

主論文の要旨

**Development and clinical evaluation of a contactless
operating interface for three-dimensional image-guided
navigation for endovascular neurosurgery**

〔脳血管内治療における3次元画像誘導ナビゲーションの非接触操作
インターフェースの開発とその臨床評価〕

名古屋大学大学院医学系研究科 総合医学専攻
脳神経病態制御学講座 脳神経外科学分野

(指導：齋藤 竜太 教授)

西堀 正洋

【緒言】

昨今の脳血管内治療では、治療開始時に血管撮影装置を用いた3次元(3D)回転血管撮影を行うことで、複雑な血管解剖を詳細に確認し、病変に適した角度の作成や治療方針の決定に用いることが必須となってきた。治療中でも状況に応じて重要な血管や角度が変わることがあり、再度ワークステーションの操作を行なう必要が生じる。術者自らワークステーションを操作すると、手術時間の延長や感染リスクが高まる可能性がある。清潔状態を維持しながら画像ナビゲーションシステムを操作できることは、手術を安全に成功させるために重要である。Kinectは、Microsoft社が開発したジェスチャー認識、深度カメラ、加速度計などの機能を備えている比較的安価なゲーミングデバイスである。Kinectを用いた非接触ジェスチャーナビゲーションシステムは、腹部手術、経皮的生検、検査の補助、リハビリテーション、医学教育などで開発・利用され既に報告されているが、脳神経外科領域では、開頭手術の画像ナビゲーションシステムの研究が唯一報告されているのみである。本研究は、愛知工科大学との共同研究として行い、清潔状態を維持したまま非接触で画像を操作できるシステムを開発すること、並びに実臨床に応用可能な精度に高めることを目的とした。

【対象および方法】

非接触かつ迅速に操作を行なうために、Kinectと音声認識デコーダソフトを用いて、手のジェスチャーと音声の両方でワークステーションの操作ができるシステムを考案した。既存のナビゲーションシステムに追加接続し、ワークステーションに表示されている3D画像を主な操作対象とした。システムの開発・導入環境の概要を表1に示す。

ハードウェアのセットアップと操作インターフェースの開発

Xbox One Kinect V2及び高性能大語彙連続音声認識デコーダソフトの”Julius”(Open source, version.4.1.3)を使用した。操作者はヘッドセットマイクを装着し、サブディスプレイ及びKinectは図1のように大型手術用ディスプレイの上に設置した。図2は、今回開発したネットワーク方式の構成を示す。術者の操作及び音声は同一PC内で動作・通信し、Bluetoothを介して既存のワークステーション(Artis Q BA Twin Syngo X workplace, Siemens Health Care)に送信される。モード変更は左手に割り当て、拡大や回転などのマウス操作は右手に割り当てた。音声は、角度保存、視点切り替えなど単純なコマンドを割り当てた。

ジェスチャーデザインインターフェース

操作に使用する手の形状は「じゃんけん」を模したものとした。図3は、手のジェスチャーと対応する機能を示す。左手を肩より高く上げると、右手の位置情報の更新を停止するように設定を追加したが、本機能により細かなズレが補正され、反応精度が向上した。

音声操作と認識

表2に音声コマンドと対応する機能を示す。音声認識率が低いSで始まる音は避け

るなど、工夫を行なうことで認識率を向上させた。音声コマンドが正しく操作されたかどうかを術者に伝えるために、術者のイヤホンに対応する英単語が聞こえるように設定した。本研究では、発音の滑らかさを考慮して日本語を採用したが、日本語以外の言語も代用可能である。

誤認識の防止

「チョキ」が「グー」と誤認識されやすいため、「チョキ」が数フレーム続くことで初めて認識されるように設定した。また周辺に立つ助手と操作者の手を区別するために、主操作者の立ち位置を意図的に狭い範囲で認識するようにした。音声コマンドについては、左手が「チョキ」の形をしているときのみ認識する設定とし、コマンドの種類を必要最小限の数に設定した。

【結果】

実際の臨床使用に先立ち、術者は 1-2 回操作トレーニングを行った。操作は一人でいき、ビデオ解析並びに時間測定は複数人で確認した。脳動脈瘤コイル塞栓術や脳動脈静脈奇形の治療など、実際の手術症例で使用し、それぞれの操作に要した時間を、録画したビデオを後方視的に解析して計測した。表 3 は、実際の治療において本システムを用いたワークステーションの操作に要した時間である。計測や画像の保存など、各ジェスチャー操作は 10-30 秒程度必要であった。音声認識については、新規に音声コマンドを認識するのに 4 秒必要であったが、連続であれば 3 秒でモード切替が可能であった。一症例に対する一連の操作は、合計 2-3 分で行うことができた。同じ動作の不要な繰り返しはなく、血管径の測定などの微細な操作も可能であった。術前のシミュレーションよりも音声認識率が若干低くなったが、周囲の会話やモニターの音などの環境音が原因であった。

【考察】

本研究では既存の 3D 画像ナビゲーションシステムに、ハンドジェスチャーと音声を用いた操作インターフェースを構築し、臨床使用に耐えうるインターフェースとなった。本研究は、非接触ナビゲーションシステムを実際の脳血管内治療に適用された初めての報告である。多くの施設においては、脳血管内治療専門医が一人で様々な役割を担っており、本システムの導入により迅速で有効な画像支援手術が可能であると思われる。従来の不潔な操作やドレープ下でマウスを使用する方法と比較して遜色ない所要時間であった。昨今 COVID-19 のため提唱されている Protected Code Stroke においては、患者や周囲の環境との接触をできるだけ少なくして治療を行なうべきとされており、本システムは感染症リスクがある状況下で有用である。従来のインターフェースと今回開発したインターフェースの構成比較を図 4 に示す。変換ユニットのソフトウェアを交換すれば、異なるナビゲーションシステムに対応することができる。

今後の課題としては、認識範囲内に人数が多い場合の応答時間延長の問題、指向性マイクによる環境音との干渉回避、単語学習機能付きの音声認識ソフトの使用などに

より、さらに洗練された認識率が得られる可能性がある。AIの発展に伴い、ジェスチャーセンサーの精度や音声認識率は日々向上しており、より操作性の高い次世代システムが感染症リスクや術者ストレスを軽減し、更に手術の質を向上させることができると考えている。

本研究の限界がいくつか存在する。施設間で手術室や血管撮影室の環境が異なり、操作性が変化する可能性がある。近年 Kinect が市販中止となったため、今後は他の深度センサー (Leap Controller, VicoVR, Orbbec camera, Intel Real Sense etc.) にも対応できるように、本システムのコンセプトを変更する必要がある。複雑なシャント病変における multi-modality の画像読影などには適しておらず、キーボード入力が必要な作業にも適していないため、このシステムがワークステーションの全ての機能を代替するものではない点である。

【結語】

既存のワークステーションシステムに Kinect と音声認識ソフトを組み合わせた非接触型操作インターフェースを開発した。脳血管内治療医は、清潔状態を維持しながら、十分な精度と操作性を持ってワークステーションを操作することが可能であった。