

実時間における写実的な 3次元顔表情のクローニング

Real-Time and Realistic 3D Facial Expression Cloning

Samsung Advanced Institute of Technology,

Multimedia Lab

Jung-Bae Kim

September 10, 2010

■ 家庭内における3次元複合現実感 – リアルな3次元複合現実感



バーチャル・スポーツ

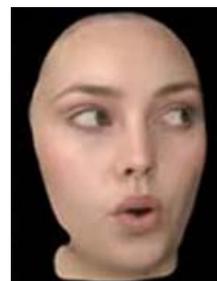
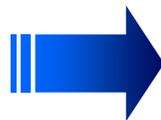
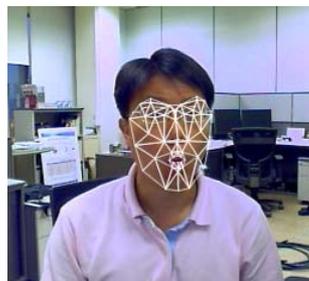


バーチャル・エンタテインメント



バーチャル・コマース

■ 目標:ユーザーのアバターに自然な表情のクローンを作成すること



問題の定義: 従来のクローニング手法の限界

- 手法 1: 映画のCGで利用される高精度な表情のクローニング
→ 実時間の処理ができず、面倒



Vicon 社のモーションキャプチャ装置
(高価な赤外線カメラ12台が用いられる。顔には約100個の赤外線マーカーが貼り付けられる。内部ノイズの除去処理に数日かかる。)

- 手法 2: テレビ会議のための実時間での表情クローニング
→ 低精度

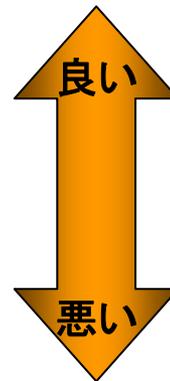
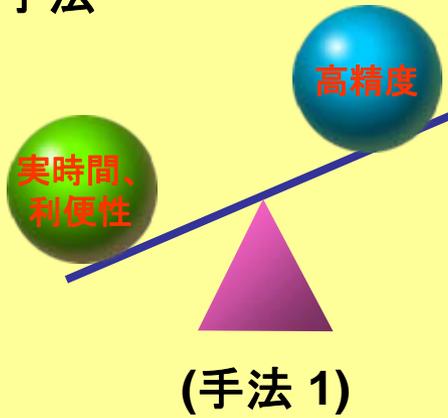


Logitech のテレビ会議S/W
(正面からの対称な表情)

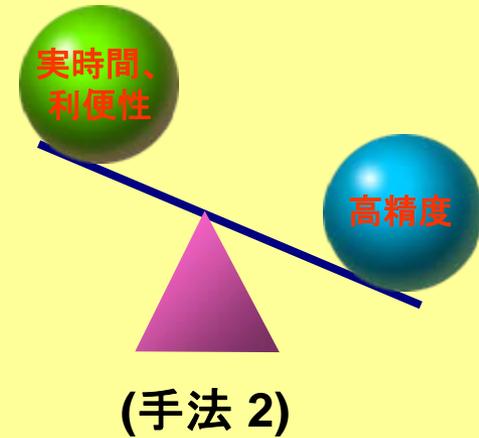
技術的矛盾



TC 1 : モーションキャプチャベースの手法



TC 2 : ビジョンベースの手法

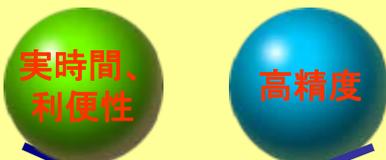


主 IFR

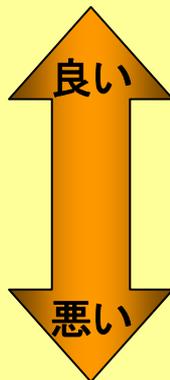


副 IFR

IFR = ビジョンベースの手法の利点
+ モーションキャプチャベースの手法の利点



(新しい手法)

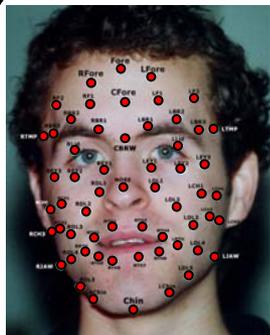


副-IFR1

<顔> それ自体が行う
<表情制御ポイント上にマーカ-を置くか取り去る>

副-IFR2

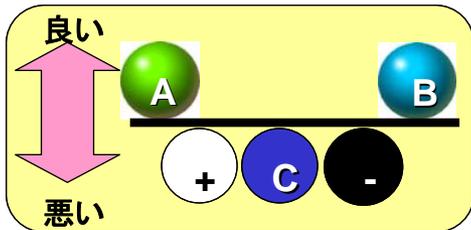
<システム> それ自体が決定する
<精度の度合い>



顔のマーカ-

- 存在する →
 - **高精度**: ユーザー間の差異、3次元的な頭部の回転、表情の変化を追跡できるようにする。
- 存在しない →
 - **実時間**: 後処理を必要としない。
 - **利便性**: カメラは1台だけ使用する。面倒なマーカ-を貼り付ける必要がない。

■ 物理的矛盾



<精度よく追跡する>ためには、
<変化に関する自由度>は **<大きく>** なければならず、
<実時間処理>のためには、
<変化に関する自由度>は **<小さく>** なければならない。

→ 変化に関する自由度: 別人物間の差異、同一人物における変化

(頭部の回転、表情)

■ 時間による分離

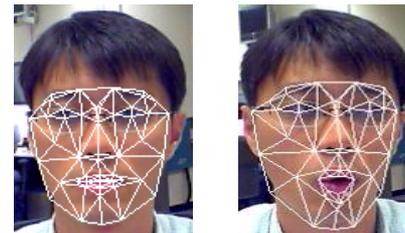
カラー画像



一般的な3次元顔モデルの
フィッティング

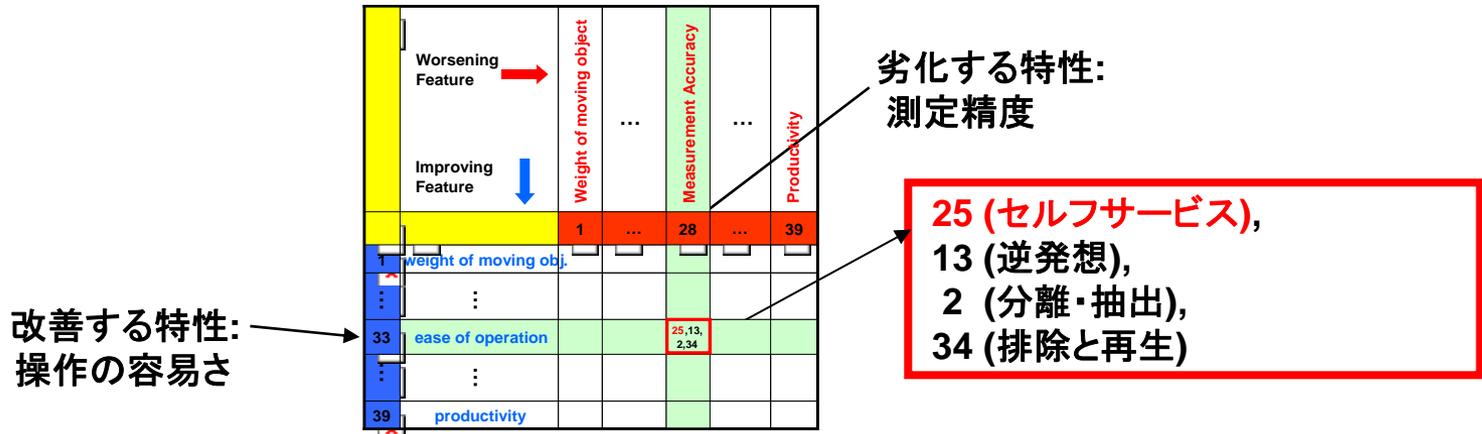


個人化された3次元表情モデルの追跡

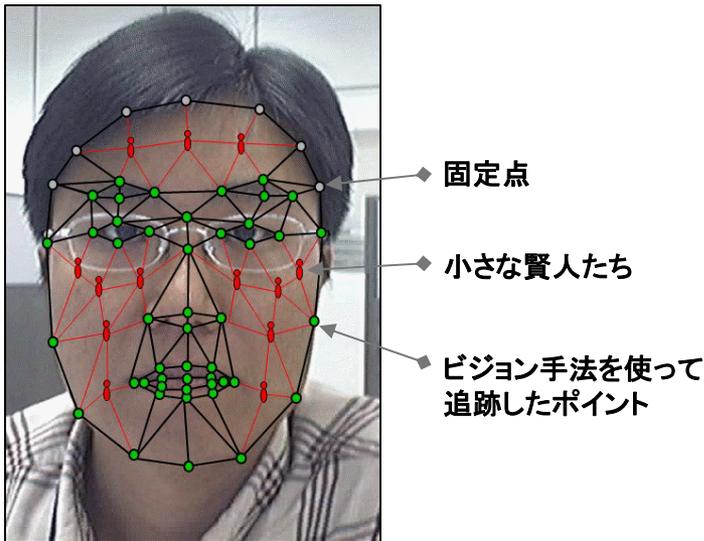


**最初の段階においては、別人物間における差異だけを考慮し、
その後に、同一人物における変化を考える。**

■ 矛盾マトリクスの利用



■ 発明原理 25 (セルフサービス), 小さな賢人たちモデル

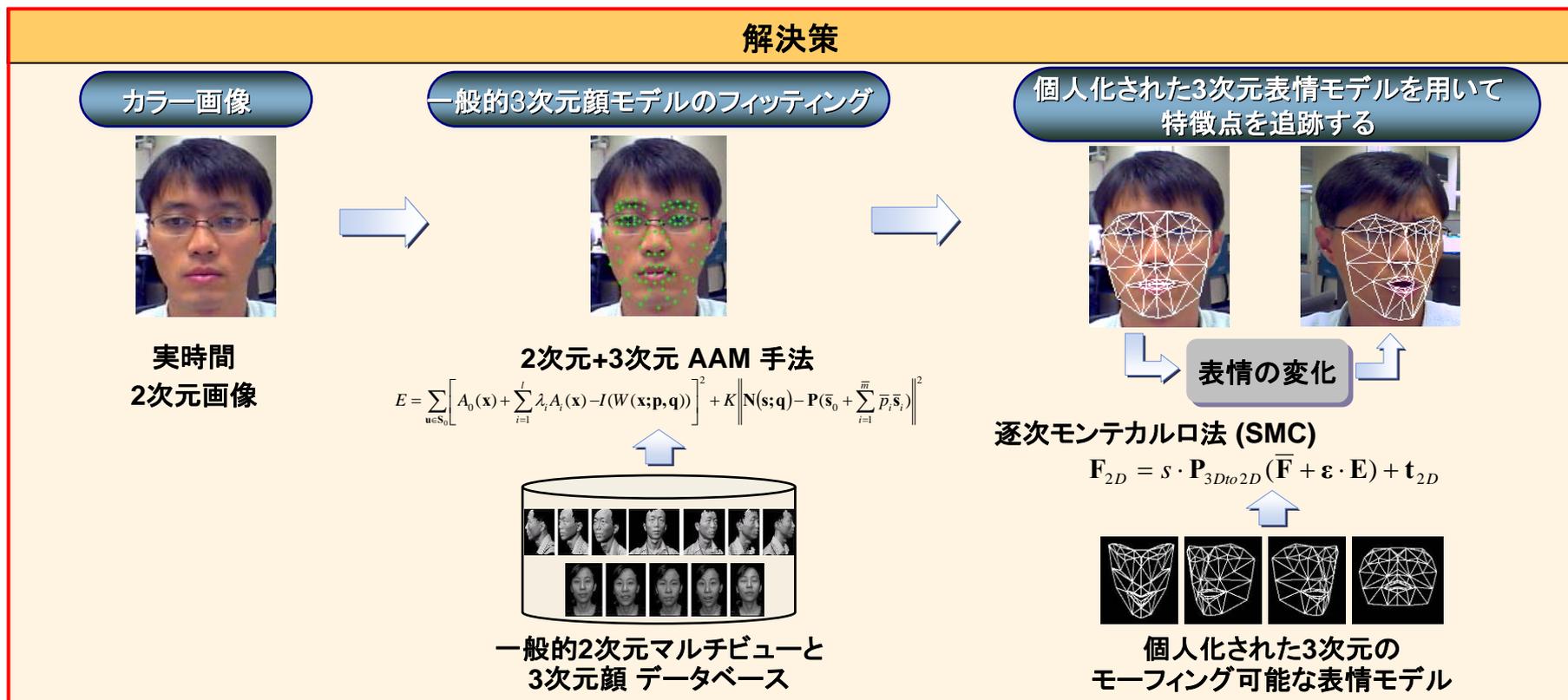
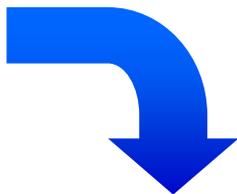
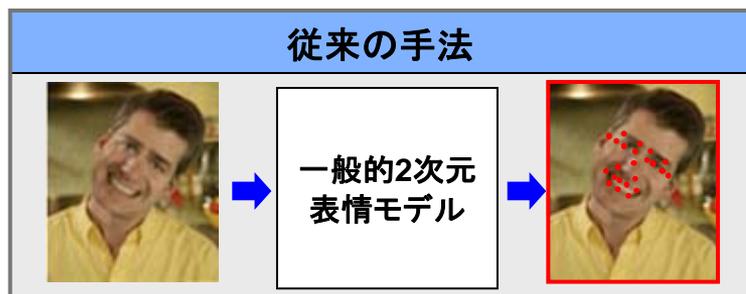


システムは、小さな賢人たち(筋モデル)を使って、額や頬に表情制御ポイントの動きを発生させる。

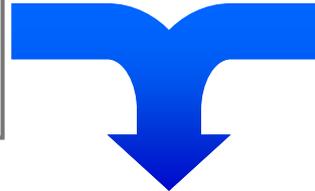
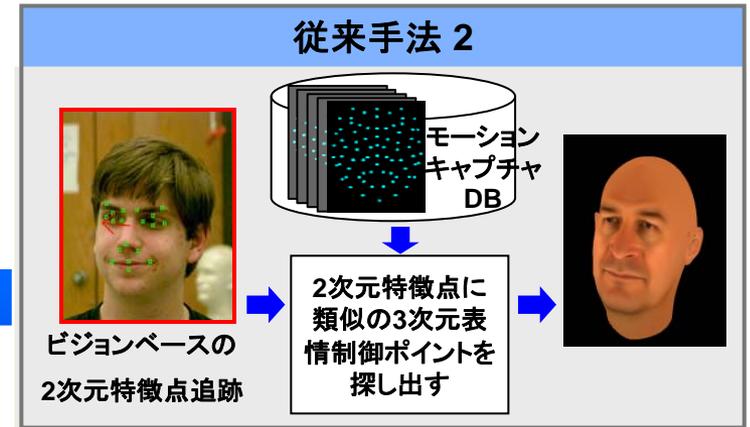


筋モデルは、モーションキャプチャデータに基づいてつくられる

解決策 (1/2): 個人化された3次元表情モデルを用いて特徴点を追跡する

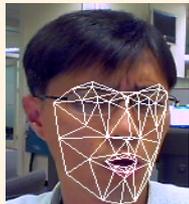


解決策 (2/2): ビジョンモデルと筋モデルに基づいて表情制御ポイントを生産させる



解決策

個人化された3次元表情モデルを用いて特徴点を追跡する



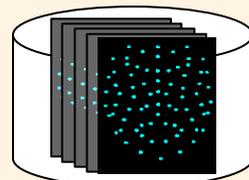
3次元特徴点

筋モデルに基づいて表情制御ポイントを生産させる

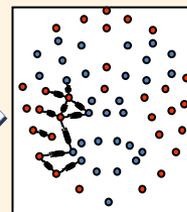
表情制御ポイント = 特徴点 + 筋モデル

$$x'_n \approx \frac{\sum_m k_m \left(x'_m - \|L_{mn}^0\| \frac{L_{mn}^{l-1}}{\|L_{mn}^{l-1}\|} \right)}{\sum_m k_m}$$

表情のレンダリング



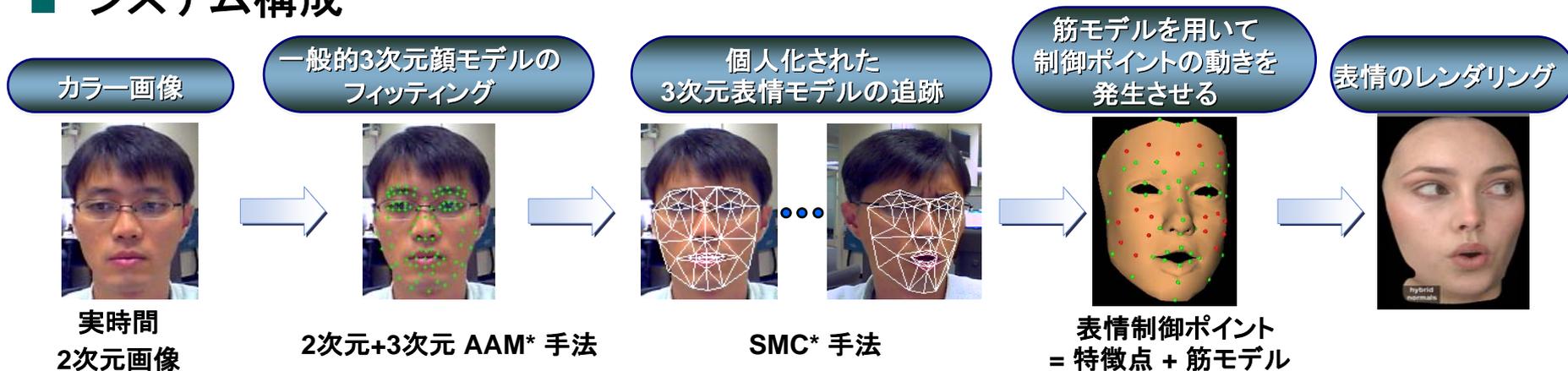
モーションキャプチャ DB



剛性に基づく筋モデル

$$\sum_m k_m \left[L_{mn}^l - \|L_{mn}^0\| \frac{L_{mn}^l}{\|L_{mn}^l\|} \right] = 0$$

■ システム構成



■ Performance

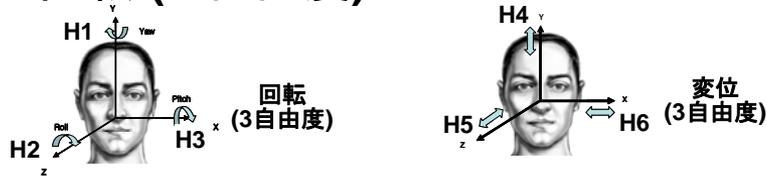
	モーションキャプチャベースの手法	カメラ1台のビジョンベースの手法	提案手法	
カメラの台数	≥ 7	1	1	} 利便性
顔に貼り付けるマーカーの数	≥ 70	0	0	
処理速度	オフライン	15fps	38.3fps	} 実時間
追跡可能な表情	すべての表情	対称な表情	すべての表情	} 高精度
頭部の回転角度	-90°~ 90° (3次元のポーズ)	-15°~ 15° (2次元のポーズ)	-90°~ 90° (3次元のポーズ)	
制御ポイントの数	≥ 60	22	75	

*AAM: Active Appearance Models, SMC : Sequential Monte Carlo Algorithm、逐次モンテカルロ法

■ 様々な表情変化 (12 動作単位)



■ 頭部の回転 (6 自由度)



■ 認識率: 90.5% (10 人、504 表情) @回転 $\leq \pm 90^\circ$; 処理時間 26.1ms @1024x768 の画像

