

複雑系の数理とシミュレーション

Mathematics and simulation of the complex system

1014250 平井 力 Chikara Hirai

1. 背景

ニューラルネットワークとは、人間の脳神経細胞を模倣する数理モデルである。人間の脳には多くの神経細胞(ニューロン)が存在し、互いにつながりを有することで複雑な処理を行う。そのネットワークを人工的に再現することで、学習能力を持たせることが可能となった。このニューラルネットワークは人間の認知的課題に応用されている。そこで、本プロジェクトの活動ではニューラルネットワークを用いた研究として、読唇術と呼ばれる技術について研究を行い、読唇術を行うために必要な情報は何かを調査することになった。

読唇術とは、口唇の動きや形状から発話内容を観察し、その発言の内容を知る方法を指す[1]。この技術は聴覚に障害を抱えている者が健常者の発話内容を予測するために利用されている。

2. 課題の設定と到達目標

本プロジェクトの目標は、人間の認知過程についての研究を、ニューラルネットワークを用いて行うことである。今年度の活動では読唇術に注目し、読唇術を実現するためにはニューラルネットワークを用いる必要があるかを調査することで、読唇術を行うためにはどういった情報が使われているのかを研究することを目標とした。

しかし、ニューラルネットワークについての知識がほとんどない状態から活動が始まったため、この目標を達成するにあたり、前期の活動ではメンバー全員のニューラルネットワークについての理解を深め、後期の活動のための準備をするということを目指して活動することになった。

後期の活動では、読唇術についての研究を行うことになったため、読唇術とは何かを理解を深め、読唇術を実現させるためにはどういった情報が必要なのかを分析、研究を行っていくことで精度の高い読唇術を行えるマシンを作成することで読唇術に必要な情報を調査することを最終目標とした。

3. 課題解決のプロセス

本プロジェクトの目標はニューラルネットワークを用いて人間の認知過程の研究を行うことであり、今年度の活動では読唇術に着目した。読唇術に関する情報を入手する為に必要なことは、読唇術を実現させるために使用するニューラルネットワークの理解を深めること、読唇術を実現するためにはどういったニューラルネットワークを用いる必要があるか、どういったアプローチをする必要があるかを考えて構築を行うことである。これらを解決するために前期と後期でそれぞれ目標を

設定し、各月の活動を行っていくことで、我々が設定した目標の達成できるように活動を行っていった。

前期の活動の目標はニューラルネットワークについて基本的な知識を深め、後期の活動ではどのような人間の認知過程の研究を行うかを決定することとした。ニューラルネットワークについての理解を深めるためにメンバー全員で指定した参考書[2]を利用して輪読を行った。輪読によって各々のニューラルネットワークについての理解を深めることができた。また、前期で学習した内容についての理解を深めるために、[2]内に記載されていた Hopfield 型ネットワーク、及び Boltzmann Machine の動作を再現するプログラムを作成し、理解を深めることができた。しかし、後期の活動でこういった活動を行っていくかを決定することまでは出来ず、候補を出すだけで前期のプロジェクト学習の活動が終了してしまっ

た。後期の活動では、前期のうちに決定する予定であった、人間の認知過程を研究する為にどういったことに着目するかということを決めるということから始まった。前期の案にも出ていた読唇術についての研究を行うということをメンバー全員が賛同したため、後期の活動では読唇術について研究していくことになった。その後は読唇術を実現させるためにはどういった情報が必要か、どういったニューラルネットワークを構築する必要があるかを調べる必要があったため、プログラム関係を担当するグループと先行研究の調査を行う小規模なグループと役割分担を行い、分担して作業を行っていった。

4. 実験

本プロジェクトではニューラルネットワークを用いて読唇術を実現するにあたって、どういった情報を用いる必要があるか、及びどのようなニューラルネットワークを用いる必要があるかを調査することを目的とし実験を行った。文献[3]によると、現代の技術では口の面積のみで完全な読唇術を行うことはできないということが述べられている。そこで今回の実験では口の面積以外にも鼻の先端から顎まで距離を追加で測定し、そのデータを利用することで読唇術を行うことにした。

読唇術を行うためのニューラルネットワークは、各単語に含まれる母音を出力し出力された母音の並びから発話された単語を予測するという方式(以下 A)と、単語を直接出力するという方式(以下 B)の2種類を用意した。それぞれのニューロン数と形状は以下の図1、図2に記す。

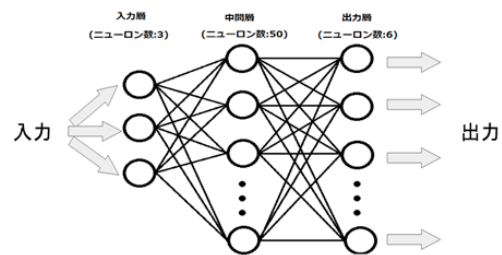


図 1:A のニューラルネットワーク

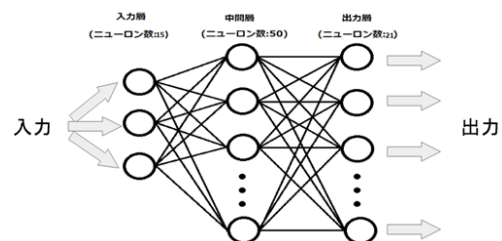


図 2:B のニューラルネットワーク

次に口の周りの特徴点のデータの取得方法について説明する。被験者 1 人 1 人に母音を発音してもらっている時に口の周りの特徴点の座標を抽出して x_1 と x_2 、 x_3 と x_4 、 x_5 と x_6 、 x_7 と x_8 間のそれぞれの距離を取得する(図 3)。この時の特徴点は Kinect を用いて取得した。また、被験者の顔と Kinect のカメラとの距離によって生じる口の大きさの違いを修正するためにどんな発音をしたときでも変化しないと考えられる目頭間の距離(図の x_1 と x_2 間の距離)で割る。

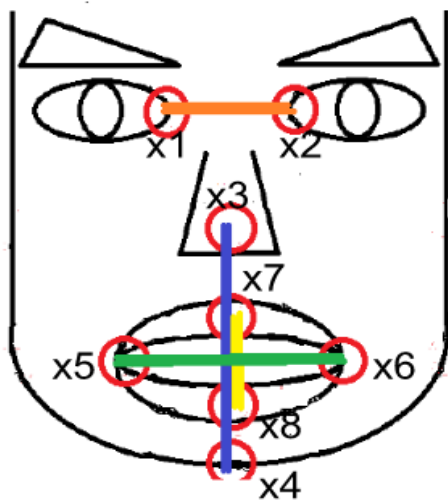


図 3:取得した顔の距離

また、今回の実験で使用した単語は次の 21 種類の 5 音の単語に限定した。

- ・あさひかわ(aaiaa)・あさごはん(aaoan)
- ・ありがとう(aiaou)・えいがかん(eiaan)
- ・おかあさん(oaaan)・おしょうがつ(ouau)
- ・おむらいす(ouaiu)・くりすます(uiuau)
- ・こんにちは(oniiia)・こんばんは(onana)
- ・さようなら(aouaa)・しいげんご(iieno)
- ・しんぶんし(inuni)・すみません(uaen)
- ・どらえもん(oaen)・ひるやすみ(iuai)
- ・ふゆやすみ(uiuai)・ぶろじえくと(uoou)
- ・ぶろぐらむ(uouau)・みらいだい(iaiai)
- ・ゆきだるま(uiaua)

実験の流れについて、今回の実験は以下の手順に従って行った。

I：ニューラルネットワークの学習を行うため、被験者にリスト内の単語を発話してもらいその単語に関する顔の座標データを取得しそのデータを利用しニューラルネットワークに学習させる。今回の実験ではそれぞれのネットワークには 20 人分の学習用データを用いた。使用するデータについては図 4 にデータを取得するための環境については記述する。

II：実験で使用するリスト内の単語のデータを被験者から取得する。この時に I と II の被験者は同一人物だと学習用データと実験で使用するデータが完全に一致してしまう可能性があるため、異なる人間からデータを収集するようにした。また、データの収集環境は I と同様の環境で行った。

III：リスト内の単語からランダムに 1 つ選出し、A、B のそれぞれのニューラルネットワークに入力し、それぞれのニューラルネットワークが出した結果と実際に入力しようとした単語が一致しているかを調べ、一致している場合は 1、不一致の場合は 0 と記録する。それぞれのニューラルネットワークは 2 つ以上の単語が予測される可能性が存在するため、複数予測された場合に、それらの中に入力した単語データと一致したものが存在する場合は一致した場合と同じく 1 と記録する。

IV：III の手順を 100 回繰り返すことで、それぞれのニューラルネットワークの正答率求め、それらを比較し A と B のどちらのニューラルネットワークのほうがより優れているかを検証する。この比較を行うにあたってそれぞれの正答率に有意な差が存在するかどうかを確認する為にカイ二乗検定を行うことで調べる。

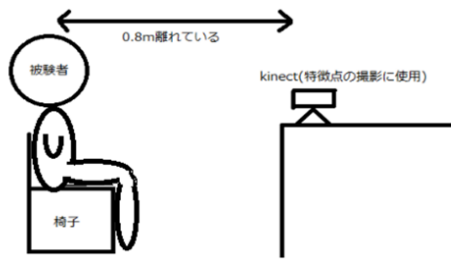


図 4:実験中の環境

5.結果と考察

測定の結果、Aの正答率は21%であったことに対して、Bの正答率は45%であった。またカイ二乗検定の結果からこの2つは有意な差が見受けられることから、AとBのパターンではBの方がより正確な読唇術を行うことができることが判明した。

このような結果になった要因として考えられることは、Bの方が入力に使用するデータが多くなり、それを元に判別を行うことからではないか、若しくは、[3]にも各母音の識別は口の面積だけでの判別が難しいという記述があり、それ以外の距離を追加した場合でも同じく口の形から母音の識別は困難であるからではないか等が考えられる。今回の実験では前者、後者共に別の実験を行って検証を行う必要があり、活動時間内に検証を行うことは厳しいということもあって、このような結果になった要因を決定できるだけのデータを用意することが出来ないという状態で終了してしまった

6.成果とまとめ

今回のプロジェクト学習における成果について、前期の活動ではニューラルネットワークについての基本的な知識を学習することが中心となっており、成果物として参考書[2]内に記述されているネットワークの例題を実際のプログラムで再現させるといったものが挙げられる。後期の活動では読唇術についての活動が中心となっており、ニューラルネットワークを用いた読唇術を行うためのプログラム、及び読唇術に関する研究結果が挙げられる。

最後に今回のプロジェクト学習を通じ、計画を細かく立てること、及び常に自身の担当部分だけをこなすのではなく、他の作業の状況を各自がしっかり理解し、お互いにコミュニケーションをとることで全体の作業状況を明確化することの重要性を学ぶことが出来た。

参考文献

- [1]ブリタニカ国際大百科事典
- [2] 麻生英樹 ニューラルネットワーク情報処理 -コネクションイズム入門、あるいは柔らかな記号に向けて- 産業図書,1998
- [3] 小林宏,原文雄.ニューラルネットワークによる人の基本表情認識.計測自動制御学会論文集 1993.