

単語予測と文章予測によるスマートグラス向けサジェストシステム

林 健太郎* 入江 英嗣* 坂井 修一*

概要. 本研究では, 単語単位の入力予測と文章単位の入力予測を組み合わせたスマートグラスのサジェストシステムを提案する. このシステムでは, 見ている対象のキーワードを選択するだけ操作で文章入力ができる. これまでのスマートグラスのテキスト入力に関する研究はできるだけ正確に速く入力することを目指したものが多く, しかし, スマートグラスでは, 入力しようとした文章を一文字ずつ時間をかけて入力するよりも, およその内容をすばやく入力できる方がユーザーにとって好ましいということが考えられる. 本論文では, 視界画像を利用した予測による提案システムを実装し, 使用場面を想定したユーザー実験を行った.

1 はじめに

スマートグラスの普及に向けた課題の一つとしてテキスト入力の問題がある. 現在, スマートグラスにおけるテキスト入力は音声入力による方法とディスプレイに表示された仮想キーボードを使用する方法が主流であるが, 音声入力は公共の場では利用できないという問題があり, 仮想キーボードを用いた方法では一回の入力にかかる時間が長く誤入力も起きやすい.

しかし, スマートグラスを使用する場面では, 時間をかけて確実に入力するよりも, およその内容をすばやく入力できる方が好ましい場面の方が多いと考えられる. そこで, 見ている対象についてキーワードを選択するだけの少ない操作で入力することを目的として, 単語単位の入力予測と文章単位の入力予測を組み合わせたサジェストシステムを提案する. 視界画像を用いた予測により, 入力内容の中心となるキーワードから入力できる. 本論文では提案システムの実装を行い, ユーザー実験によってシステムの有効性を検証した結果を報告する.

2 関連研究

スマートデバイスにおけるテキスト入力のさらなる効率化を目的として, 様々なサジェストシステムが提案されている. Arnold らは [1] 候補単語を提示するとともに予測されたフレーズを提示するサジェストシステムを用いた結果, ユーザーの入力候補を選ぶ回数が増えたと報告している. 鈴木ら [6] は, 単語単位のテキスト入力により入力回数を大きく減らした. 土屋ら [5] は画像キャプション生成モデルをインタラクティブに利用した入力補助システムにより, キーストロークを 16%削減した. Smart Reply[2] では電子メールの返信メッセージ候補を自動で生成することで, 短い返信メッセージを入力す

る手間を 10%軽減した. このような文章単位の入力予測は, 提示された候補が入力に使われる場合は非常に便利であるが, そうでない場合は全く使用されないか, 後から修正を行う必要がある.

3 直感的テキスト入力のためのサジェストシステムの提案

3.1 概要

そこで本研究では, 文章単位の入力予測に単語単位を組み合わせることで, より柔軟な候補を提示するサジェストシステムを提案する. 提案するシステムでは, 図 1 のように単語の入力候補と文章の入力候補が提示される. 単語の候補は, 入力したい内容の中心となるキーワードを予測しており, 単語を選択すると, 選択した単語が含まれる文章候補が生成される. 単語候補と文章候補はそれぞれ視界画像を用いたキーワード予測と, 文章生成によって生成される. 単語候補から入力内容の中心となる単語を選択すると選択された単語が含まれる文章が候補として提示される.

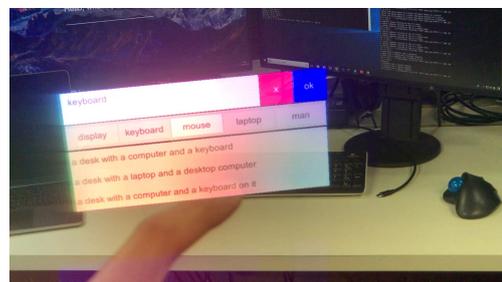


図 1: 提案システム

3.2 キーワード予測

ユーザーの意図の中心となるキーワード予測は視界情報を用いてを行う. 予測では視界画像に写っている物体を検出し, 検出された物体の名前をキーワー

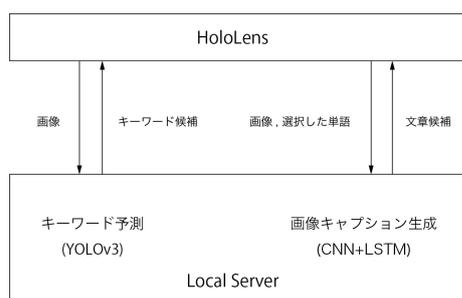


図 2: システム概要図

ドの候補として提示する。物体検出には YOLOv3[3] を用い、検出された個数が多い単語から順に 5 つ選んでキーワード候補とする。

3.3 キーワードを含んだ文章生成

文章生成は画像キャプション生成モデル [4] に視界画像を入力して行う。文章生成モデルは、文章の先頭から単語を選択して文章を生成するため、任意のキーワードを含むように文章を生成することはできない。そこで、一度生成された文章に単語を埋め込み、それ以降の文章を再生するという方法でキーワードを含んだ文章を生成した。単語の埋め込みは、生成された文章に含まれる単語のうち選択された単語と最も意味が近いものを入れ替えることで行う。

3.4 実装

システム全体の実装は図 2 のようになっている。スマートグラスでは視界画像の撮影とインターフェースの表示を行い、キーワード予測および文章生成はローカルサーバで行う。スマートグラスには Microsoft HoloLens を用いており、HoloLens で撮影された画像は HTTP 通信でローカルサーバに送られる。

4 ユーザー実験

4.1 概要

提案システムの有効性を調査するために 20 歳から 25 歳の男性 12 名に対してユーザー実験を行った。実験では、スマートグラスを用いて指定された物を見ながら自由に記述するという課題を仮想キーボードによる入力方法と提案手法による入力方法の 2 通りで課題を行ってもらった。なお、仮想キーボードとはディスプレイに表示されたキーボードである。

4.2 結果

主観評価の結果によると、75% の人がテキスト入力の総合的な負担は仮想キーボードと比べて提案システムの方が同程度あるいは小さいと回答した。提案システムによる入力が意図通りであったかという質問に対しては、半分ずつに意見が分かれた。実験

で入力された文章を比較した結果、提案システムを用いたテキスト入力では概ね意図通りの内容の文章を入力できていることが分かった。ただし、物の色や見えていない物については意図通りに入力できていなかった。物体検出以外の画像認識技術を合わせて使用したり、使用場面に応じたモデルの学習することで、実用性の向上が見込まれる。

5 おわりに

本研究では、スマートグラスにおけるテキスト入力への負担を軽減することを目的としたサジェストシステムを提案し、提案システムの実現に必要なキーワード予測と文章生成の 2 つの機能の実装を行い、ユーザー実験による主観評価を行った。

今後の課題としては、使用場面の検討、および場面に応じた予測によるサジェスト候補の向上、インターフェースの改善などが挙げられる。

謝辞

本研究の一部は TIS 株式会社の支援を受けたものである。

参考文献

- [1] On Suggesting Phrases vs. Predicting Words for Mobile Text Composition. *Proceedings of the 29th Annual Symposium on User Interface Software and Technology - UIST '16*, pp. 603–608, 2016.
- [2] A. Kannan, K. Kurach, S. Ravi, T. Kaufmann, A. Tomkins, B. Miklos, G. Corrado, L. Lukacs, M. Ganea, P. Young, and V. Ramavajjala. Smart Reply: Automated Response Suggestion for Email. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, pp. 955–964, New York, NY, USA, 2016. ACM.
- [3] J. Redmon and A. Farhadi. YOLOv3: An Incremental Improvement. *CoRR*, abs/1804.02767, 2018.
- [4] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and Tell: Lessons Learned from the 2015 MSCOCO Image Captioning Challenge. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(4):652–663, Apr. 2017.
- [5] 土屋 潤一郎, 入江 英嗣, 坂井 修一. シチュエーションに応じたテキストのインタラクティブ入力支援手法の評価. 協調とモバイルシンポジウム, p. 385–394, July 2018.
- [6] 鈴木孝宏, 美馬義亮. スマートフォン向けテキスト入力システム. 研究報告ユビキタスコンピューティングシステム (UBI), 2013(3):1–6, mar 2013.