様式Ｃ－１９

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成２４年５月１８日現在

機関番号：３２６８９
研究種目：若手研究（B）
研究期間：2009～2011
課題番号：21700207
研究課題名（和文）　２次元画像・音響特徴量の選択的利用による３次元頭部形状推定に関する研究
研究課題名（英文）　A study of 3D head geometry estimation based on selective use of 2D image and acoustic features
研究代表者
前島　謙宣（ MAEJIMA　AKINOBU ）
早稲田大学・ＩＴ研究機構・研究員
研究者番号：50454046

研究成果の概要（和文）
2 次元画像・音響特徴量を選択・相補的に利用した 3 次元頭部形状推定手法の実現を目的とし，以下の 4 つの研究成果を得た．(1) 109 人分のレンジスキャン・日本語 5 母音の音声データを含むデータベースを構築した．(2) 顔形状の事前知識を用いて特徴点の情報から個人の 3 次元顔形状を高速推定する手法を実現した．(3) 音響特徴量と 3 次元頭部形状をニューラルネットワークにより対応付けし，単母音音声から 3 次元頭部形状を推定する枠組みを構築した．(4) 3 次元頭部モデルを効率的に生成するためのレンジスキャンデータに対する非剛体レジストレーション手法を実現し論文として発表した．

研究成果の概要（英文）：
We aimed to realize a 3D geometry estimation method based on selective use of 2D image and acoustic features and obtained four results as follows. (1) We constructed a database which contains 109 range scans and 5 Japanese vowel speeches. (2) We developed a method for fast 3D geometry estimation from a face photograph using prior knowledge of 3D face geometries. (3) We developed a framework to estimate 3D head geometry from a vowel speech using a neural network trained using a pair of acoustic features and 3D head geometries. (4) We developed a registration method for non-rigid object to effectively generate a 3D head model from a range scan, and several achievements have published as journals.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
22 年度	500,000	150,000	650,000
23 年度	500,000	150,000	650,000
年度			
年度			
年度			
総　計	1,000,000	300,000	1,300,000

研究分野：総合領域
科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス
キーワード：画像特徴量，音響特徴量，顔特徴データベース，正準相関分析，ニューラルネットワーク，3 次元頭部形状推定

１．研究開始当初の背景

近年，最先端の映像制作の現場では本人と見間違うほどリアルな CG 顔を作成することが可能となってきているが，そのためには高額

で大がかりな設備，人手，時間が必要とされる場合が多い．このため，特別な機器を用いず，安価，かつ簡易的に CG 顔を生成する手法が求められている．

２．研究の目的

本研究は，このような問題を解決するために，顔画像から得られる画像特徴量，あるいは音声から得られる音響特徴量を選択・相補的に利用して，CG 顔の生成に必要な 3 次元頭部形状の推定を目的とした研究であり，この目的を達成するために，以下に示す 5 つの研究を行った．

(1) 個人の顔のレンジスキャンデータと音声を収録した顔特徴データベースの構築

(2) 3 次元顔モデルのデータベースから抽出した事前知識を用い，顔画像中の特徴点情報から 3 次元顔形状を推定する手法の研究

(3) 音響特徴量と 3 次元頭部形状との関連性を利用した，音響特徴量単独の情報から 3 次元頭部形状を推定する研究

(4) (2),(3)の素となる 3 次元頭部モデルを効率的に生成するための，レンジスキャンデータに対する非剛体レジストレーション手法に関する研究

(5) 顔画像から得られる画像特徴量，あるいは音声から得られる音響特徴量を選択・相補的に利用して，3 次元頭部形状を推定する手法の検討

３．研究の方法

(1)については，まず顔撮影用のレンジスキャナ Danae 100SP を利用して，被験者の顔のレンジスキャンデータを撮影した．次に，被験者が日本語 5 母音を連続発声した時の音声を標本化周波数 48000[Hz],量子化ビット幅 16bit で収録し，音声を母音毎にセグメンテーションした上で，各トラックを 16000[Hz]にダウンサンプリングした．これにより顔特徴データベースを構築した．また加えて，推定結果との比較用に，被験者の正面・側面顔画像データを収録した．

(2)に関しては，顔モデルのデータベースから構築される顔変形モデルと，混合ガウス分布で表される顔形状分布に基づく尤度制約を利用し，顔画像から自動検出される特徴点と顔変形モデル上の対応頂点とのユークリッド距離が最少かつモデルパラメータの尤度が最大となるようなエネルギー最小化問題を解くことで，顔画像 1 枚から画像中の人物の顔形状を推定する研究を行った．また，推定結果と本人のレンジスキャンデータの比較により提案手法の推定精度を評価した．

(3)については，個人を表す音響特徴として，基本周波数，メル周波数ケプストラム係数，Relative Spectra について検討し，正準相関分析，ニューラルネットワークを用いた 3 次元頭部形状とのマッピングに基づく，3 次元頭部形状推定手法の研究を行った．また，クロスバリデーションテストを通じて形状推定精度の評価を行った．

(4)に関しては，最適化局所アフィン変換に基づきレンジスキャンデータから 3 次元頭部モデルを自動生成する手法，および Radial Basis Functions とエネルギー最小化に基づく非剛体レジストレーション手法を組み合わせたレンジスキャンデータから半自動的に 3 次元頭部モデルを生成する手法の研究を行った．

(5)については，(2),(3)の結果に基づき，また 3 次元形状の推定に有効な画像特徴量について調査し，さらに画像・音響特徴量を組み合わせた 3 次元頭部形状推定の枠組みについて検討した．

４．研究成果

(1)については，109 人（男性 71 人，女性 38 人）のレンジスキャンデータ，日本語 5 母音の音声データ，正面・側面顔画像データから成る顔特徴データベースを構築することができた．

(2)に関しては，自動検出された顔特徴点と顔変形モデル上の対応頂点とのユークリッド距離が最少かつモデルパラメータの尤度が最大となるようなエネルギー最小化問題を解くことにより，正面顔画像 1 枚から顔画像中の人物の顔形状を推定することができた．実際に，20 人の被験者について，推定された形状と本人のレンジスキャンデータとの推定誤差を評価したところ，2.1[㎜]の精度で 3 次元顔モデルを自動生成可能であることが分かった（図 1）．

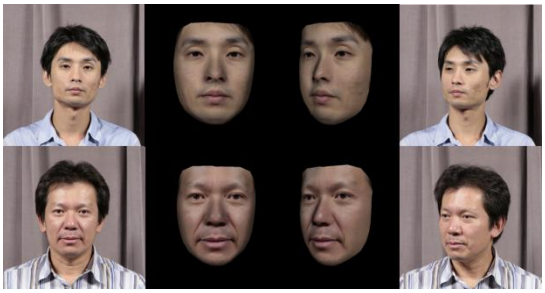

図 1. 推定結果の例（左から，入力顔画像，推定結果，推定結果・被験者を 30° 方向から眺めた様子）

また提案手法に基づき，エンタテインメントシステム「アクティブスナップショット」を開発し(図 2)，第 13 回 文化庁メディア芸術祭協賛展，先端技術ショーケース‘10 にて一般向けに展示を行い，実環境における提案手法の有効性を確かめた．

図 2. アクティブスナップショットにおけるアニメーションのフレーム例

(3)については，話者の口腔の音響特徴を表すことができるメル周波数ケプストラム係数とその 1 次，2 次差分係数を特徴量とし，正準相関分析およびフィードフォワード型のニューラルネットワークを用いて 3 次元頭部形状とマッピングした．顔特徴データベースに収録されている 70 名分の単母音「あ」のデータを用いた 10 ホールドのクロスバリデーションテストの結果から，ニューラルネットを用いた場合に音声単独から 6.7[mm]の精度誤差で頭部の形状を推定できることが分かった．図 3 に正解形状と推定された形状の例を示す．しかしながら，データベース中に存在する他人物の頭部形状をランダムに 20 個選択し，それぞれを推定結果とみなして正解形状と比較した場合の平均誤差が 7.4[mm]であったことから，マッピングすべき領域の絞り込みと，多くのサンプルを用いた更なる検証が必要であることが分かった．

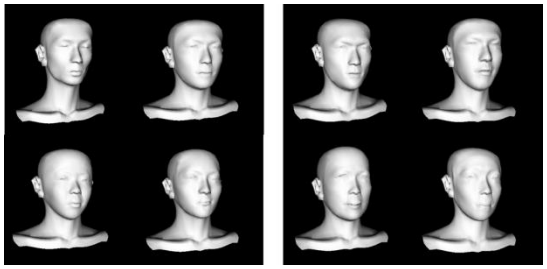

図 3. 推定結果の例

(各組左：正解形状，右：推定された形状)

(4)について，最適化局所アフィン変換に基づく自動生成手法では約 10 秒，Radial Basis Functions とエネルギー最小化に基づく半自動生成手法では，手作業による特徴点の指定を含めて 2 分程度で 3 次元頭部モデルを生成できるようになった．両手法により顔のレンジスキャンデータから効率的に 3 次元頭部モデルを生成することが可能になった(図 4)．

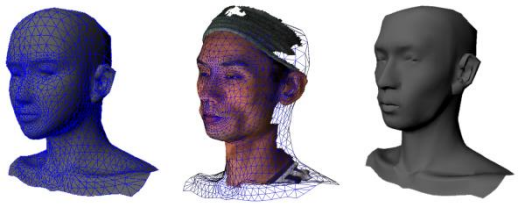

図 4. 頭部モデルの生成結果

(左：テンプレートメッシュ，中央：レジストレーションの様子，右：結果)

(5)に関しては，いくつかの知見は得たが，実際に 3 次元頭部形状推定を行うためには，音響・画像特徴量の選択，3 次元頭部形状推定における両特徴量の重みの設定方法，推定された 3 次元頭部形状に対する印象の評価について更なる研究が必要であり，これらが今後の課題として残った．

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) 前島 謙宣，“人物頭部モデル自動生成システムの実現：最適化局所アフィン変換に基づく人物頭部モデルの自動生成”，画像ラボ，21 巻，(2010)，29-34
(2) 前島 謙宣，森島 繁生，“最適化局所アフィン変換に基づく正面顔レンジスキャンデータからの頭部モデル自動生成”，画像電子学会論文誌，38 巻，(2009)，404-413，査読あり

〔学会発表〕（計 17 件）

(1) Akinobu Maejima, Shigeo Morishima, “3D Human Head Geometry Estimation from a Speech”, ACM SIGGRAPH 2012, 2012.8, Los Angeles, USA
(2) Tomoya Hara, Hiroyuki Kubo, Akinobu Maejima, Shigeo Morishima, “Fast-Automatic 3D Face Generation Using a Single Video Camera”, ACM SIGGRAPH 2012, 2012.8, Los Angeles, USA
(3) 前島 謙宣，森島 繁生，“単母音に含まれる音響特徴からの 3 次元頭部形状推定に関する一検討”，Visual Computing/グラフィックスと CAD 合同シンポジウム 2012，2012.6.24，早稲田大学
(4) 原 朋也，久保 尋之，前島 謙宣，森島 繁生，“事前知識と Structure-from-Motion

を併用した 1台のビデオ画像からの 3 次元顔モデル高速自動生成手法”, Visual Computing/グラフィックスと CAD 合同シンポジウム 2012, 2012.6.23, 早稲田大学
(5) Tomoya Hara, Akinobu Maejima, Shigeo Morishima, “Automatic 3D Face Generation from Video with Sparse Point Constraint and Dense Deformable Model”, ACM SIGGRAPHASIA 2011, 2011.12.14, China, Hong Kong
(6) 前島 謙宣, “PC 上の顔：デジタルフェイスモデリングの現状と今後”, 日本顔学会フォーラム顔学 2011, 2011.9.24, 日本歯科大学新潟生命歯学部(招待講演)
(7) Akinobu Maejima, Hiroyuki Kubo, Shigeo Morishima, “Realistic Facial Animation by Automatic Individual Head Modeling and Facial Muscle Adjustment”, HCI International 2011, 2011.7.14, USA, Hilton Orlando, Bonnet Creek
(8) Akinobu Maejima, Hiroto Yarimizu, Hiroyuki Kubo, Shigeo Morishima, “Automatic Generation of Head Models and Facial Animations Considering Personal Characteristics”, Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology (VRST2010), 2010.11.22, Hong Kong PolyU
(9) Akinobu Maejima, Hiroyuki Kubo, Shigeo Morishima, “Facial Animation Reflecting Personal Characteristics by Automatic Head Modeling and Facial Muscle Adjustment”, 2010.10.27, 明治大学
(10) 前島 謙宣, 久保 尋之, 森島 繁生, “アクティブスナップショット”, 第 15 回日本顔学会フォーラム顔学 2010, 2010.10.23, 東京歯科大学
(11) 前島 謙宣, 森島 繁生, “顔写真 1 枚からの 3 次元顔モデル高速自動生成”, CEDEC2010(CESA デベロッパーズカンファレンス 2010), 2010.8.31, パシフィコ横浜会議センター
(12) 前島 謙宣, 森島 繁生, “顔変形モデルと顔形状分布制約に基づく単一顔画像からの 3 次元顔モデル高速自動生成”, 画像の認識・理解シンポジウム(MIRU2010), 2010.7.28, 釧路市観光国際交流センター
(13) 前島 謙宣, 森島 繁生, “顔形状事前知識に基づく顔画像からの 3 次元顔モデル高速自動生成”, SSII2010 第 16 回画像センシングシンポジウム, 2010.6.11, パシフィコ横浜
(14) 前島 謙宣, 森島 繁生,“スナップ写真からの 3 次元顔モデル高速自動生成”, 電子情報通信学会 PRMU 研究会, 2010.3.15, 鹿児島大学
(15) Akinobu Maejima, Shigeo Morishima, “Human Head Modeling based on Fast Automatic Mesh Completion”, ACM SIGGRAPHASIA2009, 2009.12.18, パシフィコ横浜
(16) 前島 謙宣, 森島 繁生, “事前知識に基づく正面顔画像からの 3 次元形状復元”, ビジュアルコンピューティングワークショップ 2009, 2009.11.13, 辰口温泉
(17) Akinobu Maejima, Takanori Suzuki, Hiroto Yarimizu, Hiroyuki Kubo, Shigeo Morishima, “An individual character head modeling for instant casting”, The 9th Asian Conference on Computer Vision, Invited workshop1: vision based human modeling and synthesis, 2009.9.24, Xian, China (招待講演)

〔その他〕
ホームページ等

ホームページ
https://sites.google.com/site/maejimalaboratory

Youtube: DigiInfo Technology from Japan：Fast, autmatic generation of 3D face models from photos
http://www.youtube.com/watch?v=EGMTWL9QihM

第 13 回 文化庁メディア芸術祭協賛展, 先端技術ショーケース‘10
http://www.media.jst.go.jp/research_result/sentan10/leaflet.pdf

６．研究組織
(1)研究代表者
前島 謙宣 (AKINOBU MAEJIMA)
早稲田大学・IT 研究機構・研究員
研究者番号：50454046

(2)研究分担者
( )

研究者番号：

(3)連携研究者
( )

研究者番号：