Word segmentation of output response for sign language devices

Za'Bah, N.F., Nazmi, A.A.A.M., Azman, A.W.
Department of Electrical and Computer Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia

Abstract
Segmentation is an important aspect of translating finger spelling of sign language into Latin alphabets. Although the sign language devices that are currently available can translate the finger spelling into alphabets, there is a limitation where the output is stored in a long continuous string without spaces between words. The system proposed in this work is meant to be used together with a text-generating glove device. The system used text input string and the string is then fed into the system, one character at a time, and then it is segmented into words that is semantically correct. The proposed text segmentation method in this work is by using the dynamic programming and back-off algorithm, together with the probability score using word matching with an English language text corpus. Based on the results, the system is able to properly segment words with acceptable accuracy. © 2020, International Islamic University Malaysia-IIUM.
| # | Reference | Title | Year | Journal | Pages | DOI | View at Publisher |
|---|-----------|-------|------|---------|-------|-----|-------------------|
| 1 | Rao, G.A., Kishore, P.V.V. | Sign language recognition system simulated for video captured with smart phone front camera | (2016) | *International Journal of Electrical and Computer Engineering*, 6 (5), pp. 2176-2187. | Cited 19 times. | [http://www.iaescore.com/journals/index.php/IJECE/article/view/1144/4207](http://www.iaescore.com/journals/index.php/IJECE/article/view/1144/4207) | [View at Publisher](http://www.iaescore.com/journals/index.php/IJECE/article/view/1144/4207) |
| 2 | Lee, B.G., Lee, S.M. | Smart Wearable Hand Device for Sign Language Interpretation System with Sensors Fusion | (2019) | *IEEE Sensors Journal*, 99, | pp. 1-1. | | |
| 3 | Shukor, A.Z., Miskon, M.F., Jamaluddin, M.H., Ali Ibrahim, F.B., Asyraf, M.F., Bahar, M.B.B. | A New Data Glove Approach for Malaysian Sign Language Detection | (2015) | *Procedia Computer Science*, 76, | pp. 60-67. | Cited 26 times. | [http://www.sciencedirect.com/science/journal/18770509](http://www.sciencedirect.com/science/journal/18770509) | [View at Publisher](http://www.sciencedirect.com/science/journal/18770509) |
| 4 | Norvig, P. | Natural Language Corpus Data. | Cited 54 times. | | | | |
| 5 | Jiang, W. | | (2015) | | | | |
| 6 | Khan, S.N., Khan, K., Khan, W., Khan, A., Subhan, F., Khan, A.U., Ullah, B. | Urdu word segmentation using machine learning approaches | (2018) | *International Journal of Advanced Computer Science and Applications*, 9 (6), | pp. 193-200. | Cited 3 times. | [http://thesai.org/Downloads/Volume9No6/Paper_28-Urdu_Word_Segmentation.pdf](http://thesai.org/Downloads/Volume9No6/Paper_28-Urdu_Word_Segmentation.pdf) | [View at Publisher](http://thesai.org/Downloads/Volume9No6/Paper_28-Urdu_Word_Segmentation.pdf) |
| 7 | Tanaya, D., Adriani, M. | Dictionary-based Word Segmentation for Javanese | (2016) | *Procedia Computer Science*, 81, | pp. 208-213. | Cited 5 times. | [http://www.sciencedirect.com/science/journal/18770509](http://www.sciencedirect.com/science/journal/18770509) | [View at Publisher](http://www.sciencedirect.com/science/journal/18770509) |
| 8 | Srinivasan, S., Bhattacharya, S., Chakraborty, R. | Segmenting web-domains and hashtags using length specific models | (2012) | *ACM International Conference Proceeding Series*, | pp. 1113-1122. | Cited 14 times. | ISBN: 978-145031156-4 | [View at Publisher](http://pypi.org/project/wordsegmentationl) | [View at Publisher](http://pypi.org/project/wordsegmentationl) |
