

<b>No.10</b>	<b>高度化</b>		
<b>氏名</b>	<b>阿部 竜也</b>	<b>情報理工学部</b>	<b>2回生</b>
<b>1. 出願時のテーマ・目標を具体的に記述してください。</b>			
<p>テーマ：教師なし新規音響コンテキスト適応とSpeech-Speechブラインド音源分離</p> <p>昨年度、私は「単一チャンネルでのSpeech-Speechブラインド音源分離」の研究を行った。ここでは、音源分離を時間周波数毎の音源のクラスタリング問題として定式化し、各時間フレームをBLSTMの入力とすることで時間周波数マスクを推定、非目的音源をマスキングすることで音源分離を実現した。</p> <p>今回はこれを一般化することを目的とした研究を行う。現状の音源分離技術の問題点としては、“教師データに含まれていない音響コンテキスト（話者/性別/マイクなど）を抽出対象音源にできない”という点である。つまり、音源分離システムは私の声を抽出することに最適化されているため、抽出対象話者を切り替えるには大量の教師データを用意する必要がある。</p> <p>そこで、CADNN（コンテキスト適応型DNN）の発想を取り入れ、教師なし新規音響コンテキスト適応を可能とする音源分離技術の研究を行うこととする。</p>			
<b>2. 上述のテーマ・目標を実現するために実施した計画を具体的に記述してください。</b>			
<p>①2020年4月～6月</p> <ul style="list-style-type: none"> <li>BLSTMを用いた音源分離システムのアーキテクチャを改造 <ul style="list-style-type: none"> <li>CADNNの発想を取り入れ、メインNNとサブNNに分解する</li> </ul> </li> <li>音響コンテキストから抽出する補助的特徴量の選定 <ul style="list-style-type: none"> <li>抽出された特徴量はサブNNの入力とする</li> <li>現時点では、i-vectorやボトルネック特徴量を利用しようと考えている</li> </ul> </li> </ul> <p>②2020年7月～9月</p> <ul style="list-style-type: none"> <li>教師データを収集 <ul style="list-style-type: none"> <li>できるだけ多くの音響コンテキストのデータを用意する</li> <li>自分で用意できない分は、友人に依頼する予定（承諾済み/感染拡大対策として、友人には自宅で録音後、送信していただく）</li> </ul> </li> <li>音声の前処理について調査 <ul style="list-style-type: none"> <li>購入予定である『前処理大全』や、音響工学についての論文などを参考に、前処理を行う</li> </ul> </li> <li>学習</li> </ul> <p>③2020年10月～</p> <p>ここでは、主に学習したモデルの評価・検証やトラブルシューティングを行う。すぐには成功しないだろうと考えられるため、10月以降は細かな調整を行う。</p> <p>※ この手法では実現不可能と判断される場合には、手法を抜本的に見直す</p>			
<b>5. 今回（今年度）の取り組みについて、今後の活動展開と展望を記述してください。</b>			
<p>新規音響コンテキスト適応は保留とし、分離後の音質向上に努める。</p> <p>今回は時間周波数マスキングを行う際にはバイナリマスクを用いているが、ソフトマスクを使うなど波の連続性を維持できる手法を模索する。</p>			
<b>6. 今回（今年度）の取り組みは、今後の学びや進路にどのように影響しますか。</b>			
<p>全く知らない状況では無くなったため、今後音響工学について扱う際のモチベーションを保ちやすくなったと考えている。</p>			
<b>7. 今回（今年度）の活動が周囲に与えた影響（社会・周囲）への貢献・還元の点で記述してください。</b>			
<p>noise-speech音源分離とは異なり、speech-speech音源分離では人と人の声を分離することができる。難聴を患っている人は人の声を聞き分ける選択的聴取の機能が発揮しづらくなる場合があるため、特定の人の声のみを取り出す仕組みはこの解決に有用であると考えられる。</p>			

<b>3. 個人の成長の軌跡3-1. 取り組みの過程でどのようなことがあったのか、グラフを作成してください。</b>	
<b>3-2. グラフで書いた☆（個人がもっとも成長したと思うポイント）では、その過程で学んだこと、気づいたことについて具体的に書いてください。</b>	
<p>一般的な音源分離手法について、</p> <ul style="list-style-type: none"> <li>優決定条件：マイク数 ≥ 音源数</li> <li>劣決定条件：マイク数 &lt; 音源数</li> </ul> <p>のそれぞれにおける手法を複数の論文を読み調査した。ここでは専門用語が絡む関係上安易に翻訳ツールを使わず、音響工学用語の英語における表現方法を学ぶことができたと考えている。当然、様々なフィルタやモダンな前処理、深層学習の音響工学への応用についての知識も深まった。</p>	
<b>3-3. “今回（今年度）の取り組み”と“正課の学びや取り組み”は、どのような関連や影響（相互作用）がありましたか？</b>	
<p>正課に取組み意欲が凋落することとなった。理由として、自己学習と正課における学びのフローの違いが挙げられる。</p> <p>正課では、理論から積み立て応用へ向かうというボトムアップ的な学びが多いと感じる。私の場合、そのような手法ではモチベーション維持が困難であるため、応用から入り気になる理論を調べるといったトップダウン的な過程で学ぶ方が楽しく思う。</p> <p>結果として、正課は単位が取れる程度の参加に留める結果となった。</p>	
<b>4. 本奨学金を受給したことで、以下の項目についてどのような影響を与えたか5段階で評価してください。（該当ナンバーに○）また、併せて評価の理由も書いてください。評価例：【 1（達成できなかった） ← 3（どちらともいえない） → 5（達成できた）】</b>	
① 目標の達成度	2
<理由> シングルチャンネルでのブラインド音源分離は非常に困難で、さらにそれを新規音響コンテキストに迅速対応するという難しいテーマであったため達成できなかった。しかしながら、自身の成長には確実に寄与する活動であったため達成度を2としている。	
② 計画の達成度	2
<理由> 本テーマは春学期中に自分の力では解決できないだろうと判断したため、計画の達成度は低い結果となった。	
③ 取り組みを通じた自己成長	5
<理由> 興味あるトピックを見つけた際に、まずは手を動かす癖がついたため。	
<b>10. 今年度の取り組みを通じて最も身についたと思う力について、具体的に記載してください。9の設問で回答した力でも、それ以外でも構いません。</b>	
① 身についた力	難しい内容を短時間でわかりやすく伝える能力
② ①で記述した力について具体的に説明してください	例えば、今回のテーマ「教師なし新規音響コンテキスト適応とSpeech-Speechブラインド音源分離」を5分間で説明する能力
③ なぜその力を身につけることが出来たのか、成長を手助け・促進させた要因を記載してください	私が所属しているRCC（立命館コンピュータクラブ）では、毎週の定例会議でLTを実施しており、発表は5分間という時間制限があるため、短時間で伝える努力が義務付けられる。この活動が、難しい内容を短時間で伝える能力の向上を促進させた。