「画像の認識・理解シンポジウム (MIRU'96)」1996 年 7 月

複数の顔画像の組合せによる任意方向・任意表情の顔画像の生成

向川 康博† 中村 裕一‡ 大田友一‡
†筑波大学 工学研究科
‡筑波大学 電子・情報工学系
URL `http://image-gw.esys.tsukuba.ac.jp/`

Abstract: In this paper, we propose a method to synthesize arbitrarily oriented face views with arbitrary expressions combining multiple facial images. In this method, any views of a 3-D face can be generated without explicit reconstruction of the 3-D structure. Facial images are synthesized by mapping the weighted texture of the multiple input images to the 2-D triangular patches whose vertices are the feature points. By using the input images which has actual facial expressions, natural facial images can be synthesized.

1 はじめに

人間の顔画像は，特定個人の特徴だけでなく，表情表現など多くの情報を含んでいるため，高度なマンマシンインターフェースの実現に向けて顔画像を生成する研究が行われている．任意方向から見た任意表情の顔画像を生成する技術は，例えばＴＶ会議システム，服飾シミュレーション，仮想俳優など，様々なアプリケーションに用いることができる．

しかし，顔は人間にとって最も親しみのある対象の一つであるため，ＣＧによって合成された顔画像では，大きな不自然さを感じてしまう場合がある．そのため，実際の人物の顔画像をもとにして，より自然な顔画像を生成する方法について研究が進められている．

このためには，顔の様々な見え方の変化に対処できなければならないが，それは

(1) 方向による見え方の変化
(2) 表情による見え方の変化

に大別できる．これらの見え方の変化を再現するためには，頭部の３次元形状モデルを利用し，様々な表情変化に合わせてこのモデルを変形させる手法が一般的である [1]．

３次元形状を計測する方法としては，レンジファインダ等を利用して計測する方法 [2] や，複数の画像を用いて structure-from-motion の原理 [3] により復元する方法 [4] などが知られている．前者は正確な形状が安定して得られる反面，レーザー光等を利用した特殊な計測装置が必要となる．それに対し，後者はカメラによって入力された画像のみを入力とするが，安定した形状の復元には問題が多い．

また，様々な表情を生成するには，表情を FACS (Facial Action Coding System)[5] によって記述し，それぞれの AU (Action Unit) ごとにあらかじめ用意された変形を加えて画像を生成する方法が知られている [6]．しかし，無表情の顔を変形させて，笑った時に生じる顔のしわといった微妙な変化を再現するのは難しい．

それに対し，我々は，入力画像上の２次元座標の線形結合により任意方向の見え方の２次元座標が表現できるとする Ullman 等が示した原理 [7] を画像生成に応用し，３次元形状を陽に復元することなく，２枚の入力画像の組み合わせによって直接に任意方向の見え方画像を生成できることを示した [8]．つまりシステムの最終目的が顔画像の生成の場合には，頭部の３次元構造の復元や表情のモデル化は必ずしも必要ではなく，複数の顔画像を組み合わせることで，別な見え方を画像として生成することができれば十分であるとの考え方である．

また，この手法の持つ，２次元の画像から得られる情報をそのままの形で活用できるという利点は，方向による見え方の変化だけでなく，表情による見え方の変化にも適用が可能であり，顔の方向の変化と表情の変化を同じ枠組みで扱うことができる．本稿では，複数の画像を組み合わせることで，任意方向・任意表情の見え方を直接的に生成する手法について提案する．

Synthesis of Arbitrarily Oriented Face Views with Arbitrary Expressions from Multiple Facial Images
Yasuhiro Mukaigawa†Yuichi Nakamura‡Yuichi Ohta‡
†University of Tsukuba, Doctoral Program in Engineering
‡University of Tsukuba, Institute of Information Sciences and Electronics

2　特徴点座標の算出

2.1　顔の向きによる見え方の変化

撮影方向が未知の異なる方向から撮影された, 同一表情の2枚の顔画像 B_1, B_2 が与えられ, 顔表面上に配置された n 個の特徴点の画像間での対応付けがとれている場合を考える．画像 B_1, B_2 上での i 番目の特徴点の2次元座標をそれぞれ $(x_i^1, y_i^1), (x_i^2, y_i^2)$ とする．また, これらの特徴点は, 3次元空間中の点 (X_i, Y_i, Z_i) を回転させて2次元に射影したものであるとする．これらの特徴点の座標を一定の順で並べたベクトルと行列を以下のように定義する．

$$
\begin{aligned}
\boldsymbol{x}^1 &= [x_1^1, x_2^1, \cdots, x_n^1] && (1)\\
\boldsymbol{y}^1 &= [y_1^1, y_2^1, \cdots, y_n^1] && (2)\\
\boldsymbol{x}^2 &= [x_1^2, x_2^2, \cdots, x_n^2] && (3)\\
\boldsymbol{y}^2 &= [y_1^2, y_2^2, \cdots, y_n^2] && (4)
\end{aligned}
$$

$$
\boldsymbol{P} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \\ Y_1 & Y_2 & \cdots & Y_n \\ Z_1 & Z_2 & \cdots & Z_n \end{bmatrix} \tag{5}
$$

ここで, 話を簡略化するために平行移動成分は取り除いてあるものとし, 正射影モデルを仮定する．各画像上での特徴点座標ベクトルは次式のように, 特徴点を3次元空間中で回転させて2次元に射影する 2×3 の変換行列と, 特徴点の3次元座標との積で表される．

$$
\begin{bmatrix} \boldsymbol{x}^1 \\ \boldsymbol{y}^1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_x^1 \\ \boldsymbol{r}_y^1 \end{bmatrix} \boldsymbol{P} \tag{6}
$$

$$
\begin{bmatrix} \boldsymbol{x}^2 \\ \boldsymbol{y}^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_x^2 \\ \boldsymbol{r}_y^2 \end{bmatrix} \boldsymbol{P} \tag{7}
$$

ここで, 入力画像とは別の方向からの見え方 $\hat{B}$ における特徴点の2次元座標を一定の順に並べたベクトルを $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$ とする．また, $\hat{B}$ に対応する座標変換行列の第1行, 第2行ベクトルも同様に $\hat{\boldsymbol{r}}_x, \hat{\boldsymbol{r}}_y$ とする．

ベクトル $\hat{\boldsymbol{r}}_x$ は3次のベクトルであるので, $\boldsymbol{r}_x^1$, $\boldsymbol{r}_x^2$, $\boldsymbol{r}_y^1$ が互いに1次独立であれば,

$$
\hat{\boldsymbol{r}}_x = c_{x1}\boldsymbol{r}_x^1 + c_{x2}\boldsymbol{r}_x^2 + c_{x3}\boldsymbol{r}_y^1 \tag{8}
$$

を満たす係数 $c_{x1\sim3}$ が一意に決まる．このことは, $\hat{B}$ における全てのX座標が, 次式のように B^1, B^2 における2次元座標の線形結合で表現可能であることを意味する．Y座標についても同様である．

$$
\hat{\boldsymbol{x}} = c_{x1}\boldsymbol{x}^1 + c_{x2}\boldsymbol{x}^2 + c_{x3}\boldsymbol{y}^1 \tag{9}
$$

ただし, 2枚の入力画像 B^1, B^2 の撮影方向によっては, 線形結合の基底ベクトルが1次従属になる場合がある．そこで, 4つのベクトル $\boldsymbol{x}^1$, $\boldsymbol{y}^1$, $\boldsymbol{x}^2$, $\boldsymbol{y}^2$ の主成分分析を行い, 第1 ~ 3主成分 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_3$ を線形結合の基底ベクトルに利用することで, 式(10)(11)のようにX, Y座標は安定した正規直交ベクトルの線形結合で表現できる．また, 入力画像が3枚以上利用できる場合には, 同様にそれぞれの入力画像での x, y 座標ベクトルを全て主成分分析することにより, さらに安定した基底ベクトルが得られる．

$$
\begin{aligned}
\hat{\boldsymbol{x}} &= c_{x1}\boldsymbol{p}_1 + c_{x2}\boldsymbol{p}_2 + c_{x3}\boldsymbol{p}_3 && (10)\\
\hat{\boldsymbol{y}} &= c_{y1}\boldsymbol{p}_1 + c_{y2}\boldsymbol{p}_2 + c_{y3}\boldsymbol{p}_3 && (11)
\end{aligned}
$$

これらの線形結合の6つの係数を適切に決めることで, 任意方向の顔画像上の特徴点の2次元座標が求められる．

2.2　表情による見え方の変化

同一方向で表情の異なる m 枚の顔画像 $B_j(1 \leq j \leq m)$ が与えられ, これらをもとに様々な表情を持った顔画像を生成することを考える．もし, これらの入力画像が様々な表情パターンを十分に含んでいれば, あらゆる表情における特徴点座標は, 入力画像上の特徴点座標の内挿・外挿で近似できると考えられる．

前節と同様に, m 枚の入力画像上での特徴点のX, Y座標を一定の順に並べたベクトルをそれぞれ $\boldsymbol{x}^j$, $\boldsymbol{y}^j$ とする．任意の表情を持った見え方 $\hat{B}$ における特徴点座標を表すベクトル $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$ は次式のように線形結合で近似できるとすれば, これらの線形結合の係数 $c_1 \sim c_m$ を適切に決めることで, 様々な表情の顔画像上の特徴点の2次元座標が求められる．

$$
\begin{aligned}
\hat{\boldsymbol{x}} &= \sum_{i=1}^{m} c_i \boldsymbol{x}_i && (12)\\
\hat{\boldsymbol{y}} &= \sum_{i=1}^{m} c_i \boldsymbol{y}_i && (13)
\end{aligned}
$$

このように, 生成する画像の表情を, 入力画像の表情の内挿で指定するという方法は, 表情に関するヒューリスティックな知識を用いないという点で, また言語やFACSによる表情の指定よりも視覚的にわかり易いという点で, 有効であると期待できる．

この方法では, 生成できる表情の多様さは, 入力として用いる画像にどれだけ多くの基底となりうる表情が含まれているかに依存する．このことは逆に言えば, より多くの表情の種類を持った入力画像を用いれば, 生成できる画像の種類や自然さが向上するということである．

3　テクスチャの生成

顔表面上の特徴点の2次元座標を算出する方法については前節で述べた．次に, 顔画像を生成するために, 特徴点の間を埋める面をテクスチャマッピングにより生成する手法について述べる．

3.1　テクスチャの重み付き平均

マッピングするテクスチャとして，仮に正面向きの顔画像のみをテクスチャとして用いた場合，生成する顔の向きが正面向きから離れるにしたがって，テクスチャの歪みが大きくなる．この原因は，顔表面を多面体で近似したことによるずれや，頭部側面などの正面向きではほとんど見えていない領域を無理に引き延ばしたことなどが考えられる．しかし，生成する画像と顔の方向が似ている見え方を入力画像から選択してテクスチャとして用いれば，テクスチャの変形が少ない分だけ，これらの歪みの問題が軽減でき，より自然な顔画像を生成することができる．

また，表情変化を持った顔画像を生成する場合にも同様のことが言える．様々な表情を持った顔画像のテクスチャを生成するには，無表情の顔を無理に変形するよりも，もともと表情を持った画像を無理な変形なしに用いる方が，より自然な表情のテクスチャが得られる．特に，表情によって生じる顔のしわなどは，単に無表情のテクスチャを伸縮するだけでは再現できない．

そこで，生成するテクスチャは複数の入力画像から得られるテクスチャの重み付け平均を用いる．生成する画像と似た方向・表情の入力画像に対する重みは大きく設定する．これにより，入力画像の持つテクスチャ情報を最大限に活用できる．

この方法では，入力画像と同じ方向・表情の顔画像を生成する時には，入力画像と全く同じ画像が得られるという利点がある．また，表情の内挿においても，眼や口などの形状だけでなく，テクスチャ自体も，その中間の見え方が生成され，粗いパッチモデルを利用した場合でも視覚的に満足な画像が得られる．

3.2　重みの設定

顔の向きによる重み

入力画像は，特徴点の2次元座標がわかっているだけで，撮影方向は未知であるが，因子分解法[4]によって，大まかな撮影方向を推定することは可能である．まず，入力画像から鼻の頂点や眼の両端など表情に依存しない数点の特徴点を選択する．これらの特徴点を用いて因子分解法を行い，各入力画像に対する撮影方向，すなわち顔の向きを推定する*．また，生成する画像でも同様に顔の向きを推定する．各入力画像に対する重み係数は，生成する画像との顔の向きの角度差の2乗に反比例するように設定する．

表情による重み

表情に関する重み係数については，式(12) (13)の係数 $c_1 \sim c_m$ と同じ比率とし，総和が1になるように正規化する．特徴点座標と同時にテクスチャも内挿するという点では，原理的にモーフィングと同じである．

3.3　2次元のテクスチャマッピング

テクスチャマッピングは，特徴点を頂点とする3角形パッチごとに行う．テクスチャは，各パッチ毎に複数の入力画像から3角形領域を切り出し，生成する画像での特徴点座標をもとにアフィン変換を行い，重み付け平均をとった後，マッピングされる．ただし，特徴点は2次元であるため，パッチモデルも2次元となり，パッチ間の前後関係はわからない．そこで，各3角形パッチの頂点の記述順から表裏の判定を行い，裏側のパッチにはテクスチャを描画しない．人間の頭部のような凸多面体に近い形状では，この程度の表裏判定だけで実用上十分である．

4　実験結果

4.1　顔表面上の特徴点とパッチモデル

顔画像上に86点の特徴点を配置する．特徴点は，眼や口などの顔部品の端点や，表情による移動量が多い頬などに配置した．これらの特徴点の対応付けは，現時点では顔表面にマーキングをつけ，手入力により与えている．また，図1に示すような，特徴点を頂点とする158面の2次元の3角形パッチモデルを作成した．

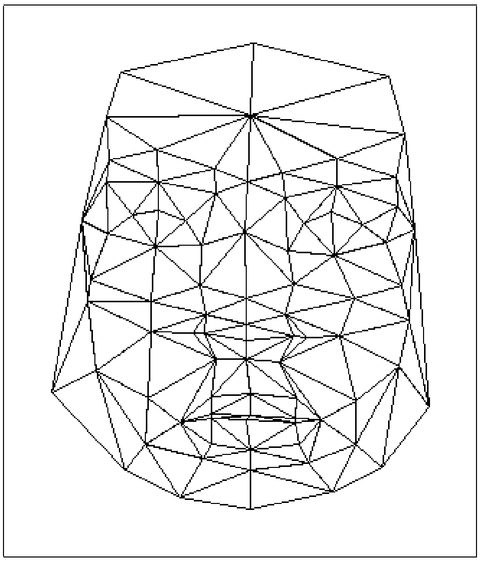

図1: 2次元の3角形パッチモデル

---

*因子分解法によって撮影方向と同時に大まかな頭部の3次元形状も推定できるが，入力画像は様々な表情を含んでおり，推定結果は不安定になる．そのため本手法では，因子分解法によって推定された頭部の3次元形状は利用せず，撮影方向のみをテクスチャの重み係数を決めるために利用している．

図 2: 2 枚の入力画像（無表情）

4.2 画像生成

まず, 図 2に示す 2 枚の顔画像を入力とし, 顔の向きを変化させた画像を生成する実験を行った．生成する顔の向きに対応する線形結合の係数を決定するための一方法として, 別に用意した眼の端・鼻頂点・耳つけ根の 5 点の 3 次元データを用いて, 正面顔から左右に $-30°$, $0°$, $30°$, 上下に $-20°$, $0°$, $20°$ 回転させた場合の 5 点の 2 次元座標を算出し, これらの 5 点が画像上で一致するように線形結合の係数を最小自乗法により求めた．生成された顔画像を図 3に示す．

画像の単なる 2 次元の中割り [9] では, このような 3 次元的な回転を含む見え方の変化を正確に表すことはできないが, 本手法では正しく画像が生成されていることがわかる．

次に, 図 4に示す, 無表情と「怒り」表情の 2 枚の画像を用いて, 表情の内挿画像を生成する実験を行った． 2 枚の入力画像に対応する線形結合の係数を, 等間隔に線形に内挿し, 内挿画像の特徴点座標を算出した．

テクスチャの重み付け平均の有効性を確認するため, 図 4に示す内挿画像では上から順に, 無表情の画像のみを用いる場合,「怒り」の画像のみを用いる場合, 双方のテクスチャの重み付け平均を用いる場合の結果を示す．一方の表情のテクスチャを用いた場合, 変形の度合が大きくなるにつれて, 不自然さが増加する．それに対し, テクスチャの重み付け平均を用いた場合には, 両端の見え方と内挿の見え方の双方とも, 自然な顔画像が得られている．

さらに, 図 5に示す, 様々な方向と表情を持った入力画像を用いて顔画像を生成する実験を行った． 図 6(a) は, 生成したい画像の顔の向きを示すサンプル画像であり, この画像と同じ向きになるように新しい顔画像を生成する．サンプル画像は, 入力画像の人物とは別人である．サンプル画像から, 個人性や表情変化に影響を受けにくい, 眼の端・鼻頂点・耳つけ根・顎の先の 6 点の 2 次元座標を抽出し, これらの座標をもとに線形結合の係数を決定した．なお, 片方の耳の付け根が隠れている場合は, 残りの 5 点だけで係数を求めている．

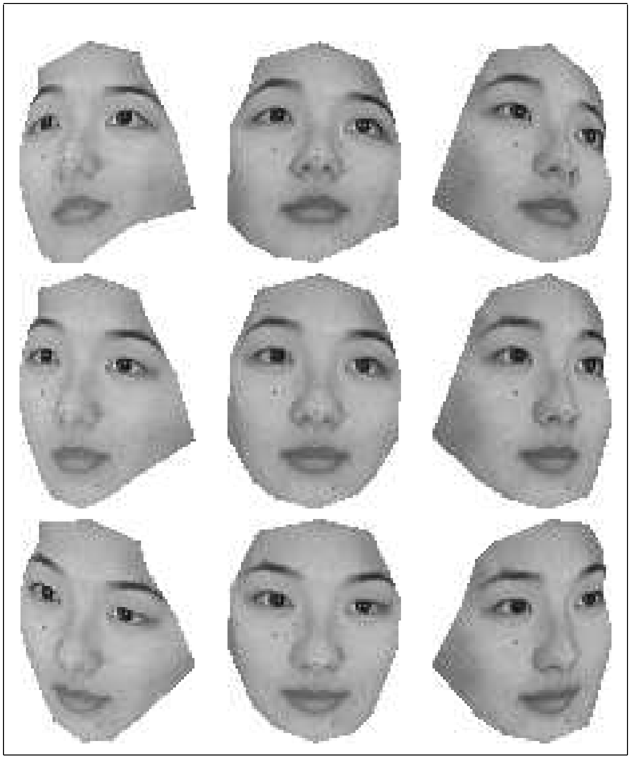

図 3: 様々な方向の顔画像の合成例

図 6(b)∼(e) は, (無表情, 笑顔, 怒り顔) に対する係数をそれぞれ順に $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(0.5, 0.5, 0)$ に設定して, 生成された表情である．特に (e) では, 無表情と笑顔の中間の表情という, 新たな表情の見え方が生成できていることがわかる．

この実験例では, 入力画像の各表情の顔の方向が同じであるが, それぞれ任意方向で撮影方向も未知で構わない．また, 入力画像に含まれていない顔の向きや表情の顔画像を新たに生成できることが確認できた．

5 まとめ

頭部の 3 次元構造の復元や, 表情の精密なモデル化を行うことなく, 任意方向・任意表情の顔画像を生成する枠組みについて提案した．入力画像上での特徴点座標の線形結合と, テクスチャの重み付け平均によって, 顔画像を生成することにより, 方向による見え方の変化と, 表情による見え方の変化を, 同じ枠組みで統一的に扱うことが可能であることが確かめられた．また, 2 次元の入力画像から得られる情報をそのままの形で活用することで, 自然な表情を持った顔画像を生成できることを, 実験により確認した．

現時点での実験では, 生成できる表情の種類に制限があるが, 入力画像に「幸福」,「悲しみ」,「怒り」,「恐怖」,「嫌悪」,「驚き」の基本 6 表情を用いることで, 任意の表情変化を持った顔画像を生成することが今

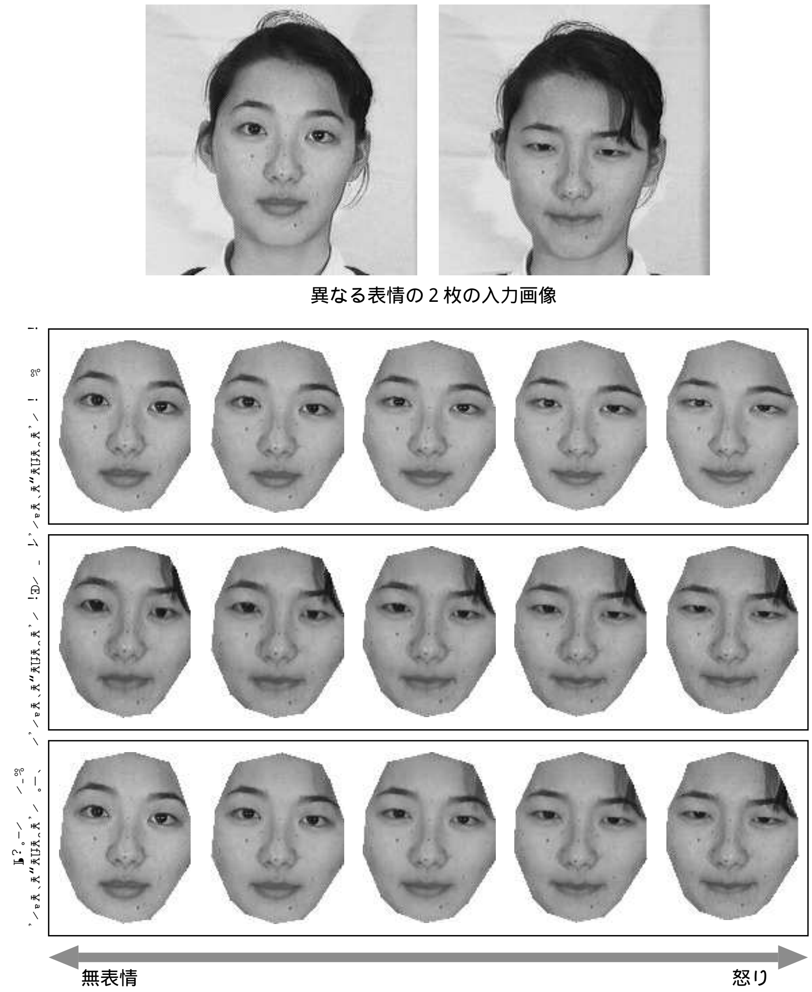

図 4: 表情の内挿

後の課題である.

参考文献

[1] 小松 功児, “顔画像からの顔の３次元形状の推定とその表情の変更”, 信学論, Vol.J73-D-II, No.5, pp. 707–716, (1990)

[2] 末永 康仁, 渡部 保日児, “３Ｄ形状と輝度（色）の同時計測が可能なスキャナとその顔画像入力への応用”, 情処研報, CV 67-5 (1990)

[3] S.Ullman: The Interpretations of Visual Motion, MIT Press, Cambridge, MA (1979)

[4] Carlo Tomasi, Takeo Kanade, “The Factorization Method for the Recovery of Shape and Motion from Image Streams”, *In Proceedings of Image Understanding Workshop*, pp 459–472 (1992)

[5] P.Ekman, W.W.Friesen (工藤 訳), 『表情分析入門』, 誠信書房 (1987)

[6] 崔 昌石, 原島 博, 武部 幹, “分析合成符号化における表情の記述と合成について”, 信学技報, IE 88-84 (1988)

[7] S.Ullman, R.Basri: Recognition by Linear Combinations of Models, *IEEE Trans. PAMI*, Vol.13, No.10, pp. 992–1006 (1991)

[8] 向川康博, 中村裕一, 大田友一, “２枚の顔写真を用いた任意方向の顔画像の生成”, 情処論, Vol.37, No.4, pp. 635–644 (1996)

[9] R.Skerjanc, and J.Liu: Computation of Intermediate Views for 3DTV, *Proc. 5th Workshop 1992 on Theoretical Foundations of Computer Vision*, pp. 190–201 (1992)

(a) 無表情　(b) 笑顔　(c) 怒り顔

図 5: ３種類の表情を含む入力画像

(a) サンプル画像　(b) 無表情　(c) 笑顔　(d) 怒り顔　(e) 無表情と笑顔の内挿

図 6: 様々な方向と表情の顔画像の合成結果