

高精度な残響除去技術を実用化に成功

～ナレーションなどに適した聞き取りやすい音声の制作が簡単に～

日本電信電話株式会社（以下NTT、本社：東京都千代田区、代表取締役社長：三浦惺）は、録音素材に含まれる残響を高精度で推定する基盤技術を開発し、残響があるため聞き取りにくかった録音素材から、残響の低減された聞き取りやすい高品質な音声を取り出すことに成功しました。

この基盤技術をもとに、エヌ・ティ・ティ ラーニングシステムズ株式会社（以下NTTLS、本社：東京都港区、代表取締役社長：岡田昭彦）は、「NTT MEDIA LAB」として運営しているポストプロダクション※1（以下 ポスプロ）業務で培ったノウハウを活かし、国内外のポスプロで広く利用されている音声編集ツール※2のプラグインソフト※3「NML RevCon-RR」を開発しました。本製品を使うことで、従来難しかった、録音素材に含まれる不要な残響の除去が可能となり、CMや映画の種々の場面の演出意図に沿った最適な残響特性をもつ音声の制作が簡単にできるようになります。

1. 残響除去技術の概要

部屋の中で離れたマイクに向かって人が話した場合、その音声は、部屋の形や壁面の材質に起因する残響を伴った音として録音されます。今回開発した技術は、この録音音声に含まれる残響を高精度かつ簡単に推定・除去することを可能にした基盤技術です。また、本基盤技術に基づいた本製品によって、映画やドキュメンタリー番組の演出で利用される響きの無い音声や、アフレコで収録したようなはっきりとした聞き取りやすい音声を、より簡単に制作できるようになります。

2. 開発の経緯

NTTのコミュニケーション科学基礎研究所（以下、NTTの研究所）では、いつでもどこでも利用可能な音声インターフェース技術の実現を目指して、日常環境下で聞き取りやすい音声を抽出する技術の開発の一環として、録音音声に含まれる雑音や残響を取り除く技術の研究を行ってきました。その結果、録音音声に含まれる残響を効果的に低減させ、音声を聞き取りやすくするソフトウェアの基盤技術を開発しました（図）。

一方、CMや映画を制作するポスプロでは、録音音声に残響が含まれていたために制作者の意図する演出を実現できず、俳優を再度手配して録音を

やり直す場合も多々ありました。このような状況を解決するため、NTTLSは、ポスプロ業務で培ったノウハウをNTTの研究所で開発した残響除去技術に取り入れ、ポスプロ業界向け音声編集ツールの残響除去プラグインソフトを開発しました。

3. 技術のポイント

部屋の残響の特性や、残響の影響を受ける前の声の情報等が予めわかっているならば、残響を取り除いた音声を得ることは容易です。しかし、録音音声の情報しか利用できない状況で残響除去を行うことは、非常に困難とされていました。

この処理を実現するためにNTTの研究所は、残響を含む録音音声（直接音に相関の高い「原音」（直接音および直接音から30～50ミリ秒以内に録音される反射音）と、直接音と相関の低い残響※4にわかれることに着目し、マルチステップ線形予測※5を適用することで、録音音声のみから残響を高精度に推定することを可能にしました。

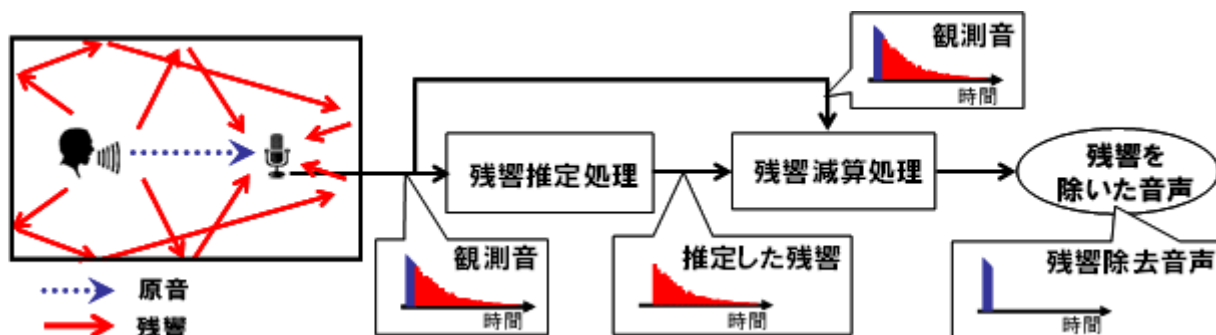


図. 残響除去の処理概要

4. 製品化のポイント

ポスプロでは、聞き取りにくいセリフやナレーションを聞き取りやすくするために、録音音声を調整、整音することがあります。録音音声に不要な残響が含まれる場合、従来はエンジニアが雑音抑圧やボリュームコントロール等様々なツールを駆使し、膨大な時間をかけ、出来る限り残響を目立たなくするような作業を行ってきました。

NTTLSは、ポスプロ業務で培ったこうしたノウハウをNTTの研究所で開発した残響除去技術に取り入れ、残響除去処理後の音声品質の向上と使い勝手の向上を図り、プロが求める品質と、録音素材ごとの調整を適切かつ簡単にできる機能を備えたプラグインソフトとして製品化しました。

5. 今後の展開

NTTLSは、本技術を用いた製品「NML RevCon-RR」を2008年11月19日

(水)～21日(金)に幕張メッセで行われる国際放送機器展(Inter BEE 2008)に参考出品し、ポスプロスタジオのエンジニアなどからフィードバックを得て、機能の改良や使い勝手の向上を図っていきます。なお、本製品の発売開始は2009年1月、価格は12万円程度を想定しております。

またNTTは、残響で困っている様々なシーンを解決するため、民生用機器等への応用など、本技術を利用した技術・製品・サービスの検討を進めていきます。

<用語解説>

- ※1 放送やパッケージメディア、映画などの映像作品の制作における撮影後の作業(例えば、映像の編集・調整、音の編集・調整、フォーマットの変換など)の総称。
- ※2 複数の録音素材をデジタル化してコンピュータ上に取り込み、ボリュームや音色等を加工・調整し、最終的な作品としてまとめあげるためのソフトウェア。なお、本製品では、Digidesign社のPro Toolsを対象としています。
- ※3 アプリケーションソフトが本来備えていない機能を追加するために、後から組み込むソフトウェアのこと。
- ※4 音声信号には、現信号(直接音)と残響を形成する過去の信号の相関が低くなる性質があります。
- ※5 マルチステップ線形予測は、現在からNサンプル遡った過去の信号を用いて現在の信号を予測する手法です。本文に述べた残響除去技術では、Nを適切に定めることで、録音音声に含まれる残響の高精度な推定を可能にしました。

【本リリースに関する報道機関からの問い合わせ先】

NTT 広報室

TEL: 03-5205-5550

エヌ・ティ・ティ ラーニングシステムズ株式会社

総務部 広報室

担当 友部 芹澤 03-5440-2716

【本リリースに関する報道機関以外からの問い合わせ先】

NTT研究企画部門 プロデュース担当

担当 平松 03-5205-5765

【製品に関する問い合わせ先】

エヌ・ティ・ティ ラーニングシステムズ株式会社

映像制作事業部 メディアプロデュース部
担当 上谷 03-5440-5521

NTT ニュースリリース 

Copyright(c) 2008 日本電信電話株式会社