

指差しと口述説明を理解する幾何学問題学習支援システム Learning Support System for Geometric Problem Solving Using Finger Pointing and Speech

清原侃太[†] 西村良太[†] 北岡教英[†]
Kanta Kiyohara Ryota Nishimura Norihide Kitaoka

1. はじめに

人は様々なモダリティを用いて、自分の意図を伝えようとする。例えば、音声や指差し、視線などが挙げられる。このようなモダリティを複数組み合わせ合わせたマルチモーダルインタラクションはユニモーダルなインタラクションと比較して、タスク効率の向上やエラー率の低下が期待されている[1]。

また近年、ICT の活用による教育、医療、介護・福祉などの社会分野への貢献が期待されており、ICT デバイスの普及と処理能力の向上に伴い、マルチモーダルインタラクションの導入が容易になっている。

そこで本研究では、ICT を利用した教育の一環として、音声および指差し操作によるマルチモーダルインタラクションを用いた幾何学問題解答アプリケーションの開発を行う。幾何学問題の解答の説明は、音声と指差しを用いたマルチモダリティの典型である。また、我々は、人のより自然な動作でかつ汎用的な解答が行えるようにシステムの構築を目指している。

2. 関連研究

Bolt の “Put-That-There” [2] から始まり、音声とジェスチャーを組み合わせ合わせたマルチモーダルインタラクションに関する研究は数多くある。Miki らによる研究[3]では、音声認識および位置センサーを用いたジェスチャー認識によって幾何学問題を解答するシステムを紹介している。

しかし、音声とジェスチャーやタッチ操作などのポインティング操作を組み合わせ合わせたマルチモーダルインタラクションの実現において課題となるのは、それらモダリティの統合理解である。人は「そこ」と言われても「そこ」がどこなのかおおよそ理解できるが、コンピュータは「そこ」と入力されただけでは「そこ」がどこなのか全く理解できない。そこで、指差しによるジェスチャーなどで認識を補助することでコンピュータは「そこ」がどこなのかを理解させることができる。例えば、Bolt による研究では、音声のみでは理解できないスクリーン上の位置（「そこ」）を指差しによって認識の補助をすることで理解させている。このように、音声とポインティングを組み合わせ合わせたマルチモーダルインターフェースを構築するためには、それらのモダリティをコンピュータに統合理解させることが必要不可欠である。

3. アプリケーションの構築手法

本章では、本アプリケーションの構築手法について述べるが、その前に、本研究で考慮している 2 つの事項について述べる。1 つは、主に中学生向けの簡単な幾何学問題を取り扱っていること。もう 1 つは、技術職ではない教員、

つまりプログラミングやコーディング等の専門知識を必要とせずに問題作成ができること。これらの 2 つの事項を考慮した上で簡単な幾何学問題に対して汎用性を持ちつつ、特別な技術を必要とせずにシステムに入力できることが望ましいと考えている。

また、本研究で構築するマルチモーダルインターフェースは、ICT 教育向けの一般に使用できるアプリケーションを想定している。そのため、前述の Bolt や Miki らの研究のように特殊なデバイス（グローブや位置センサーなど）からの入力を使用せずに、幾何学問題を解答できるようにシステムを構築した。

3.1. 幾何学問題作成手法

上記で述べたように、幾何学問題を特別な技術を必要とせずに作成してシステムに入力するために、我々が注目したのは、「GeoGebra」[4]である。「GeoGebra」は動的数学ソフトウェアであり、その特徴として、操作が容易であることや作図したオブジェクトの数式が「動的に」表示されることなどが挙げられる。この「GeoGebra」を使用して幾何学問題を作成し、アプリケーション内で描画処理をするために「Asymptote」形式のテキストファイルへと変換する（「Asymptote」形式への変換は、「GeoGebra」が持つ機能をそのまま使用している）。変換したテキストファイルを本アプリケーションで読み込み、「Asymptote」形式に則り、問題の図の描画などの処理をしていく。このようにすることで、一般的な学習アプリケーションでは、設計者または開発者が事前に作成した問題にのみ解答することができるが、本アプリケーションでは、「GeoGebra」を介して教員などのユーザが作成した独自の問題を使用することができる。

3.2. マルチモーダル操作

本アプリケーションでは、「音声」およびタッチパネルによる「指差し」の 2 つのモダリティを入力として使用する。

音声認識には Julius[5]を用いる。音声認識では、指示語や角度、数式（「この角」、「角 ABC」、「=」）などの認識を行い、後述の指差し操作と共に統合理解しながら認識結果に応じた処理を行っていく。

指差し操作では、指示語などの音声入力をコンピュータに理解させるための認識補助や、「補助線を引く」などのユーザ操作によって行われる描画処理などを行う。指差し操作について、例えば、Bolt の研究では、3D 空間における指差しのジェスチャー認識によって音声認識の補助を行っているが、触覚的なフィードバックが、エラー率を減少させるという研究報告もある（[6],[7]）。そのため、本研究では 3D 空間におけるジェスチャー認識よりも触覚的なフィードバックが得られやすいタッチパネルによるタッチ操作によって入力を行う。また、現在のタッチスクリーンは、キーボードのキーを押すことに比べて、触覚的なフィードバックを提供しない。したがって、触覚的なフィード

[†] 徳島大学

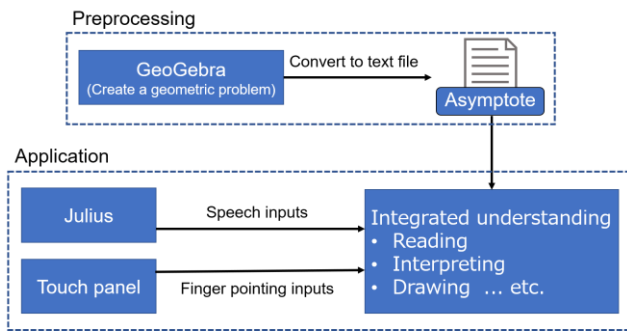


図 1. システム構成図

バックに限らず、適切なフィードバック（例えば、視覚や音など）の使用が、タッチスクリーンインタフェースを設計する際に不可欠であるとされている[8]。そのため、ユーザが音声や指差しによる入力を行った際に、短時間の音（SE）または視覚的に効果のあるフィードバック（例えば、色情報）を追加することも考えている。

4. システム構成

本研究で構築したシステム構成を図1に示す。まず、GeoGebraを使用して問題作成者（例えば、教員など）が図2のように問題を作成し、GeoGebraが持つ機能を使用してAsymptote形式へと変換する。その後、変換されたテキストファイルをアプリケーションに読み込ませることで問題を描画する。この時点で解答するための準備が整ったため、解答者（例えば、生徒など）に問題を音声と指差しによって解答してもらい、それらの入力をアプリケーション内で統合理解し、反映することで解答を行っていく。図2で示した幾何学問題を読み込み、解答を行っている際のアプリケーション・インターフェースを図3に示す。画面の上部には読み込んだ幾何学問題が表示される。右側のテキストボックスには、解答者の音声入力による認識結果（例えば、“この角とこの角が等しい”）が表示される。左側のテキストボックスには、解答者の音声入力と指差し入力がシステム内部で統合理解された結果（例えば、“ $\angle ADC = \angle ABC$ ”）が表示される。

5. おわりに

本稿では、音声と指差しを用いて幾何学問題を解答できるマルチモーダルシステムについて述べた。音声と指差しの入力をコンピュータに統合理解させることで、問題内の角度の認識や補助線を引くなどの操作を行えるようにした。構築したアプリケーションは、ICT教育への導入に向けてより実践的な問題を解答できるようにシステムの改善を目指している。

今後の課題として、多種の幾何学問題への対応と複数の解答方法を行えるように、より汎用的なシステムの構築を考えている。そのため、実際に本アプリケーションを使用している評価実験とそのデータの収集が必要であると考

参考文献

- [1] Oviatt, S. L., “Mutual Disambiguation of Recognition Errors in a Multimodal Architecture”, Proc. of the ACM

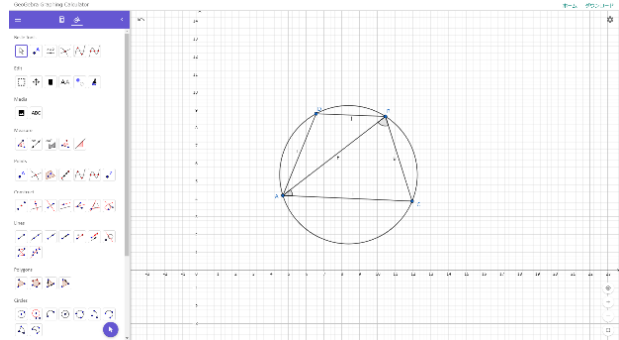


図 2. GeoGebra を用いて作成した幾何学問題

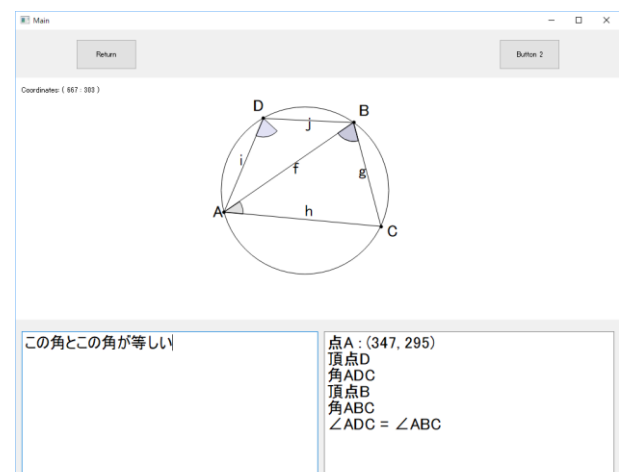


図 3. 幾何学問題解答時のインターフェース

CHI 99 Human Factors in Comp. Sys. Conf., pp. 576-583., 1999.

- [2] Richard A. Bolt, "Put-That-There: Voice and Gesture at the Graphics Interface", ACM Computer Graphics, Vol.14, No3, pp.262-270, 1980.
- [3] Madoka Miki, Norihide Kitaoka, Chiyomi Miyajima, Takanori Nishino and Kazuya Takeda, "Improvement of multimodal gesture and speech recognition performance using time intervals between gestures and accompanying speech", EURASIP Journal on Audio, Speech, and Music Processing 2014, 2014.
- [4] GeoGebra - Dynamic Mathematics, Geogebra, <https://www.geogebra.org/>
- [5] Julius, <http://julius.osdn.jp/>
- [6] Hoggan, E., Brewster, S.A., and Johnston, J. "Investigating the effectiveness of tactile feedback for mobile touchscreens", Proceedings of the SIGCHI conference on Human factors in computing systems, ACM (2008), 1573-1582.
- [7] Koskinen, E., Kaaresoja, T., and Laitinen, P. "Feel-good touch: finding the most pleasant tactile feedback for a mobile touch screen button", Proceedings of conference on Multimodal interfaces, ACM (2008), 297-304.
- [8] Bachl, S., Tomitsch, M., Wimmer, C., & Grechening, T. (2010). "Challenges for Designing the User Experience of Multi-touch Interfaces", Proceedings of the ACM Symposium on Engineering Interactive Computing Systems, 1-6.