

Android 端末を用いた音声認識による照明システムのユーザインタフェース

水上 雅博

1 はじめに

我々はオフィス環境を改善する研究の一環としてユーザの要求に応じた明るさを提供する知的照明システムの研究を行っている¹⁾。現在、知的照明システムは Web ユーザインタフェース (Web User Interface : Web UI) を通して各執務者の個人 PC から操作を行っている。このユーザインタフェース (User Interface : UI) では、照明を操作するために PC を起動しなければならないという問題があった。これを解決するため、Android 端末を通して照明を操作できる、新たな UI を開発する。本研究では Android 端末を用いた音声認識による知的照明システムの UI を提案し、その有用性について検討する。

2 知的照明システム

知的照明システムとは、それぞれの照明器具の協調動作によってユーザの要求する照度を満たし、かつ省エネルギーを実現するシステムである。現在、我々は知的照明システムの実用化に向けて東京の新丸の内ビルをはじめとした多数のオフィスへ導入を行い、検証実験を行っている。検証実験においても、知的照明システムはユーザが要求する照度に合わせて制御が行われ、省エネルギーの観点から従来の照明の消費電力と比較して 50% 程度の削減が可能であるという良好な結果を出している²⁾。

しかしながら、知的照明システムを利用するには PC を用いる必要があり、照明操作のために執務者が PC を起動しなければならないという点が問題となっている。

3 UI

3.1 知的照明システムにおける UI

まず、知的システムの操作に用いる UI については、誰にでも見やすく使いやすい直接的コントロールが可能な UI が求められており³⁾、さらに、PC のみでなく様々な端末を用いて知的照明システムを利用可能であることが望ましい。

特に、Android 端末などのマイクやタッチパネルといった様々な種類のインタフェースを利用可能な端末においては、それらインタフェースの特徴に合わせた UI をそれぞれ実装していることが望まれる。本研究では、Android 端末の音声認識機能を生かした知的照明システムの音声認識 UI の導入について検討する。

3.2 知的照明システムへの音声認識の導入

今までオフィスにおいて音声認識 UI が取り入れられてなかった理由が大きく二つ存在する。一つは、照明を操作するためにマイクを持つことが現実的でなかったことが挙げられる。もう一つは、マイクを持たずに音声認

識を行えるよう、オフィスの様々な場所にマイクを設置するという手法が認識率を確保するために大量のマイク必要とし、また、マイクの設置数を少なくした場合は大幅に認識率を下げってしまうという問題を解決できなかったことである。その点、Android 端末を用いることは音声認識の認識率、オフィス環境における適切さと汎用性を解決する手法として効果的なものの一つであると言える。

以下に知的照明システムの音声認識 UI と知的照明システムの構成を図 1 に示す。

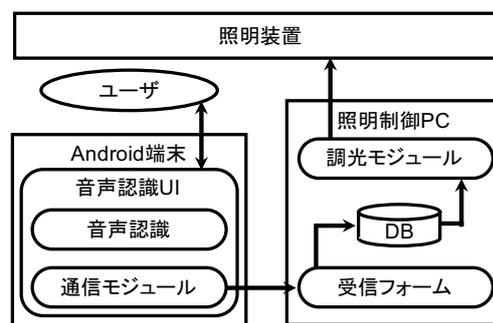


Fig.1 音声認識 UI と知的照明システムの関連

音声認識 UI は従来の知的照明システムに対してのアドオン (拡張機能) 形式で提供する。知的照明システムの照明制御 PC に Android 端末からの情報を受信し、データベースに書き込むためのフォームを追加する事で従来の UI と音声認識 UI を併用することが可能となる。

4 音声認識 UI の実装

4.1 Android 端末における音声認識の問題

音声認識は Android 1.5 以降に標準搭載されている音声認識フレームワークである Recognizer Intent を利用する。この Recognizer Intent は Android 1.5 以降の Android OS が搭載された端末ならば利用することが可能なフレームワークであり、端末に搭載されているマイクを通して音声を取得し、サーバ上のデータをもとに音声認識を行うことができる。このフレームワークを用いる際の問題として、音声認識結果の候補が複数与えられるという問題がある。音声認識 UI では、ユーザの意図した照明の操作を示す文章を複数の認識結果の中から選び出さなければならない。

4.2 音声認識の問題の解決

音声認識によって得られた複数の結果の中から、適切な認識結果を採用しなければならない。その採用の手法として照明の操作に関連する語句ごとに得点を与え、その得点の総計から尤度を計算する得点評価法を実装した。この尤度の計算に、文法解析などを用いた意味解析など

を用居なかった理由は、Android 上で茶筌などの形態素解析を行う事や、大量のコーパスを取り扱う事は性能上難しいと考えたためである。

得点評価法は照明の操作に関する形態素ごとに得点を定める。特に操作に重要な形態素には高い得点を、操作に関連しているが操作に関係ない場合でも登場する語には低い得点をつける。その形態素と得点の関係を 1 とし以下に示す。

Table1 得点評価法における得点の割当表

形態素	得点	形態素	得点
全	5p	点灯	10p
前	5p	消灯	10p
後	5p	つけ	10p
右	5p	消し	10p
左	5p		

上記の Table 1 に従って音声認識の結果に得点を与え、その中で最も得点の高かった語を照明操作を表す結果として採用する。

4.3 Android 端末からサーバへのデータ送信の実装

Android 端末とサーバ間の通信には HTTP 通信における POST メソッドを用い、Android 端末から照明制御 PC に「どのような照明の操作を行うか」という命令をエンティティとしてリクエストを行う*1。

5 実験と考察

情報機器の取り扱いに慣れていない学生 5 名 (年齢 21~23 歳, 全員男性) を被験者として実験を行う。実験手順としては、音声認識 UI の操作方法について説明を行い、被験者に Android 端末 (MOTOROLA XOOM, TBi11M) を渡す。そして、全点灯や全消灯、ある一部のみ点灯、消灯などの操作命令を行ってもらい、実際にその操作命令が認識されたどうか確認する。18 種類の操作を 1 セットとし、これを 2 回繰り返してもらった。以下に各操作の 1 回目および 2 回目の認識率をそれぞれ Table 2 として示す。

Table2 音声認識の実験結果

操作命令	認識率 (1 回目)	認識率 (2 回目)	認識率 (合計)
点灯	100 %	100 %	100 %
消灯	100 %	100 %	100 %
位置指定 1 語+点灯	90 %	95 %	92.5 %
位置指定 1 語+消灯	45 %	65 %	55 %
位置指定 2 語+点灯	90 %	70 %	80 %
場所指定 2 語+消灯	55 %	65 %	60 %

*1 POST メソッドによってリクエストが送られる URL
[http://\(IP アドレス:ポート番号\)/form.php](http://(IP アドレス:ポート番号)/form.php)
 また、ここにはエンティティボディとして、全点灯の場合は下記のようなエンティティが付加される
 ./form.php?cmd=ALLON

Table 2 より得られる結論は大きく 3 つある。1 つ目に、語数の増加に伴って認識率は減少する傾向にあること。これは今後、より多くの操作命令に対応していく上で、一度の操作命令に対して、どの程度の語数を含むと操作に支障が出るかを認識率および操作者の両方の視点から検証する必要がある。また、認識する語句の数と認識率の関係についても検証を行っていく。

2 つ目に、「点灯」という語に比べて「消灯」という語の認識率が著しく低いことがあげられる。「消灯」という操作命令がどのような語に誤認識されたのかを調査したところ、「消して」を「決して」や「で指定」、「消灯」を「消防」や「食堂」、「しょうたろう」などに誤認識していた。これを改善するためには、音声認識の際に用いるコーパスを修正する方法が最も効率的であるが、今回は Android 標準の音声認識フレームワークを利用しているため、コーパスは修正できない。そのため、誤認識を修正するための辞書データなどを用意し、誤認識された結果を修正する形となる。また、音声認識の結果から尤度の高いものを採用するために策定した得点評価法についても、形態素と得点の対応を見直し、認識率の向上を図る。

3 つ目に、音声認識の認識率についてである。現在広く普及している音声認識ソフトウェアとして Julius が挙げられるが、この音声認識ソフトウェアの認識率はおよそ 80~95% であった⁴⁾、対して今回の音声認識 UI の認識率は 55~100% と低い。しかしながら、2 つ目の結論として挙げた「消灯」という語の認識率が「点灯」という語の認識率と同様の数値に向上すれば、Julius と同程度の 80~100% という認識率を得ることが出来ると考えられる。

以上より、Android 端末を用いた音声認識 UI については改善点が残るものの、その性能の基礎評価としては UI として問題なく、今後は更なる音声認識フレームワークの性能についての検証実験と、操作者側の目線に立った主観的な音声認識 UI の操作性、有用性について検証を行い、利便性の向上を図る。

参考文献

- 1) 三木光範. 知的照明システムと知的オフィス環境コンソーシアム. 人工知能学会誌, Vol.22, No.3, pp.399-410, 2007.
- 2) 三木光範, 加來史也, 廣安知之, 吉見真聡, 田中慎吾, 谷澤淳一, 西本龍生. 実オフィス環境における任意の場所にユーザが要求する照度を提供する知的照明システムの構築. 電子情報通信学会論文誌, Vol.J94-D, pp.637-645, 2011.
- 3) 三木光範, 廣安知之, 池田聡. 知的照明システムにおけるユーザインタフェースの構築 (音声認識およびタッチパネルを用いた照明コントロール). 第 67 回情報処理学会, 2005.
- 4) 1999 年度 大語彙連続音声認識 評価実験結果.
<http://winnie.kuis.kyoto-u.ac.jp/pub/julius/result99/tmp.html>.