

日常利用の拡張現実感環境における タッチタイピング可能な文字入力システム

川瀬 寛也^{*1} 澤邊 太志^{*1} 藤本 雄一郎^{*1} 神原 誠之^{*1} 加藤 博一^{*1}

Touch-Typable Text Entry Systems for Augmented Reality in Everyday Life

Hiroya Kawase^{*1}, Taishi Sawabe^{*1}, Yuichiro Fujimoto^{*1}, Masayuki Kanbara^{*1} and Hirokazu Kato^{*1}

Abstract – 本研究では、HMDを用いたAR環境においてスマートウォッチを入力装置としたタッチタイピング可能(手元のキーを見ること無く入力が可能)な文字入力システムを提案する。まず、本システムにおける仕様を決めるため、「キーレイアウト」「スマートウォッチが上下反転したときの挙動」「入力候補の選択方法」に関する3つの予備実験を行った。次に、日常生活で起こり得るシナリオでの利用を想定した姿勢で本システムの利用評価実験を行った。評価実験の結果、普段からQWERTY配列キーボードを利用し、日常的にスマートウォッチを身に付けている、AR-HMDに慣れたユーザを対象に日常利用における提案手法の有用性が確認された。

Keywords : 文字入力, タッチタイピング, 拡張現実感, スマートウォッチ, HMD

1 はじめに

拡張現実感(以下, AR)環境は、ヘッドマウントディスプレイ(以下, HMD)によっても実現されているが、現状では未だ日常利用までには至っていない。ハードウェア的な課題として、HMDの視野角や体積、重量があるが、これは現在研究及び開発が盛んに行われている[1]。一方で、ソフトウェア的課題の1つとして、日常利用が想定されていない文字入力インターフェースが挙げられる。HMDを用いた多くの3Dアプリケーションでは、英数字の入力が重要であり、この環境では従来とは異なる別のシステムが必要である[2]。これを解決する為、状況に制限されず快適に文字を入力出来るシステムを実装する必要がある。

HMDを用いたAR環境における既存の文字入力手法として、従来のハードウェアキーボードを使った入力手法、中空ジェスチャ、ポインティング操作や音声認識による文字入力などもあるが、これらの問題点としてユーザに不快感を与える可能性がある点や入力中に他の人の邪魔になる可能性もある点などが挙げられる[3]。上記の理由より、HMDを用いたAR環境における既存の文字入力手法は、どれも日常的に利用するためには課題がある。これらの問題を解決し、加えて入力速度と精度が保証されている新しい文字入力システムを実装する必要がある。

よって本研究では、スマートウォッチを入力装置としたタッチタイピング可能な文字入力システムを提案する。

2 関連研究

2.1 少数キーを用いた文字入力手法

少数キーを用いた文字入力には、大きく分けて2つの手法がある。1つ目のアプローチは、マルチタップ方式または非予測方式と呼ばれる方法である。これは、ユーザが複数のストロークを使用して入力する文字を一意に示すことで曖昧さを解消する方法であり、概念自体はタイプライターが発明された時から存在したものの、Smithら[4]が押しボタン式電話機でアルファベット入力を行う手法を提案したことをきっかけに改めて注目を集めた。

2つ目のアプローチは、予測方式と呼ばれる方法である。これは、システムが曖昧さを解消し、順序付けられた候補のリストからユーザが選択する方法であり、1970年代に押しボタン式電話機のキーパッドを使用した文字入力に適用された[5]。Tegic Communicationsが開発したT9[6]は携帯電話のキーパッドを用いて英語の入力が可能な技術であり、スマートフォンやスマートウォッチにおける文字入力での一般的なレイアウトとなっている。

2.2 AR環境における文字入力手法

Fashimpaurら[7]は、QWERTY配列キーボードのレイアウトに関するユーザーの既存の知識と、親指と他の指による簡単な操作を組み合わせた文字入力システムであるPinchTypeを提案した。入力速度は平均12.54WPM(Words per Minute)であり、エラー訂正の時間を除けば20.07WPMに到達している。しかし、この研究で用いられているハンドトラッキングシステムは、外部に設置されたカメラを用いたマーカーベース・トラッキングであり、現状のHMDでは単体

^{*1}奈良先端科学技術大学院大学

^{*1}Nara Institute of Science and Technology

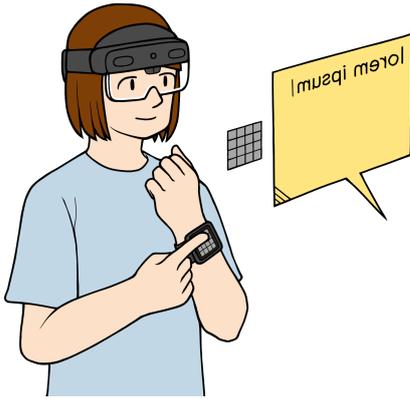


図1 AR環境におけるスマートウォッチを入力装置としたタッチタイピング可能な文字入力システム。

Fig.1 Touch-typable text entry system using a smartwatch for AR.

で動作させることが出来ない。Luら[8]は、ハンズフリーの不可視キーボードを利用した文字入力手法であるiTextを提案した。瞬きによるキー入力を利用した方法では、平均13.76 WPM、エラー率1.5%を達成している。

2.3 ウェアラブルデバイスを用いた文字入力手法

Xuら[9]は、親指の先端の微小なジェスチャーを用いた文字入力手法であるTipTextを提案した。これは親指の先で人差し指の先端をタップすることで、文字が入力出来る。入力速度は平均11.9 WPMであった。しかし、文字入力のための専用のデバイスが必要である。Ahnら[10]は、スマートウォッチの側面に3つのボタンを追加し、それを用いてキーボードを切り替えることで全てのアルファベットを入力可能にする文字入力手法を提案した。この手法の特徴は、スマートウォッチの側面にボタンを追加する必要があるものの、特別な外部入力デバイスを携帯する必要が無いことである。通常時の入力速度は9.08 WPMであった。歩行状態では7.66 WPMであった。

3 提案手法: 日常利用の拡張現実感環境におけるタッチタイピング可能な文字入力システム

本研究では、図1のように、AR環境においてスマートウォッチを入力装置としたタッチタイピング可能な文字入力システムを提案する。

3.1 デバイスと実装環境

ジャイロスコープを搭載し傾きが測定出来るスマートウォッチとしてApple Watch Series 7-45mmと、視線追跡が可能なAR-HMDとしてMicrosoft HoloLens 2を用いる。また、本研究では実装の簡単化の為、Apple WatchとHoloLensの双方向通信にiPhoneを介する

方法を利用する。Apple WatchはiPhone 12 Pro (iOS 15)とペアリングされており、Apple Watch-iPhone間の通信はWatch Connectivityによって行う。iPhone-HoloLens間の通信はOSCによって行う。

Apple Watch上のボタンを押してからその信号がHoloLensに送信されるまでの遅延をデバッグコンソールを用いて簡単に計測したところ、50ms程度であった。

3.2 キー選択手法

少数キーを用いた文字入力手法のうち、予測方式を採用する。予測方式は面積の限られた領域での文字入力の問題を対処するために採用され、キーボードをサポートするために必要なキーの数を減らすことを可能にする一般的な手法の為である。単語の曖昧性解消アルゴリズムは、American National Corpus(以下、ANC)[11]とMacKenzie Phrase Set[12]に基づいて実装した。まず、ANCからアルファベット文字を含まない単語を全て削除し、それを登場頻度でソートする。更に、上位からMacKenzie Phrase Set内の全ての単語を含むように取得した単語群を言語モデルとした。入力候補の選択肢には、キーが入力される度に対応する単語をその言語モデルからANCにおける使用頻度順に表示する。

また、アルファベットの入力の決定はスマートウォッチのディスプレイから指を離れたときに行われる。即ち、ユーザはキーの位置を記憶出来ない学習状態であっても、AR-HMDに表示されたキーパッドのプレビューを確認しながらディスプレイ上で指を滑らせるように動かし、正しいキーの位置で指をディスプレイから離すことで正確に入力を行うことが可能である。更に、スマートウォッチのディスプレイやキーパッドのプレビューを視覚により確認せずとも文字入力フェーズに正常に移行していることが触覚で認識出来るように、いずれかのキーが押下された又は押下しているキーが切り替わったタイミングでスマートウォッチを振動させるという実装を行った。

3.3 キー配列

本システムは手元やキーパッドを確認することなく入力することを想定している為、キーの位置を記憶し易い配列であることが望ましい。そこで、キー配列は既存のQWERTY配列をベースにしたものを採用する。これは、文字入力を行う際、ユーザは新たなレイアウトの学習をする意欲が無いと示されているためである[13]。

4 予備実験

本システムにおける仕様を決めるため、3つの予備実験を行った。それぞれの予備実験は、普段からQWERTY配列キーボードを利用し、スマートウォッ



図2 予備実験1で検証したキーレイアウト (左から3×3, 3×4, 4×3, 4×4, 5×4).

Fig. 2 Key layouts in exploratory experiment 1 (from left to right: 3×3, 3×4, 4×3, 4×4, 5×4).

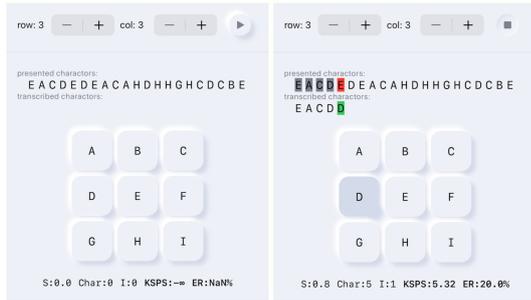


図3 予備実験1で用いた計測用アプリケーション (3×3レイアウト).

Fig. 3 Measurement application used in exploratory experiment 1 (3×3 layout).

チを身につけている, AR-HMDに熟練したユーザを想定し, 筆者 (年齢 24 歳, 身長 170cm, 男性, 右利き, 眼鏡有) によって行われた.

4.1 予備実験 1: キーレイアウトが入力速度とエラー率に与える影響

スマートウォッチのディスプレイ面積は限られているため, より多くのキーに分割すると入力の誤りが多くなり, 入力誤りを減らそうとするとキーストロークの速度低下が予想される. しかしキー数が増えれば1つのキーに含まれる文字数を少なく出来, 入力操作に対応する単語数が減るため, 文字入力速度は向上する. そこで, キーの数やレイアウトを変えてキーレイアウトを決定することを目的とした予備実験を行った.

4.1.1 設計及び実装

図2のような記号の入力が可能な複数のボタンを持つキーパッドを Apple Watch へ実装した. 簡略のため, キーパッドのプレビューや入力文字列及び計測データ等の情報を表示するためのデバイスとして, iPhone を用いた. iPhone の画面上部に図3が表示される. 実験では, 3×3, 3×4, 4×3, 4×4, 5×4 の5つのキーレイアウトについて検証した.

4.1.2 実験手順

図3に示すように, ランダムなアルファベットが表示された文字列を転写するタスクを行った. 椅子に座り, 腕を机の上に置いた姿勢で実験は行われた. 目線の高さ, 50cm 先にスマートフォンを設置し, それに

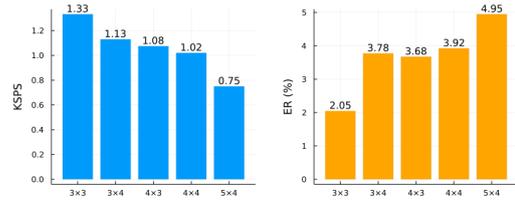


図4 予備実験1の結果: それぞれのレイアウトでの KSPS と ER.

Fig. 4 Results of exploratory experiment 1: KSPS and ER in each layout.

表示されたキーパッドのプレビューを確認しながら, それぞれのキーレイアウトに対して3分間の入力を行った. 21個のアルファベットを入力する毎にその文字列は更新され, 誤った入力に対する削除・修正は出来ない.

4.1.3 評価指標

記号の入力を評価する指標として, 既存の評価指標である KSPS (Keystrokes Per Second)[14] と ER (Error Rate)[15] を採用した.

KSPS は式 (1) のように表せる.

$$KSPS = \frac{|IS| - 1}{S} \quad (1)$$

ここで, IS は入力ストリーム (Input Stream) であり, $|IS|$ はその数である. IS には入力操作における全てのキー入力が含まれる. S は秒数であり, 最初の文字の入力から最後の文字の入力までの時間である. KSPS を使用すると, 入力された全ての文字が正しいと仮定することにより, 経験的な上限を推定出来る.

ER は式 (2) のように表せる.

$$ER = \frac{INF}{|T|} \times 100\% \quad (2)$$

ここで, INF は入力された文字のうち修正されなかった誤字 (Incorrect Not Fixed) の数である. T は最終的に入力された文字列 (Transcribed text) であり, $|T|$ はその数である.

4.1.4 実験結果と考察

結果は図4のようになった. キー数が多いほど KSPS は小さく, ER は大きいことが確認出来る. 3×4, 4×3, 4×4 はレイアウトの違いによる KSPS と ER の差が他のレイアウトと比較して小さいことが確認出来る. また, 5×4 では, KSPS が1を下回っている.

前述したように, キーの数を増やすことで1つのキーに含まれる文字数が少なくなるため, パフォーマンスの低下率が著しくない限りはキー数が多いほうを採用すべきである. 以上のことから, 以降の実験では 4×4 レイアウトを採用する.



図5 予備実験2で検証した2種類のパターン。
Fig.5 Two patterns verified in exploratory experiment 2.

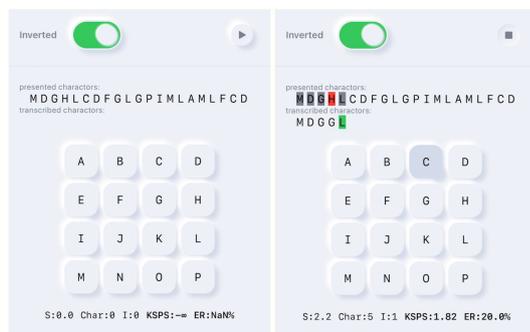


図6 予備実験2で用いた計測用アプリケーション。
Fig.6 Measurement application used in exploratory experiment 2.

4.2 予備実験2: スマートウォッチが上下反転したときの挙動を決定する

日常生活で取り得る姿勢には、手のひらが自身の方に向けてスマートウォッチが上下反転した状態になることもある。このとき、キーパッドはそのままである(パターン1)か垂直方向に反転する(パターン2)かの2つの挙動が考えられる。パターン1はスマートウォッチを装着している腕に対してキーパッドが常に同じ方向を向いているのに対して、パターン2はキーパッドが鉛直方向に対して同じ方向を向く。本実験では、パターン1とパターン2のどちらが本システムに適しているかを評価し、挙動を決定する。

4.2.1 設計及び実装

図5のような4×4のキーパッドをApple Watchへ実装した。予備実験1と同様に、キーパッドのプレビューや入力文字列、計測データ等の情報表示をiPhoneへ実装した。パターン1とパターン2はiPhone上に実装されたトグルボタンによって切り替えを行う。

4.2.2 実験手順

図6に示すように、ランダムなアルファベットが表示された文字列を転写するタスクを行った。椅子に座

り、スマートウォッチを鳩尾の前に構え外側に向け、スマートウォッチの0時方向から右手で入力するような姿勢で実験を行った。目線の高さ、50cm先にスマートフォンを設置し、それに表示されたキーパッドのプレビューを確認しながら、それぞれのキーレイアウトに対して3分間の入力を行った。

4.2.3 評価指標

予備実験1と同様、 $KSPS$ (式(1))と ER (式(2))を使用する。

4.2.4 実験結果と考察

パターン1では、入力速度は0.77KSPSで、 ER は12.9%であった。一方、パターン2では入力速度は0.64KSPSで、 ER は19.7%であり、どちらもパターン1より悪い結果となった。これは、パターン2はスマートウォッチが完全に上下反転した場合にはメンタルモデルと合致するためこの挙動が有効であるものの、実際の入力時には方向がまちまちであることが多く、スマートウォッチに対して常に同じ方向を向いているほうがキーの位置を記憶しやすいからであると考えられる。以上からスマートウォッチが上下反転してもキーパッドはそのままの挙動を採用する。

4.3 予備実験3: 入力候補の選択方法を決定する

予測方式を用いた文字入力では入力したい単語を確定するために入力候補から選択する動作が必要となる。Apple Watchにはアプリ操作のためにDigital Crown(竜頭型ボタン)が搭載されており、同様の機能を持つボタンは他のスマートウォッチにも搭載されている。また、HoloLensは視線追跡が可能で、それを用いて画面上に表示された選択肢を選ぶことができる。入力候補の選択方法として、スマートウォッチのDigital Crownを利用するもの(方法1)とAR-HMDの視線追跡を利用するもの(方法2)を実装し、そのどちらが本システムに適しているかを評価した。

4.3.1 設計及び実装

予備実験1からレイアウトは4×4と決定された。削除や確定、空白キー等の特殊キーを最下段に置き、句読点キーを用意する場合、残りは11キーである。図7(b)に示すような言語モデル内での各アルファベットの登場頻度を考慮すると、図7(a)のレイアウトが適切であると考えた。このレイアウトで本システムの曖昧性解消アルゴリズムを用いると、言語モデル内の任意の単語に対して92.34%の単語が入力候補の1番目に現れ、入力候補の6番目までには全ての単語を入力出来ることを確認した。これをApple Watchに実装し、HoloLensにキーパッドのプレビューや入力文字列の表示の実装を行った。また、Apple WatchとHoloLensを双方向に接続するiPhoneにも図8のような実装を行った。選択候補の切り替えはiPhoneから

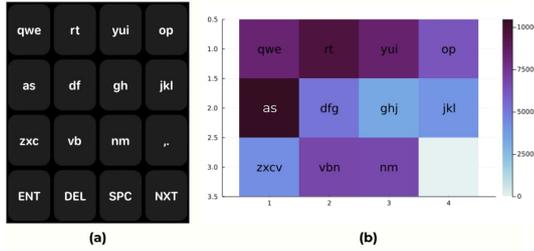


図7 (a) 言語モデル内の各アルファベットの登場頻度から決定されたレイアウト。
(b) ANC 上位 10,000 語中のアルファベットの登場頻度。
Fig. 7 (a) Layout determined from the frequency of alphabets in the language model.
(b) Frequency of alphabets in the top 10,000 ANC words.

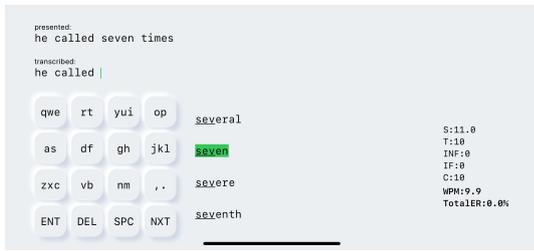


図8 予備実験3で用いた計測用アプリケーション。
Fig. 8 Measurement application used in experimental experiment 3.

行う。

4.3.2 実験手順

転写するフレーズは MacKenzie Phrase Set[12] からランダムに選択したものをを用いた。椅子に座り、腕を机の上に置いた姿勢で実験は行われた。目線の高さ、約 50 cm 先に表示された HoloLens2 のキーパッドのプレビューを確認しながら、それぞれの選択方法で3分間の入力を行った。

4.3.3 文字入力に用いる評価指標

文字の入力を評価する指標として、既存の評価指標である WPM(Words Per Minute) と TotalER(Total Error Rate)[14] を採用した。

WPM は式 (3) のように表せる。

$$WPM = \frac{|T| - 1}{S} \times 60 \times \frac{1}{5} \quad (3)$$

T 及び S は式 (1), 式 (2) と同様である。

TotalER は式 (4) のように表せる。

$$TotalER = \frac{INF + IF}{C + INF + IF} \times 100\% \quad (4)$$

ここで, INF は式 (2) と同様であるが, その値は P を入力すべき文字列 (Presented text) としたとき, T を

P へ変換するのに必要な最小の操作回数 MSD(P, T) と同値である (Minimum String Distance)[16]. IF は, 入力中に削除キーを入力した数 (Incorrect Fixed) である. C は, T のうち正しく入力した文字の数 (Correct) であり, $\max(|P|, |T|) - MSD(P, T)$ と同値である. $\max(\cdot, \cdot)$ は 2 つの引数のうち大きな値を返す。

4.3.4 実験結果と考察

方法 1 では入力速度は 12.9 WPM であり, Total ER は 3.2% であった. 方法 2 では入力速度は 15.5 WPM であり, TotalER は 0.9% であった. 単語の入力途中, 視線追跡では目的の単語が現れた時にその単語の確定が簡単に行えるのに対して, Digital Crown では目的の単語が現れた時にその単語の確定のため距離に応じて操作する必要がある. そのため Digital Crown を用いることなく 1 番目に入力候補が現れたときのほうが結果的に早く入力出来ることがあり, これが Digital Crown の WPM と TotalER の悪化の原因であると考えられる. 以上から, 入力候補の選択には視線追跡を用いた方法を採用する。

5 日常生活シナリオでの利用評価実験

5.1 日常生活におけるシナリオ

スマートウォッチに対する腕の方向は 0 時, 3 時, 6 時方向があり, 姿勢は仰向け, 直立, 着席状態が考えられる. 日常生活で自然に取り得る姿勢は様々存在するが, これらを網羅するシナリオとして以下の 4 つで評価する。

シナリオ 1 仰向けの姿勢で, 両手を胴体の上に置いて入力する. 6 時方向から操作する。

シナリオ 2 直立の姿勢で, 両手を腹部の前に置いて入力する. 6 時方向から操作する。

シナリオ 3 椅子に座った姿勢で, 机に両肘を付き両手が鳩尾辺りの高さになる位置で入力する. 0 時方向から操作する。

シナリオ 4 直立の姿勢で, 両手を鳩尾に置いて入力する. 3 時方向から操作する。

それぞれは図 9 にあるような実環境におけるシナリオを想定している。

5.2 実験概要

実験機器, 実験手順 (それぞれのシナリオを模した姿勢で, 3 分間の入力), 評価指標は予備実験 3 と同様である。

5.3 実験結果と考察

図 10 に実験結果を示す. シナリオ 3,4 はスマートウォッチが上下反転した状態であるため, 他と比較し

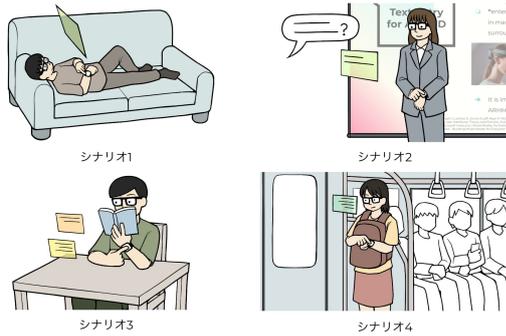


図9 本実験で評価する4つのシナリオ。

Fig. 9 Four scenarios evaluated in this experiment.

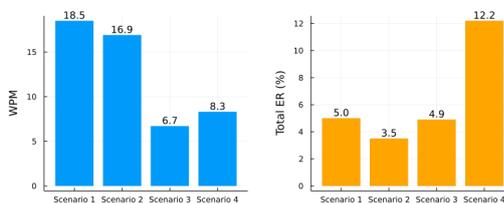


図10 シナリオ毎のWPM及びtotalER。

Fig. 10 WPM and totalER for each scenario.

てWPMが小さい。フィーチャーフォンのT9入力では、初心者は約9WPM、熟練者は約20WPMでテキストを入力出来る[17]ことから、携帯電話に迫る入力速度で文字入力が可能であると言える。日常利用における提案手法の有用性が確認された。

6 まとめ

本研究では、HMDを用いたAR環境においてスマートウォッチを入力装置としたタッチタイピング可能な文字入力システムを提案するために、「キーレイアウト」、「スマートウォッチが上下反転したときの挙動」、「入力候補の選択方法」を決定する3つの予備実験を行った。その後、日常生活で起こり得るシナリオを模した姿勢で本システムの利用評価実験を行った。評価実験の結果、普段からQWERTY配列キーボードを利用し、日常的にスマートウォッチを身に付けている、AR-HMDに慣れたユーザを対象に日常利用における提案手法の有用性が確認された。

今後の課題として、キーパッドのプレビューの改善、複数人による長期的な実験、そして数字や記号などの特殊文字の入力機能の追加などが挙げられる。

参考文献

[1] Ye, Bi, et al. "Cross-talk elimination for lenslet array near eye display based on eye-gaze tracking." *Optics Express* 30.10. (2022): 16196-16216.

- [2] Bowman, Doug, et al. "3D User interfaces: theory and practice." Addison-Wesley, (2004).
- [3] Hsieh, Yi-Ta, et al. "Designing a willing-to-use-in-public hand gestural interaction technique for smart glasses." *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. (2016).
- [4] Smith, Sidney L., and Nancy C. Goodwin. "Alphabetic data entry via the Touch-Tone pad: A comment." (1971): 189-190.
- [5] Desautels, Edouard J., and Stuart B. Soffer. "Touch-tone input techniques: Data entry using a constrained keyboard." *Proceedings of the 1974 annual conference-Volume 1*. (1974).
- [6] Grover, Dale L. "Reduced keyboard disambiguating computer." Patent no. us5818437 (1998).
- [7] Fashimpaur, Jacqui, Kenrick Kin, and Matt Longest. "Pinchtype: Text entry for virtual and augmented reality using comfortable thumb to fingertip pinches." *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. (2020).
- [8] Lu, Xueshi, et al. "iText: Hands-free text entry on an imaginary keyboard for augmented reality systems." *The 34th Annual ACM Symposium on User Interface Software and Technology*. (2021).
- [9] Xu, Zheer, et al. "Tiptext: Eyes-free text entry on a fingertip keyboard." *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. (2019).
- [10] Ahn, Sunggeun, Seongkook Heo, and Geehyuk Lee. "Typing on a smartwatch for smart glasses." *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*. (2017).
- [11] Macleod, Catherine, Nancy Ide and Ralph Grishman, "The American National Corpus: A Standardized Resource for American English." LREC, (2000).
- [12] MacKenzie, I. S. and Soukoreff, R. W. "Phrase sets for evaluating text entry techniques." *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems* (2003): 754-755.
- [13] MacKenzie, I. Scott, Shawn X. Zhang, and R. William Soukoreff. "Text entry using soft keyboards." *Behaviour & information technology* 18.4 (1999): 235-244.
- [14] MacKenzie, I. Scott, and Kumiko Tanaka-Ishii. "Text entry systems: Mobility, accessibility, universality." Elsevier, (2010).
- [15] Arif, Ahmed Sabbir, and Wolfgang Stuerzlinger. "Analysis of text entry performance metrics." *2009 IEEE Toronto International Conference Science and Technology for Humanity*. IEEE, (2009).
- [16] Soukoreff, R. William, and I. Scott MacKenzie. "Measuring errors in text entry tasks: An application of the Levenshtein string distance statistic." *CHI'01 extended abstracts on Human factors in computing systems*. (2001).
- [17] James, Christina L., and Kelly M. Reischel. "Text input for mobile devices: comparing model prediction to actual performance." *Proceedings of the SIGCHI conference on Human factors in computing systems*. (2001).