

対話型生成 AI を用いた接客システムの文字数制限による反応時間の検証

福知山公立大学情報学部 32045078 平木健太郎
指導教員 山本吉伸

1. はじめに

対話型生成 AI が店頭に立ち人間の代わりに接客を行うことがある。一般的な対話型生成 AI が行う接客は店舗やレジに設置されたディスプレイに映る 2D キャラクターが行うことが多く、対話型生成 AI の接客はお客様の発話を待つ受動的なものが多い。しかし、受動的な接客だけでなく客引きなどの店員から能動的に行う接客も重要なタスクであり、対話型生成 AI が人間に代わり接客を行うためには必要不可欠なタスクとなる。この能動的な接客を対話型生成 AI が行う場合は以下の問題があげられる。(1)お客様が入店した時に対話型生成 AI が単純に話しかけてお客様が自身に話しかけられたか判断できず無視される。(2)対話型生成 AI による返答が遅い場合はお客様に悪い印象を与えてしまう。

そこで本研究では店舗に設置した対話型生成 AI 「AI ふうこ」の付近にいるお客様の顔画像を切り出し対話するシステムの開発を行った。本システムはディスプレイに表示した 2D キャラクター「ふうこ」が付近にいる人間の顔を認識し画像として切り出す。この切り出した画像をディスプレイ上に表示し発話を始めるシステムとなっている。本稿では AI ふうこの発言に文字数制限を設定し、対話型生成 AI による返答までに生じる時間の差について実験を行った。

2. 実験

2. 1. 実験環境

本システムは「ふうこカフェ」に導入することを目的とし開発を行った。また、AI が発言する文字数を制限することで返答までにかかる時間に差があるため、本実験では 10、20、30、50 文字以内の 4 種類で実験を行う。

2. 2. 実験手順

本実験の手順は以下の通りである。

(1)AI ふうこ付近にお客様が近づくと顔を認識し、発話を行う。(2)お客様は AI ふうこの発話に対して「こんにちは」と言う。(3)AI ふうこはお客様の音声の認識直後「はい」と発話し、返答を考える。(4)お客様は AI ふうこからの返答を待ち、返答後またあらかじめ決めた発言を行う。(5)(3)～(4)を繰り返す。(6)(4)でお客様からの発言が 20 秒間ない場合は対話終了する。手順(4)でお客様が発言する言葉は順に「メニューは何がありますか。」、「コーヒーをください。」、「ありがとうございました。」とする。

2. 3. 収集データ

実験では各文字数に以下の 3 つのデータを収集した。(1)お客様の発言後の音声認識までの時間(2)認識した音声に対する返答文章を対話型生成 AI の ChatGPT から受け取るまでの時間(3)ChatGPT の返答文章を音声ファイルに変換し、再生するまでの時間

3. 3. 実験結果

実験結果より収集データ(1)は安定して 2 秒以内であるが、単純に文字数の上限が増えると収集データ(2)と収集データ(3)の時間が延びることが分かった。また、これと同時に文字数の上限が増えると一度に伝えられる情報が増えることが分かった。図 1 から文字数制限が増えると平均的に収集データ(2)の時間が長くなっていることがわかる。

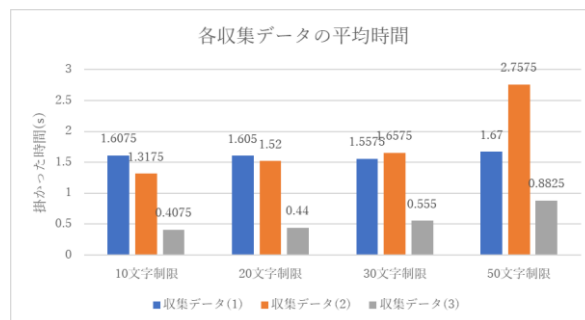


図 1 各文字数制限の収集データの平均時間

4. 評価

実験結果より、収集データ(1)は安定して 2 秒以内であることが示された。志和らの研究[1]ではコンピュータの反応時間は 1 秒が最も効果が高いと示されているため改善する余地がある。

収集データ(2)(3)の合計は文字数制限が増えると比例して長くなることが示された。どの文字数制限でも 1 秒以上の時間が掛かるため改善する余地がある。図 1 より収集データ(2)の時間の増大が大きいことが分かった。

返答内容は文字数制限が増えるほど一度に伝える情報が増えることが示唆された。AI ふうこが実際に接客を行う際には環境や役割に合わせて文字数制限を決める必要がある。

参考文献

[1] 志和敏之, 神田崇行, 今井倫太, 石黒浩, 萩田紀博, 安西裕一郎, “対話ロボットの反応時間と反応遅延時間における間投詞の効果”, 日本ロボット学会誌, Vol. 27, No. 1, pp. 87~95, 2009.