine which input interface is preferable

Limitation

Not yet been completed:
- Annotation process
- Evaluation experiments

Cannot have validity of expressions

Not targeted:
- facial expressions, other grammatical elements

Annotation labels is not high of when generalized

Contents

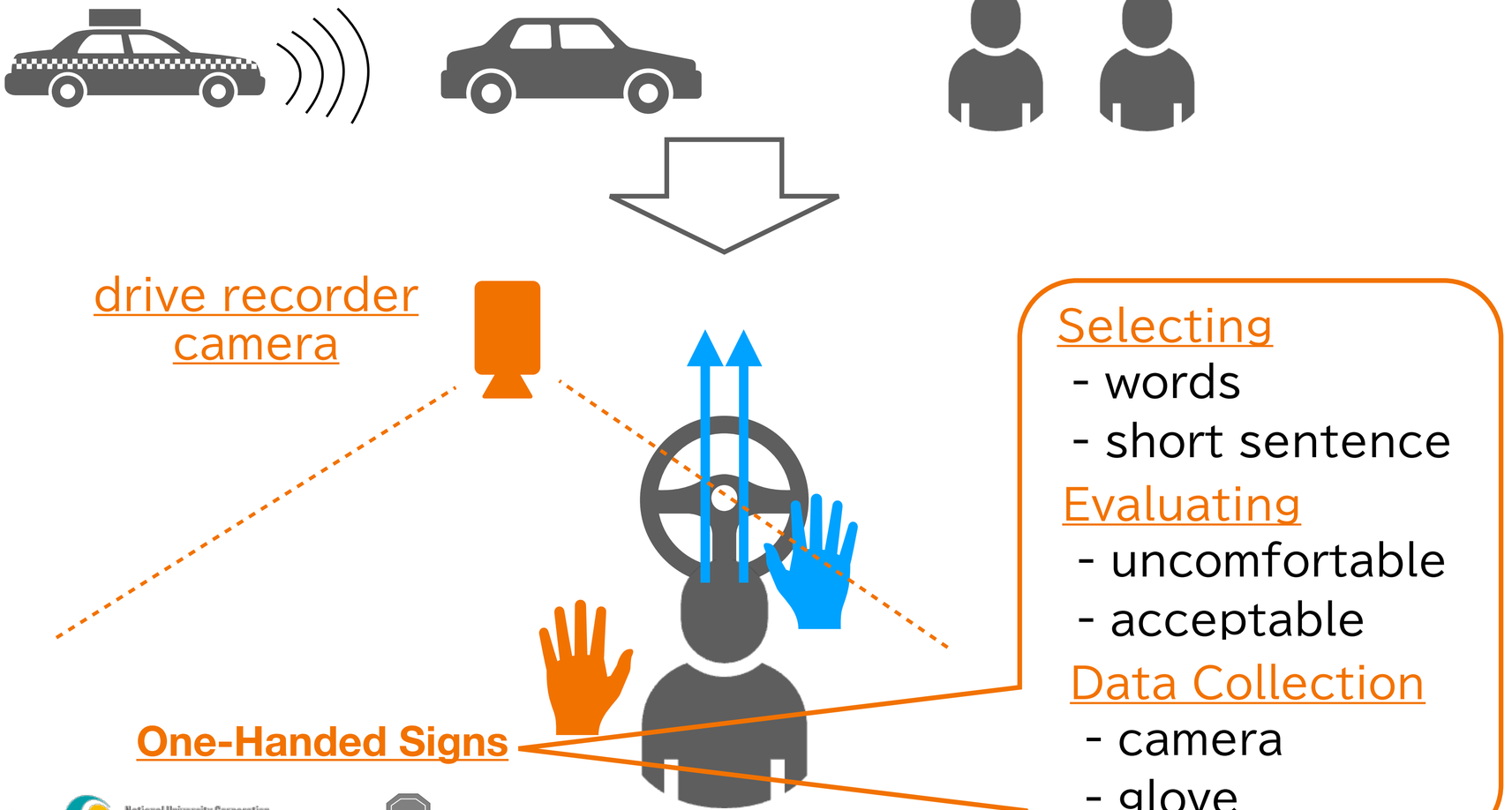
1. Introduction
2. Related Work
3. Needs Survey
4. Proposal of Vehicle Interfaces Standard
5. Data Collection and Annotation Rules
6. Discussion
7. **Conclusion and Future Work**

Conclusion and Future Work

Conclusion

environmental sounds
(emergency information etc)

Communication issues



Selecting
- words
- short sentence

Evaluating
- uncomfortable
- acceptable

Data Collection
- camera
- glove

Conclusion and Future Work

Future Work

1. Expand the dataset
   Robustness and generalizability of the sign language recognition
2. Develop and implement real-time recognition
   Assist d/Dhh drivers in communication and navigation
3. Conduct user studies
   Evaluate effectiveness and usability of the proposed system
4. Refine annotation rules
   Based on feedback
   from the d/Dhh community and research findings
5) Investigate multimodal approaches
   Explore the integration of other modalities,
   such as facial expression or eye-tracking

Acknowledgment

This work was supported by:

- Woven by Toyota, Inc.
- JSPSKAKENHI Grant Number #19K11411
- Promotional Projects for
  Advanced Education and Research in NTUT