

学 位 論 文 題 名

画像音声伝送品質向上のための適応信号処理手法と その応用に関する研究

学位論文内容の要旨

本論文は適応信号処理を用いて、画像・音声伝送の品質を向上する研究の成果をまとめたものである。

近年、テレビ電話・会議など画像・音声を用いた通信が幅広く実用化されている。このような通信において、画像・音声の伝送品質向上は非常に重要であり、多くの研究開発が行われている。一方、品質向上の鍵となる信号処理では、LSI の高性能化などにより、データの統計的性質や外部環境の変化に柔軟に対応できる適応信号処理が、従来、実現困難であった画像分野にまで、その応用範囲を広げている。

本論文では、このような背景から、適応信号処理を用いた画像・音声伝送品質向上に関する研究成果を、ステレオ音声伝送品質向上を目的とした第一の研究と、離散コサイン変換を用いた画像符号化のブロック境界歪みを軽減する第二の研究に大別してまとめる。

まず、第一の研究である、ステレオ音声伝送の品質向上では、ステレオ音声符号化の品質向上とステレオ音響エコーキャンセラの打ち消し性能向上を図る。前者に関しては、MPEG オーディオなどステレオ楽音を対象とした符号化方式に関する研究が活発に行われている。しかしながら、楽音の符号化を第一の目的として開発されたものであるため、ステレオ音声伝送の観点からは、必ずしも最適な符号化とは言えない。また、後者のステレオ音響エコーキャンセラでは、ステレオ音声伝送のように左右の音声間に強い相関 (チャンネル間相関) が存在する場合、多次元適応フィルタの収束速度が大幅に劣化するといった課題がある。このため、従来方式として、左右に、独立した非線形歪みを与えて相関を弱める方式や、過去の複数サンプルにおけるチャンネル間相関の揺らぎを利用して適応フィルタの係数修正をアフィン変換で直交化する方式等が提案されている。しかしながら、前者は音声品質に劣化を与えるという課題があり、後者は、遠端発言者の変化に起因するチャンネル間相関の大きな変動を活用しようとする逆行行列演算量が膨大になるといった課題がある。

そこで、第一の研究では、会話型のステレオ音声伝送を前提とし、その音声生成系に着目することにより、上述の複数の課題を解決する。会話の大部分を占める単独発言時には、単一音源音声は左右のマイクロフォンに入力することにより、仮想的にチャンネル間の伝達関数を定義できる。本提案方式では、このチャンネル間伝達関数を適応信号処理の代表的な実現手法である適応フィルタを用いて推定し、活用する。まず、提案方式の基本となる疑似ステレオ音声符号化方式では、単独発言時のステレオ音声を、片側チャンネルの音声 (主音声) とチャンネル間伝達関数で符号化する。通常、発言者位置や室内音響特性の変動は音声そのものの変動より大幅に小さいため、他のチャンネルの音声 (副音声) を、そのまま符号化するより、チャンネル間伝達関数を符号化する方が、より少ない

情報量で符号化できる。しかしながら、この方式は、会話の大部分を占める単独発言では有効であるが、複数同時発言のある、完全なステレオ音声に、そのまま適用できない。そのため、提案方式では、疑似ステレオ音声符号化に、チャンネル間相関を除去した残差信号を符号化して追加することにより、完全なステレオ音声符号化に拡張する。この原理を用いたステレオ ADPCM 方式では、さらに、このチャンネル間伝達関数を別チャンネルで伝送することなく、残差信号伝送のみで伝送する。この方式の実験では 32kbps を主チャンネル、16kbps を副チャンネルに割り当て、チャンネル間相関除去により 6-10dB 程度セグメンタル SNR が改善することを確認した。

次に、このチャンネル間伝達関数をステレオ音響エコーキャンセラに適用し、前述の課題であるチャンネル間相関による収束速度の劣化を改善する方式について説明する。基本方式である疑似ステレオ音声用音響エコーキャンセラは、主音声のみを入力とするモノラル適応フィルタであり、副音声をチャンネル間伝達関数に主音声を入力して得られる出力信号とみなすことにより、主音声のみを入力する適応フィルタで音響エコーを抑圧できる。このエコーキャンセラは、発言者位置変化により、チャンネル間伝達関数が変化すると、対応する複数の推定チャンネル間伝達関数とともに方程式を解くことにより、真のステレオ・エコーパス特性を推定できる。しかしながら、この方式は、疑似ステレオ音声符号化同様、複数同時発言時などには適用できない。そこで、本研究では、多次元適応フィルタを用いたステレオ音響エコーキャンセラをエコー抑圧に用い、疑似ステレオ音声エコーキャンセラは、チャンネル間伝達関数の変化を利用したステレオ・エコーパス特性の推定動作のみを用いる。これにより、従来の最小二乗誤差やアフィン変換を用いたステレオ音響エコーキャンセラの収束が、遠端発言者の変化時点で飛躍的に改善する。

次に、第二の研究では、画像圧縮伸長処理の品質向上を目的した研究成果について述べる。JPEG や MPEG など国際標準方式に広く採用されている離散コサイン変換を用いた画像符号化方式では、帯域圧縮が大きい場合、中間値を用いる従来の逆量子化手法では、直流近傍の誤差によるブロック歪みが顕著になり画像品質を劣化させる。このため、従来方式では、復号後の画像のブロック境界に低域フィルタを適用することにより、ブロック境界の平坦化を図る。しかしながら、この場合、低域フィルタにより高域成分も損なわれるため、画像がぼけるという問題がある。また、低域フィルタの特性を画像の局所特性で変化させ、最適な高域抑圧を行う方式も最近提案されているが、処理が複雑になるといった課題がある。本研究では、離散コサイン変換係数の逆量子化値をパラメタとする適応フィルタを想定し、ブロック境界の不連続性の評価関数を MSDS (Mean Difference of Slope) として定義する。この MSDS を最小化する逆量子化値を、上述の適応フィルタの最適化問題として解くことによりブロック歪みを抑圧する。この方式では、ブロック歪みを引き起こす要因となる低域成分のみの離散コサイン変換係数を対象に最適な逆量子化パラメタを求めるため、画像がぼやける原因となる高域成分を保存でき、復号画像の品質を向上することができる。さらに、復号側で処理するため、通信の互換性には影響を与えない。

本論文では、第一章で、研究の背景や基本となる適応信号処理の概要を述べたのち、第二章でチャンネル間伝達関数推定を用いたステレオ音声伝送方式を、第三章では、チャンネル間伝達関数推定を用いたステレオ音響エコーキャンセラを提案する。第四章では、適応信号処理を用いた画像のブロック境界歪みを抑圧する方式を提案し、第五章で結びとする。

学位論文審査の要旨

主 査 教 授 長谷山 美 紀

副 査 教 授 山 本 強

副 査 教 授 荒 木 健 治

学 位 論 文 題 名

画像音声伝送品質向上のための適応信号処理手法と その応用に関する研究

本論文は適応信号処理を用いて、画像・音声伝送の品質を向上する研究の成果をまとめたものである。

近年、テレビ電話・会議など画像・音声を用いた通信が幅広く実用化されている。このような通信において、画像・音声の伝送品質向上は非常に重要であり、多くの研究開発が行われている。一方、品質向上の鍵となる信号処理では、LSI の高性能化などにより、データの統計的性質や外部環境の変化に柔軟に対応できる適応信号処理が、従来、実現困難であった画像分野にまで、その応用範囲を広げている。

本論文では、このような背景から、適応信号処理を用いた画像・音声伝送品質向上に関する研究成果を、ステレオ音声伝送品質向上を目的とした第一の研究と、離散コサイン変換を用いた画像符号化のブロック境界歪みを軽減する第二の研究に大別してまとめる。

まず、第一の研究である、ステレオ音声伝送の品質向上では、ステレオ音声符号化の品質向上とステレオ音響エコーキャンセラの打ち消し性能向上を図る。前者に関しては、MPEG オーディオなどステレオ楽音を対象とした符号化方式に関する研究が活発に行われている。しかしながら、楽音の符号化を第一の目的として開発されたものであるため、ステレオ音声伝送の観点からは、必ずしも最適な符号化とは言えない。また、後者のステレオ音響エコーキャンセラでは、ステレオ音声伝送のように左右の音声間に強い相関（チャンネル間相関）が存在する場合、多次元適応フィルタの収束速度が大幅に劣化するといった課題がある。このため、従来方式として、左右に、独立した非線形歪みを与えて相関を弱める方式や、過去の複数サンプルにおけるチャンネル間相関の揺らぎを利用して適応フィルタの係数修正をアフィン変換で直交化する方式等が提案されている。しかしながら、前者は音声品質に劣化を与えるという課題があり、後者は、遠端発言者の変化に起因するチャンネル間相関の大きな変動を活用しようとする逆行列演算量が膨大になるといった課題がある。

そこで、第一の研究では、会話型のステレオ音声伝送を前提とし、その音声生成系に着目することにより、上述の複数の課題を解決する。会話の大部分を占める単独発言時には、単一音源音声は左右のマイクロフォンに入力することにより、仮想的にチャンネル間の伝達関数を定義できる。本提案方式では、このチャンネル間伝達関数を適応信号処理の代表的な実現手法である適応フィルタを用いて推定し、活用する。まず、提案方式の基本となる疑似ステレオ音声符号化方式では、単独発言時のステレオ音声を、片側チャンネルの音声（主音声）とチャンネル間伝達関数で符号化する。通常、発言者位置や室内音響特性の変動は音声そのものの変動より大幅に小さいため、他のチャンネルの音声（副音声）を、そのまま符号化するより、チャンネル間伝達関数を符号化する方が、より少ない情報量

で符号化できる。しかしながら、この方式は、会話の大部分を占める単独発言では有効であるが、複数同時発言のある、完全なステレオ音声に、そのまま適用できない。そのため、提案方式では、疑似ステレオ音声符号化に、チャンネル間相関を除去した残差信号を符号化して追加することにより、完全なステレオ音声符号化に拡張する。この原理を用いたステレオ ADPCM 方式では、さらに、このチャンネル間伝達関数を別チャンネルで伝送することなく、残差信号伝送のみで伝送する。この方式の実験では 32kbps を主チャンネル、16kbps を副チャンネルに割り当て、チャンネル間相関除去により 6-10dB 程度セグメンタル SNR が改善することを確認した。

次に、このチャンネル間伝達関数をステレオ音響エコーキャンセラに適用し、前述の課題であるチャンネル間相関による収束速度の劣化を改善する方式について説明する。基本方式である疑似ステレオ音声用音響エコーキャンセラは、主音声のみを入力とするモノラル適応フィルタであり、副音声をチャンネル間伝達関数に主音声を入力して得られる出力信号とみなすことにより、主音声のみを入力する適応フィルタで音響エコーを抑圧できる。このエコーキャンセラは、発言者位置変化により、チャンネル間伝達関数が変化すると、対応する複数の推定チャンネル間伝達関数とともに方程式を解くことにより、真のステレオ・エコーパス特性を推定できる。しかしながら、この方式は、疑似ステレオ音声符号化同様、複数同時発言時などには適用できない。そこで、本研究では、多次元適応フィルタを用いたステレオ音響エコーキャンセラをエコー抑圧に用い、疑似ステレオ音声エコーキャンセラは、チャンネル間伝達関数の変化を利用したステレオ・エコーパス特性の推定動作のみに用いる。これにより、従来の最小二乗誤差やアフィン変換を用いたステレオ音響エコーキャンセラの収束が、遠端発言者の変化時点で飛躍的に改善する。

次に、第二の研究では、画像圧縮伸長処理の品質向上を目的した研究成果について述べる。JPEG や MPEG など国際標準方式に広く採用されている離散コサイン変換を用いた画像符号化方式では、帯域圧縮が大きい場合、中間値を用いる従来の逆量子化手法では、直流近傍の誤差によるブロック歪みが顕著になり画像品質を劣化させる。このため、従来方式では、復号後の画像のブロック境界に低域フィルタを適用することにより、ブロック境界の平坦化を図る。しかしながら、この場合、低域フィルタにより高域成分も損なわれるため、画像がぼけるという問題がある。また、低域フィルタの特性を画像の局所特性で変化させ、最適な高域抑圧を行う方式も最近提案されているが、処理が複雑になるといった課題がある。本研究では、離散コサイン変換係数の逆量子化値をパラメタとする適応フィルタを想定し、ブロック境界の不連続性の評価関数を MSDS (Mean Difference of Slope) として定義する。この MSDS を最小化する逆量子化値を、上述の適応フィルタの最適化問題として解くことによりブロック歪みを抑圧する。この方式では、ブロック歪みを引き起こす要因となる低域成分のみの離散コサイン変換係数を対象に最適な逆量子化パラメタを求めるため、画像がぼやける原因となる高域成分を保存でき、復号画像の品質を向上することができる。さらに、復号側で処理するため、通信の互換性には影響を与えない。

これを要するに、著者は、適応信号処理を用いた画像・音声伝送品質向上を目的として、ステレオ音声符号化品質とステレオ音響エコーキャンセラの打ち消し性能の向上を図り、さらに、画像符号化のブロック境界ひずみの軽減を実現したものであり、情報処理工学に貢献するところ大なるものがある。よって著者は、北海道大学博士(情報科学)の学位を授与される資格あるものと認める。