

外耳道の加速度を利用した頭部及び頭部内の各器官のリアルタイムな動作弁別

distinguishment between movements of head and organs in the head by measured accelerations from external auditory canal

学 鳥居航 (大阪大) 正 近藤大祐 (大阪大)
正 安藤英由樹 (大阪大) 正 前田太郎 (大阪大)

Wataru Torii, Osaka University, torii.wataru@marino.ise.end.osaka-u.ac.jp
Daisuke Kondo, Osaka University
Hideyuki Ando, Osaka University
Taro Maeda, Osaka University

Although a lot of interfaces exist and almost of them require our limbs and eyes to operate, it is difficult for us to treat multiple interfaces at the same time as we have only one pair hands and feet. On the other hand, we can move head, tongue and mouth independently from limbs. Consequently, we think it will improve our ability of manipulation interfaces by using these head or organ in the head while operation the interface requiring limbs and eyesight. Though such devices were investigated for the disabilities, the studies which treat multiple organ's movements as input are limited. The system we developed can distinguish stably multiple movements of head, tongue, blink and the shape of mouth by applying Independent Component Analysis(ICA). We calculated the generalized separating matrix from some subjects and verified that it separate movements of an third subject.

Key Words: ウェアラブルセンサ, 人間計測, 外耳道, ICA, インターフェース

1 概要

1.1 はじめに

人間の四肢を利用するインターフェースは数多く存在するが、手足はそれぞれ一対しかもたないために四肢を要求するインターフェースを複数同時に利用することは出来ない。それに対し、舌や頭部は四肢とは独立した動作が可能であり、特に舌に関しては四肢以外で複雑な随意運動が可能で唯一の器官である。このことから、頭部及び頭部に存在する各器官を利用するインターフェースの存在は、複数のインターフェースの同時操作を可能とし、人間の操作能力を向上させると期待できる。こういった舌や頭部を利用するインターフェースは主に身障者に向けて研究開発が進められており、視線を扱うもの [1, 2, 3, 4, 18, 19], 舌を扱うもの [5, 6], 発話を扱うもの [15], 頭部を扱うもの [9] と多岐に渡る。しかしその一方で、複数の入力を同時に扱った研究は [3] などに代表されるが限定的である。また、画像診断による視線入力に関しては頭部前面に装置が必要で携帯性が悪く、EOG による視線検出も外乱に弱いなどの問題がある。舌を扱うものに関しては口腔内に装置を置く、埋め込むなど利用者への負荷が大きいことが挙げられる。これらの問題を解決するインターフェースが必要であり、これを解決するような研究として、[3] や [10] ではイヤホン型デバイスを用いて外耳道の変形及び耳内圧の変化によって動作抽出を行っている。これは外耳道の変形によって頭部内の各器官の動作を抽出できる可能性を示唆している。

そこで本研究では、外耳道の耳壁変形を三軸加速度計で計測することにより、頭部及び頭部内各器官を複数まとめて扱うことの可能なインターフェースの開発を行った。また、インターフェースとして運用するためにはリアルタイムな動作弁別が必要であり、処理や演算が複雑でないものが望ましい。

1.2 独立成分分析

各動作の弁別手法には多変量データ解析手法の一つである独立成分分析 (Independent Component Analysis, ICA) を用いた。ICA は、特定の独立な事象が生じた際の標本集合に対し、線形分離を適用することで、事象を独立な成分として抽出するものである。本研究では、両耳それぞれの外耳道変形を 3 軸加速度計で計測するため、合計で六個の独立観測点を持つことになる。入力する時間情報を $I(t)$ 、求めたい動作の加速度変化の時間波形を s としたとき、独立な各事象同士を分離する演算子を用いて

$$s = AI(t) \quad (1)$$

表現でき、ICA はこの写像 A を求める手法である。

また、加速度の時間波形の変化に加えて、加速度を積分すると速度になることから、加速度の時間波形の一階微分及び積分したものを並べた、18 成分からなるベクトル

$$I(t) = \left[\begin{array}{c} \dot{x}(t) \quad x(t) \quad \int_{\Delta t} x(t) \end{array} \right] \quad (2)$$

を得る。この $I(t)$ に対して ICA を適用して独立成分を抽出すると、

$$X(t) = WI(t) \quad (3)$$

を満たす変換行列 W 及び独立成分 $X(t)$ が得られ、この変換行列 W は式 1 における A である。

式 3 によって得られた変換行列 W によって入力信号を線形分離することが可能となる。 $I(t)$ の中から選択した事象 $s(t)$ を取り出すには、変換行列 W の中の事象 s を独立成分を抽出する列



Fig.1 装置装着図

ベクトル W_s を用いて

$$s(t) = W_s I(t) \quad (4)$$

と演算することで求められる。

2 計測及び解析

2.1 計測

計測にあたり、ポリウレタン素材内部に加速度計 [KXP84-2050] を包んだ耳栓型デバイスを開発した。ポリウレタン素材と加速度計を固定してしまうと境界面が生まれることによってインピーダンスが高くなるため、ポリウレタン素材内部における加速度計の位置が変わらないよう配線部のみ固定を行ったものを図 1 にあるように両外耳道に挿入した状態で計測を行った。

頭部運動、舌運動、瞬目、「あいうえお」の母音を発話する際の口形形成、噛み合わせ運動、また口角挙上や眉間にしわを寄せることによる表情など合計 18 動作を入力動作とし、3 人の被験者から各動作、5 試行を 1 セットとし、4 セット繰り返すことで合計 20 試行計測を行った。

計測した動作集合から、ピークフィット及び波形の DP マッチングを行い、被験者間で安定して計測された動作、全ての被験者からは安定して計測されなかったが計測された被験者内では安定して計測された動作、計測不能だった動作の 3 つに分類を行った。

その結果、眉間にしわを寄せる動作、視線の水平移動及び「う」「お」の発話口形に関しては計測不能であったことから、これらを除く合計 14 動作を弁別の対象動作とした。図 2 に、ある被験者における実際に舌を前に突き出す動作を計測した際の 5 試行分の加速度変化の時間波形を示すが、このように、両耳の加速度計において、y 軸方向すなわち加速度計の挿入方向と垂直で地面に水平な方向について加速度が変化していた。

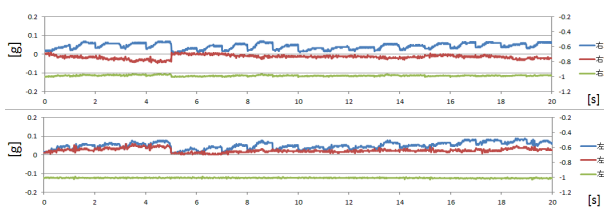


Fig.2 舌を前に突き出す動作の加速度時間変化

2.2 解析

ICA には入力として、式 2 に示したように、加速度変化の時間波形を加速度計の各軸ごとの時間軸方向につなげたものを 1 つの観測点情報として ICA に与え、3 軸の加速度成分それぞれに対

して微分項積分九を合わせて 9 成分、両耳合わせて合計で 18 成分を入力しているために、得られる独立成分は最大で 18 となる。そのため、理想状態では 17 動作を一度の線形分離で分離できるが、成分の強度や初期値によってその性能は左右される。また、事象同士が斜交している状態では、分離は可能であるがより分離性能が低下することが考えられる。そこでここでは、全事象の集合に対して一度 ICA を適用し、ICA によって独立成分として取り出された事象集合及び独立成分として取り出されなかった事象集合、それぞれに対して、事象を独立成分として分離出来なくなるまで繰り返し ICA を適用していくことによって動作の事象分離を行った。それから得られた結果は図 3 である。

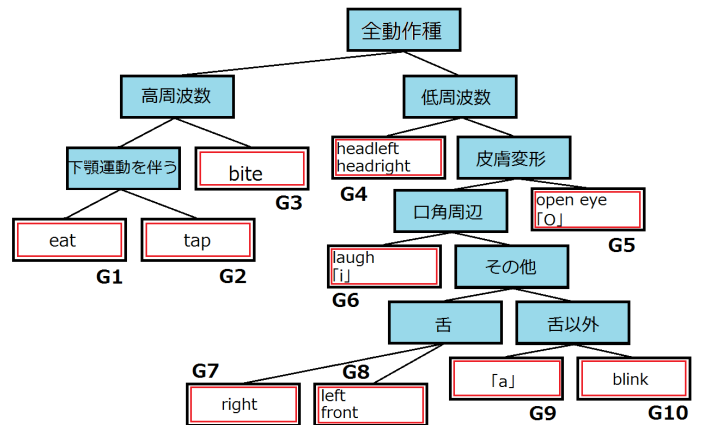


Fig.3 ICA による分離樹形図

図 3 の青色に塗られた節にある記述は、そこに示されている項目にしたがって動作が分離されていることを示している。なお、図 3 に見られる分岐は結果として青色枠の記述に従って分離されているが、これはトップダウン方式による ICA によって導かれた分岐であり、枠内の記述に関しては筆者が後からラベリングをしたものである。しかし、「その他」から下の葉にある動作集合は、ICA を適用しただけでは明確に各事象を分離することができなかった。そこで、瞬目及び「あ」の発話口形に対して舌の 3 動作が引き起こす皮膚変形は互いに類似していると考え、筆者の手により舌動作酒豪とその他から成る 2 動作集合に分割した各動作集合に対して ICA を適用した結果を示している。

事象が似通っており斜交しているものよりも事象同士がより直交に近い関係であるものが分離が容易であることに注意すると、筆者が記述した青色枠内の記述によれば、低周波数での動作と高周波数での動作の分離が非常に容易であることがわかる。そして木が深くなるほどに分離が難しく、事象として非常に似通ってくることになり、舌動作及び「あ」の発話口形形成、瞬目の分離が難しいことがうかがえる。

全動作を時間軸方向につなげて ICA を適用した結果得られた、高周波数での事象と低周波数での事象の分離の結果を実際に図 4 に示した。上で述べたように、独立成分は最大で 18 成分抽出されたが、ここでは特に、高周波数の事象をより良く独立成分として抽出している成分を抜き出した。なお、ICA によって得られた独立成分は成分であり、単位を持たない。

図 4 に示した独立成分は、高周波数での動作群と低周波数での動作群の二つの動作群のうち、高周波成分での事象についてのみ反応した。また、図 5 には、舌の左右方向への移動及び口外への

突出しの3動作にICAを適用した結果を示しているが、これは逆に、図5では舌の突出し動作に比較的大きく反応するもの全事象に反応しており、分離が難しいといえる。

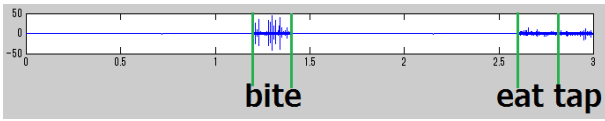


Fig.4 周波数での事象分離

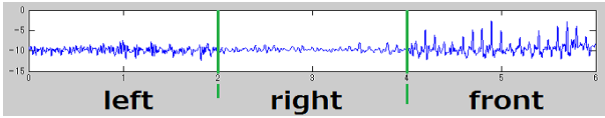


Fig.5 舌3動作の事象分離

2.3 変換行列の汎用化

三人の被験者それぞれから得た変換行列から汎用変換行列を作成した。この汎用変換行列 W_G の動作分離性能を定量的に評価する必要がある。本来独立成分分析は各事象が互いに独立であった場合に、その各事象をそれぞれに分離するものであるため、各動作に対応した部分変換行列によって完全に分離される。

ある動作 s の時間波形 $I_s(t)$ とそれ以外の動作群の時間波形 $I_{\bar{s}}(t)$ に対して、動作 s を取り出す変換行列の列ベクトルからなる変換行列の部分行列 A_s 及び、動作 s 以外の動作集合 \bar{s} を取り出す変換行列の列ベクトルからなる変換行列の部分行列 $A_{\bar{s}}$ を考えたとき、式5から8にある式で分離性を確認した。これらの式における“ \rightarrow ”は、たとえば式5においては、入力動作の時系列情報 $I_s(t)$ に対して変換行列の中の、動作 s を抽出する部分変換行列を作用させたときに動作 s として判別されることを示している。

上で求めた汎用変換行列は三人の被験者から得た加速度の時間波形情報から生成したものであるため、汎用行列の性能評価にこの被験者3人の加速度情報を使用するのは適当でない。

そこで、新たな被験者として成人男性を一人用意し、この被験者から前三人と同様の実験手順で3にある動作を、性能検証のために各動作につき30試行を計測した。

計測した動作を時系列に並べたものに対して汎用変換行列と、この被験者のみの情報から得た変換行列を適用して性能を比較した。この際に用いた評価は、式5に関しては樹形図に沿って分離を行った際の正答率を、式6、式7に関しては、変換行列の各動作抽出に対応する部分行列を対象とする動作の全試行に作用させた時の誤認率であり、実際には式9、式10に示した処理を行った。また、8に関しては他3式から導かれる。

その結果を表1に示した。なお、表中のG1-G10は図3のそれと対応している。

$$I_s(t)A_s \rightarrow s \quad (5)$$

$$I_s(t)A_{\bar{s}} \rightarrow \bar{s} \quad (6)$$

$$I_{\bar{s}}(t)A_s \rightarrow s \quad (7)$$

$$I_{\bar{s}}(t)A_{\bar{s}} \rightarrow \bar{s} \quad (8)$$

$$\frac{n(\{x(t)|x(t) \in \{I_s(t)\}, x(t)A_s \rightarrow I_s(t)\})}{n(\{I_s(t)\})} \quad (9)$$

$$\frac{n(\{x(t)|x(t) \in \{I_{\bar{s}}(t)\}, x(t)A_{\bar{s}} \rightarrow I_{\bar{s}}(t)\})}{n(\{I_{\bar{s}}(t)\})} \quad (10)$$

Table 1 各動作の認識率

動作種	汎用変換行列			個人変換行列		
	式5	式6	式7	式5	式6	式7
G1	1.00	0.00	0.00	0.73	0.06	0.00
G2	0.47	0.00	0.00	0.97	0.10	0.00
G3	0.72	0.15	0.00	0.68	0.08	0.00
G4	1.00	0.20	0.00	0.90	0.29	0.00
G5	0.80	0.82	0.27	0.69	0.75	0.42
G6	0.68	0.81	0.32*	0.87	0.86	0.15*
G7	0.03	0.97	0.87*	0.57	1.00	0.1*
G8	0.90	0.92	0.87*	0.92	0.94	0.10*
G9	0.97	0.33	0.50*	0.70	0.67	0.10*
G10	0.91	0.86	0.36	0.90	0.97	0.30

これより、*のついた箇所からわかるように個人から導出した変換行列に対して誤認率が比較的高くなっている項目が見られる一方、太字で示した箇所から正答率に関しては、被験者自身から導いた変換行列と汎用変換行列で顕著な大きな差は見られなかった。音声認識のインターフェース研究 [20] によれば、インターフェースとして運用する際には85%以上の正答率が必要とされているが、一部の動作において9割前後の正答率が出ていることから、入力動作が誤認識された際、即時に誤認識された入力を取り消し、再入力を行うことによってインターフェースの操作性が確保される可能性を示唆している。

また、表1の、被験者4自身から求めた変換行列 W_4 に対する汎用変換行列 W_G の性能を表2に示す。これより、第三者の変換行列集合から求めた W_G は W_4 に対してほとんどの動作種について、正答率(式 refreq:compare1)に関してその性能を維持しており、各動作によって両外耳道に生じる加速度変化の時間波形が被験者によらないものであるといえる。しかし W_4 に対して、式7であらわされる誤認率が高くなっていた。

Table 2 各動作の認識率

動作種	W_G/W_4		
	式5	式6	式7
舌左, 舌前	0.98	0.98	1.00
舌右	0.05	0.97	1.00
口角拳上, 「い」	0.78	0.94	0.21
見開き, 「お」	1.00	1.00	0.64
瞬目	1.00	0.89	1.00
噛み合わせ	1.00	1.00	0.00
頭部左, 頭部右	1.00	0.69	0.00
「あ」	1.00	0.49	1.00
咀嚼	1.00	0.00	0.00
口蓋を叩く	0.48	0.00	0.00

本実験では、非コントロール動作として咀嚼を弁別対象の動作に組み込んだが、これが他動作と混同されなかったことは、他の日常の動作に関してもコントロール動作と区別される能性を示唆していると考えられる。

2.4 複合動作の分離検証

ICAによって求められた変換行列は、変換行列内の各列に分離可能な動作が一つ対応している。これは、異なる複数の動作を同時に行って得られる加速度変化の時間波形に変換行列を適用することで、同時に行った各動作をそれぞれ取り出せる可能性を示唆している。インターフェースが複数の同時入力を取り扱えることによって、操作能力をさらに向上させることが可能になると考えられる。そこで、汎用行列の性能検証に用いた被験者に、歯の噛み合わせ動作及び口角拳上動作に関して、単一動作及び同時に行う複合動作の3種類を各20試行行ってもらい、得られた加速度変化の時間波形に上で求めた汎用行列を適用した。

そして、歯の噛み合わせ動作及び口角拳上動作の独立成分に対応する成分を抜き出したものを図6に示す。噛み合わせ動作のみを行った際には噛み合わせ動作の独立成分に対応する(i)が、口角拳上動作に関しては(ii)が、そして複合動作に関しては(i)(ii)両方に対応成分があらわれるはずであるが、口角拳上の独立成分には歯の噛み合わせ動作の成分が強く立ちあらわれていることが見て取れる。

これは、本研究ではトップダウン方式によって動作分離を行ったために、口角拳上動作の独立成分と歯の噛み合わせ動作の独立成分が完全に独立である保証がないこと、また、汎用変換行列では各動作に対応する部分変換行列から抽出した特徴が丸まることなどが原因として考えられる。

しかし、図にもみられるように、各事象に対応する独立成分における波形は抽出されているために、波形マッチングなどから複数の入力を扱える可能性がある。

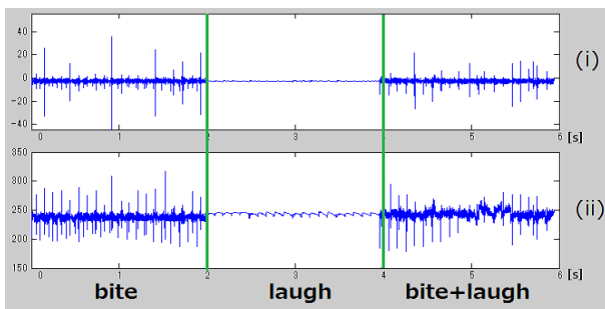


Fig.6 複合動作の分離

3 おわりに

本研究では、四肢と同時且つ独立に利用可能なインターフェースとして頭部及び頭部上各器官に注目し、独立成分分析を用いて各動作の弁別が可能であること、及び求めた汎用変換行列によって第三者の動作を正答率を維持して弁別できることを確認した。

しかし、舌の内部で起始停止する筋を用いた動作及び眼球動作を取り扱うことができなかった。これらの動作も弁別が可能とする改良などが考えられる。加えて、舌以外にも頭部にある各器官の動作は本研究で扱ったもの以外にも多くの種類があり、また測定箇所についても外耳道以外にも多くの部位が考えられ、実際に、対耳珠の動きを観察することによる瞬目動作の検出や二極 EOGでの視線移動の検出を既に確認している。そこで、観測点の追加、取り扱う動作の検討及び計測する情報の追加の検討を行い、よりインターフェースとして再現性の高い安定した動作の探索などが課題として挙げられる。

本研究では、インターフェースを大量生産できることも重要であると考え、被験者毎に装置を調整することなく単一のデバイスでの計測安定性、動作弁別性を検証した。しかし、複数人の動作を計測したが外耳道の径には個人差があるために被験者によっては計測不能な動作が存在したため、複数の被験者からの計測の安定化について課題が残っている。また、上で見たように複数の入力を同時に扱える可能性があり、インターフェース処理能力の大きな向上が期待できることから複数動作の同時入力の検証及び弁別の安定化などが考えられる。

References

- [1] 阿部清彦, 大内幹夫, 大井尚一, 大山実: 画像解析による強膜反射法を用いた視線入力システム, 2003, 映像情報メディア学会誌, Vol.57, No.10, pp.1354-1360
- [2] 真鍋宏幸, 福本雅明: ヘッドホンを用いた EOG 法による視線入力インターフェース, 情報処理学会論文誌, 2011, Vol.52 No.4, 1515-1526
- [3] 谷口和弘, 西川敦, 小林英津子, 宮崎文夫, 佐久間一郎: みみスイッチ: 外耳の動きを入力情報とする常時着用型入力装置, 情報処理学会シンポジウム論文集, 2010, No.4 pp.243-246
- [4] 真鍋宏幸, 福本雅明: 常時着用インターフェースの実現を目指したイヤホンによる視線検出技術, NTT DOCOMO テクニカルジャーナル, Vol.18 No.3, pp.13-17
- [5] 一ノ瀬裕, 本多清志, 佐藤準二, 和久本雅彦, 東輝明: ワイヤレス式舌圧センサを用いたヒューマンインターフェースと電動車椅子の制御への応用, 電子情報通信学会論文誌, 2003/2, Vol.J86-D-No.2, pp.364-367
- [6] Lotte N.S., Andreasen Struijk: An Individual Tongue Computer Interface for Control of Computers and Assistive Devices, IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, 2006/12, Vol.53, No.12, pp.2594-2597
- [7] 安藤英由樹, 渡邊淳司, 杉本麻樹, 前田太郎: 独立成分分析による爪装着型センサの接触力と屈曲の分離, 2003, TVRSJ, Vol.8, No.4, pp.379-387
- [8] Cynthia Lau, Stephanie O'Leary: Comparison of Computer Interface Devices for Persons With Severe Physical Disabilities, 1993, Am J. Occupat. Ther., Vol.47, pp.1022-1030
- [9] Y.Chen, Application of tilt sensors in human-computer mouse interface for people with disabilities, 2011/9, IEEE Trans, Rehab. Eng., Vol.9, No.34, pp.289-294
- [10] Khondaker A. Mamun, Michael Mace, Lalit Gupta, Carl A. Verschuur, Mark E. Lutman, Maria Stokes, Ravi Vaidyanathan, Shouyan Wang: Robust real-time identification of tongue movement commands from interferences, 2012/3, Neurocomputing, Vol.80, pp.83-92
- [11] 村田昇一: 入門 独立成分分析, 東京電機大学出版局, 2004
- [12] John H. Warfel: The Extremities (6th edition) and The Head, Neck, and Trunk: Muscles and Motor Points (6th Edition) by Lea & Febiger (1993). (図説 筋の機能解剖 第4版, 医学書院, 1993)
- [13] 岡田慎矢: 人間の行動意図に基づく動作伝達手法の提案とロボット操縦における実現可能性の検討; 大阪大学大学院情報科学研究科 修士学位論文, 2010.
- [14] 丹羽真隆, 飯塚博幸, 安藤英由樹, 前田太郎: つもり制御: 人間の行動意図の検出と伝送によるロボット操縦, 2012/3, 日本バーチャルリアリティ学会論文誌, No.17, pp.3-10
- [15] 瀧澤由佳子, 樋口行平, 前中一介, 才木常正: 歯装着型加速度センサを用いた音声認識の提案, 2012/7, 電気学会論文誌. C, 電子・情報・システム部門誌, No.132, pp.1206-1207
- [16] 寺田賢治, 宮原宏幸, 久保靖: 阿波踊りの動作の定量化の試み, 2011/3, 画像電子学会, Vol.34, No.3, pp.220-227
- [17] 高橋長雄: からだの地図帳, 講談社, 第44刷, 2008
- [18] 久野悦章, 山羊透, 藤井一幸, 古賀一男, 内川嘉樹: EOG を用いた視線入力インターフェースの開発, 情報処理学会論文誌, 1996, Vol.39, No.5, pp.1455-1462
- [19] 千葉修平, 堀潤一: Development of EOG-Based Character Input Interface Controlled by Three-Directional Eye Movements, 電子情報通信学会技術研究報告. WIT, 福祉情報工学, 2010, Vol.110, No.53, pp.39-44
- [20] 南條浩輝, 秋田祐哉, 河原達也. 音声認識を利用した会議録・講演録の作成支援システムの設計と評価. 日本音響学会秋季発表発表会講演論文集, 17-13, 2005.