

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 16 日現在

機関番号：12612

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25730107

研究課題名(和文)高齢者のコミュニケーション意欲を促進するマルチメディア通信システムの開発

研究課題名(英文)Development of multimedia system to motivate elderly persons' communication

研究代表者

松本 光春(Matsumoto, Mitsuharu)

電気通信大学・情報理工学(系)研究科・准教授

研究者番号：70434305

交付決定額(研究期間全体):(直接経費) 3,200,000円

研究成果の概要(和文): 本研究では、超高齢社会における人と人とのコミュニケーションの支援を目標としたマルチメディア通信システムの開発を目指した。本研究にかかる成果として、画像処理分野においては、エッジ保存平滑化技術と自己商フィルタの特性を考慮した非線形フィルタの拡張について検討した。音響処理分野においては音声認識システムを評価システムとして陽に利用したパラメータ調整方法について検討し、その有効性を確認した。

研究成果の概要(英文): In this study, we aim to develop a system to help human-human communication for super-aging society. We developed edge-preserving filtering, which is robust for light fluctuation in image processing area. We also investigated the effectiveness of parameter setting using speech recognition system directly.

As results, we improved a filter combining edge-preserving filter and self-quotient filter. We also confirmed the effectiveness of the parameter setting using speech recognition system.

研究分野：知覚情報処理・知能ロボティクス

キーワード：知覚情報処理 画像処理 音響処理 マルチメディアシステム

1. 研究開始当初の背景

年を取ってもいつまでも若くありたい。それは人間にとって普遍的な願いである。しかしながら、加齢による身体的な衰えを止めることは容易ではない。高年齢化に伴い現れるしわやしみなどといった外見上の変化は、時に高齢者の他者とのコミュニケーションに対する意欲を失わせてしまう。また、加齢による視覚・聴覚の衰えは他者とのコミュニケーション自体を困難にする。このような高齢者のコミュニケーション支援システムとして、本研究ではテレビ電話システムに注目する。従来のテレビ電話システムでは、鮮明な画像や音声を如何に忠実に伝送するかに焦点が合わされてきた。これに対し、本研究ではたとえ見かけ上ではあっても高齢者をできるだけ若く、できるだけ生き生きと見せられるようなマルチメディア通信システムを構築し、その視覚・聴覚的な衰えを機械システムで支援することで、高齢者の他者とのコミュニケーションに対する意欲を高めることを目指す。

2. 研究の目的

本研究は、超高齢社会における人と人とのコミュニケーションの支援を目標とし、聴覚の衰えのため聞き取りにくい会話を補助するための音声認識システムによる補助対象者をより若々しく見せるためのしわ、シミの除去システムを搭載したマルチメディアシステムの開発を目指す。本目的の実現のため、対象領域のコントラストの違いによらないエッジ保存平滑化技術の開発や認識システムを陽に利用したフィルタリングシステムのパラメータ設定に関する技術開発、リアルタイム・マルチメディアシステムの開発などに取り組む。

3. 研究の方法

研究目標となっているマルチメディアシステムの開発をめざし、複数の研究課題に取り組んだ。

画像処理系の研究については申請者が開発した自己商参照非線形フィルタについてそのフィルタ性能の改良を行った。

申請者が開発した自己商参照非線形フィルタは照明変動のある画像に対するエッジ保存平滑化が可能であるが、グレースケールの画像に対する性能のみ評価されており、カラー画像への適応の評価が行われていなかった。研究期間には、この問題を解決し、カラー画像への適用可能性について調査した。

図1にフィルタ処理の流れを示す。

図1に示すように提案手法では、画像の成分を強度成分とカラー成分とに分離し、強度成分について自己商フィルタを適用する。次に参照画像として自己商フィルタを用いたエッジ保存平滑化フィルタをかける。最後に得られたフィルタ出力に対し、カラー画像と合成することで出力画像を得る。

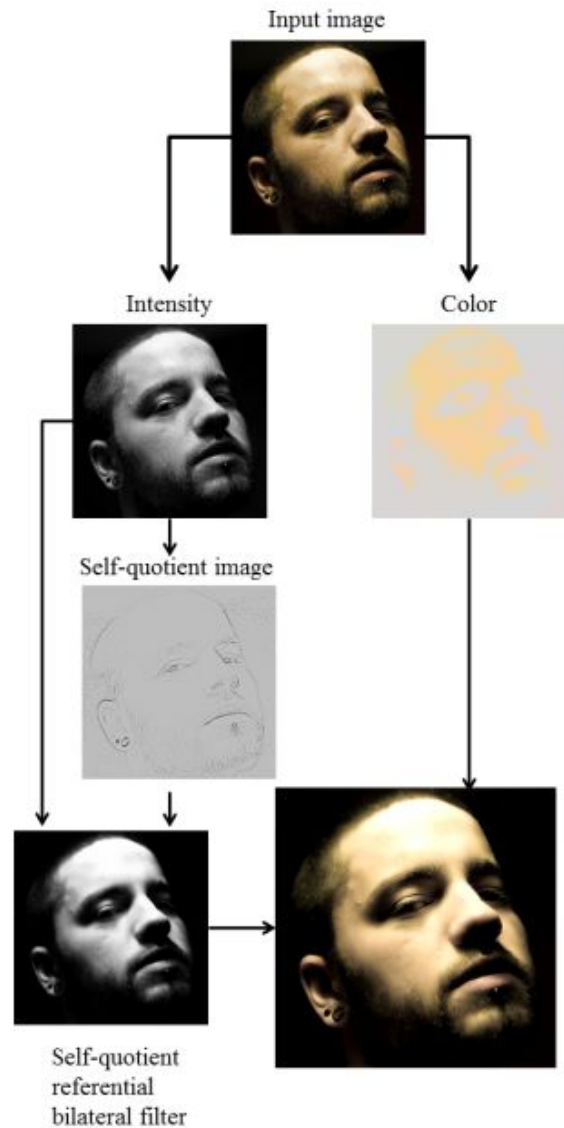


図1 フィルタリングの流れ

音響系の研究としては特にパラメータ設定に取り組む。雑音除去法にはスペクトルサブトラクション法や α -フィルタなど様々な手法が存在するが、多くの方法では状況に適したパラメータの設定が必要であり、通常は目的音声と雑音除去後の音声間の波形の誤差が小さくなるようなパラメータを用いる。しかし、実際の利用場面では目的音声の波形は通常既知ではなく、パラメータの最適化は必ずしも容易ではない。また、誤差が最小であったとしても認識率が良いとは限らないため、音声認識との連携を考えたとき、実用的であるとは言えない。これに対し、申請者は目的画像が未知の場合における画像雑音除去の方法として、意味内容を与えることで認識率を測定しその値が良くなるような雑音除去のパラメータを設定する、画像処理での枠組みを提案している。本研究では、この手法の原理を音声に適用し、認識すべき文字列に対して認識率が最良

になるようなパラメータを設定することで前述の問題点を解決する。雑音除去パラメータを事前実験により、あらかじめ決めていたとしても、そのパラメータが実環境の中で必ずしも最適であるとは限らない。そこで、本研究では、提案する枠組みを基に環境に合わせて最適なパラメータを設定できるようなシステムの構築を目指す。そのとき設定されるパラメータは雑音の大きさの変化に耐性があるようなものであることを目指す。

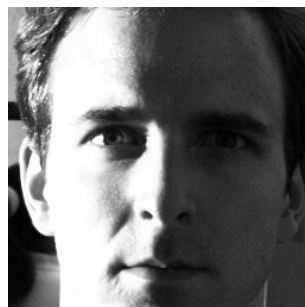
4. 研究成果

まず、リアルタイム処理に向けた研究の準備段階として、Microsoft Kinect を利用した画像・音響システムのための開発環境の整備やプログラム開発を行うとともに、特に画像処理系について、動画内に撮像された人物のしわ、しみ除去のための技術要素の開発に取り組んだ。

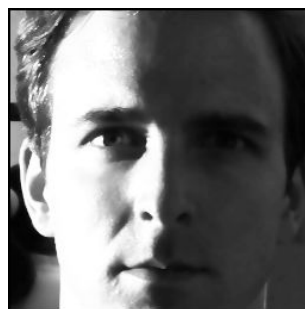
画像処理系研究について、申請者はこれまでに照明変動を含む画像に対する効率的なしわやシミの除去システムとして、エッジ保存平滑化技術と自己商フィルタの特性を考慮した非線形フィルタの開発を行ってきた。本研究では特に上記の非線形フィルタのカラー画像への適用可能性について検討した。モノクロ画像に対する結果を図2、図3に、カラー画像に対する結果を図4、図5にそれぞれ示す。図2から図5に示すように従来手法では高コントラスト領域におけるエッジ保存平滑化はうまくいっているものの低コントラスト領域においてはエッジ保存平滑化がうまく働いておらず、画像がぼけてしまっていることが確認できる。これに対し、提案手法ではモノクロ画像、カラー画像双方に対し、高コントラスト領域だけでなく、低コントラスト領域においてもエッジ保存平滑化が実現できていることが確認できた。結果として照明変動のあるカラー画像にも対応可能なエッジ保存平滑化フィルタの開発に成功し、1枚の画像のみを用いたしわ、しみ除去システムの実現可能性を確認した。一方、音響系システムにおいては環境内の雑音変動に頑健な音声認識システムの実現のため、音声認識システムを評価システムとして陽に利用したフィルタリングのパラメータ調整方法について検討した。

実験条件

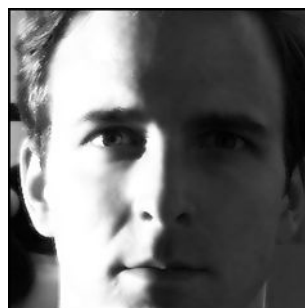
実験環境を図6に示す。また、実験条件を表1に示す。ロボットへの指示音声は「まえ」と「うしろ」の2語とし、認識システムにはこの2語を登録した。実験時の指示音声には20代女性、20代男性各1名の発話を録音して用いた。ロボットは起動状態を初期条件とした。そのため実験中は常に60dB(Kinect付近測定値)のロボットの動作音がしている状態だった。環境の変化要素としての雑音にはホワイトノイズと換気扇の音を録音した



(1)入力画像

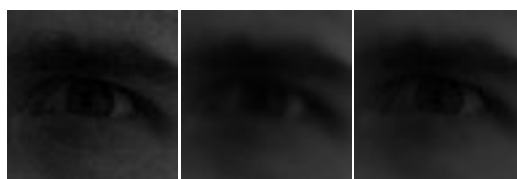


(2)既存手法の出力



(3)提案手法の出力

図2 提案手法と既存手法の比較
(モノクロ画像)

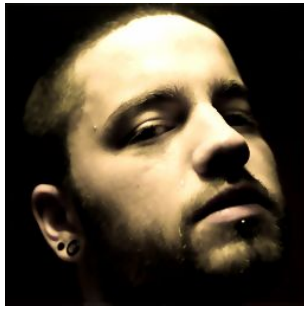


(1)入力画像 (2)既存手法 (3)提案手法
図3 提案手法と既存手法の比較(拡大図)

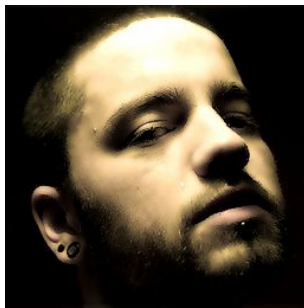
一般雑音を用い、音量を60dB、65dB、70dB、72dBと変化させた。また、雑音の種類も変化させ、パラメータの雑音種変化耐性も確認した。ここで言う雑音とは、図6中のスピーカーBから発生させるものである。加えて、汎用性を評価するため学習時と認識時の話者を変え、一度(一人)のパラメータ設定で複数人が使用可能であるかも確認した。以上を組み合わせた実験パターンを表2に示す。



(1)入力画像

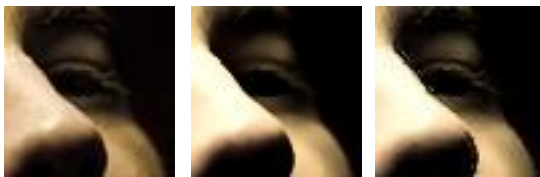


(2)既存手法



(3)提案手法

図4 提案手法と既存手法の比較 (カラー画像)



(1)入力画像 (2)既存手法 (3)提案手法
図5 提案手法と既存手法の比較 (拡大図)

表2に示した実験パターンについて、パラメータを固定した場合及び提案手法で決定したパラメータを用いた場合で実験を行い、結果を比較した。

パターンごとの詳細な条件を説明する。パターン(ア)では、ロボットの動作音のみがしている状況(初期条件)でどの程度認識可能であるかを調べた。認識には話者20代女性(以下、Aさん)の目的信号「まえ」と「うしろ」を用いた。パターン(イ)では、提案パラメータが話者の変化に耐性があるかを調べた。パターン(ア)と同様にロボットの動作音のみがしている状況で実験を行った。

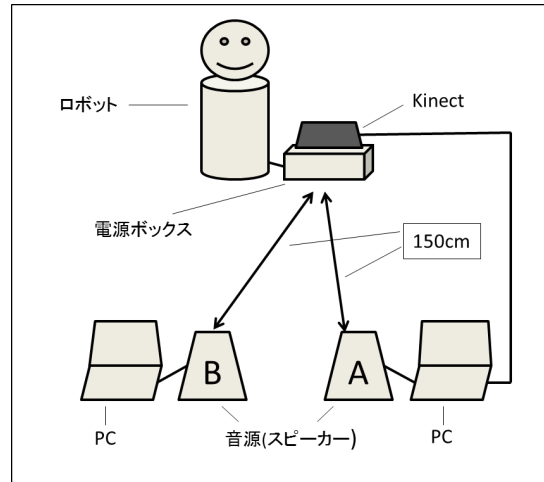


図6 実験環境

表1 実験条件

目的信号	ロボットへの指示音声 (まえ, うしろ)
話者	20代女性(1名) 20代男性(1名)
雑音信号	ホワイトノイズ, 換気扇
雑音信号の大きさ[dB]	60, 65, 70, 72
雑音除去パラメータ	0 ~ 2.0(0.1きざみ)

表2 実験パターン表

	雑音	雑音量変化	雑音種変化	話者変化
パターン(ア)	×			×
パターン(イ)	×			
パターン(ウ)			×	
パターン(エ)				

表3 パターン(エ), 話者Aにおける認識結果

話者	Aさん							
	まえ				うしろ			
目的信号								
雑音量[dB]	60	65	70	72	60	65	70	72
提案手法(A = 1.4)								×
A = 1							×	×

学習には話者Aさんの目的信号「うしろ」を用い、認識には話者20代男性(以下、Bさん)の目的信号「まえ」と「うしろ」を用いた。パターン(ウ)では、提案パラメータが雑音量の変化に耐性があるかを調べた。ロボットの動作音に加え、雑音が発生している状況で実験を行った。雑音にはホワイトノイズを用いた。雑音量は、学習時は65dBとし、認識時は60dB, 65dB, 70dB, 72dBと変化させた。また、学習には話者Aさんの目的信号「うしろ」を用い、認識には話者AさんとBさんの目的信号「まえ」と「うしろ」を用いた。パターン(エ)では、提案パラメータが雑音の種類かつ雑音量の変化に耐性があるかを調べた。ロボットの動作音に加え、雑音が発生している状況で実験を行った。学習時の雑音には65dBのホワイトノイズを用いた。認識時の雑音には録音した換気扇の音を用い、60dB, 65dB, 70dB, 72dBと変化させた。また、

学習には話者 A さんの目的信号「うしろ」を用い、認識には話者 A さんと B さんの目的信号「まえ」と「うしろ」を用いた。

実験はパターン(A)~パターン(E)を順に行った。各パターンでは最初に雑音の記憶とパラメータ設定の学習フェーズを行い、その後各パターンの条件に従って認識フェーズを行った。固定パラメータを用いる際はプログラム中で値を指定した。ここでは、雑音、及び、話者が変化するパターン(E)における実験結果の一部を表3に示す。

表3に示すように固定パラメータでは認識できなかった雑音量の領域においても提案手法では認識できることが確認でき、提案手法が、学習時と認識時で雑音量が変化した場合の耐性を持つことが確認できた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計2件)

[1] Mitsuharu Matsumoto, "Self-quotient Referential Bilateral Filter and Its Application to Real-Time Face Beautification System," International Journal of Computer and Electrical Engineering, pp.91-100, Vol.7, No.2, 2015 (査読有)

[2] Mitsuharu Matsumoto, "Light Robust Edge Preserving Smoothing using Self-quotient Referential ϵ -Filter," Journal of Intelligent Computing, pp.163-175, Vol.4, No.4, 2013. (査読有)

[学会発表](計4件)

[3] Mitsuharu Matsumoto, "Parameter tuning of robot audition using speech recognition system as evaluation function," The 2015 International Conference on advanced technologies for communications, pp.260-263, Ho Chi Minh, Vietnam, Oct. 14 - Oct. 16, 2015. (査読有)

[4] Mitsuharu Matsumoto, "Self-quotient Referential Bilateral Filter and Its Application to Real-Time Face Beautification System," 2014 International Conference on Computer Science and Information Technology (ICCSIT2014), Barcelona, Spain, Dec. 22 - Dec. 24, 2014. (査読有)

[5] 平山奈央子, 松本光春, "音声認識結果を規範とした雑音除去システムのパラメータ設定," 計測自動制御学会 システム・情報部門 学術講演会 2014, pp.10-12, 岡山, 2014年11月21日-23日. (査読無)

[6] Mitsuharu Matsumoto, "On face beautification of color images under light varying condition using self-quotient referential ϵ -filter," Third International

Conference on Innovative computing technology (Intech2013), pp.485-490, London, United Kingdom, Aug. 29 - Aug. 31, 2013. (査読有)

6. 研究組織

(1) 研究代表者

松本 光春 (MATSUMOTO, Mitsuharu)

電気通信大学・情報理工学研究科・准教授

研究者番号: 70434305