

# [3ZA-07] 顔画像生成のための人間の空間周波数特性の利用

吉岡 史樹\* 尾田 政臣† 磯野 勝宣‡ 赤松 茂‡

\*立命館大学文学部 Institute †立命館大学理工学部/ATR 人間情報通信研究所 ‡ATR 人間情報通信研究所

## 1 まえがき

コンピューターを利用した画像検索システムの研究が、動作速度や検索効率、検索インタフェースなどの観点から行われている。なかでも顔画像の検索をめぐるのは、画像処理やプログラミング以外からでも、顔の認知や表情認識などの認知心理学の視点から検索という問題を取り扱っている。

既存の顔表情データベースから好みに近いものを取り出す「検索」[1]と、顔の形状をコンピュータ処理によって変形させて新たな顔画像を提示しながらユーザーが好みの顔画像を「生成」する方法を同じようなインタフェースで行うことができることが示されている。また遺伝的アルゴリズム(GA)を用いた表情画像生成システムの試みもなされており[2]、それによってユーザーがより多様な顔画像の生成を行うことができる。

顔画像を評価・生成というインタラクティブな過程を経てよりイメージに近い表情画像を生成へと絞り込んで行くことが出来る。今回は遺伝的アルゴリズムを用いた二次元画像に基づく表情生成システム[3]に、空間周波数処理を取り入れて、提示される顔画像の空間周波数成分が、生成の進度に応じて段階的に増加する方法を試みた。ユーザーのイメージが曖昧である生成初期の段階では、ユーザーの生成行動が顔の詳細な情報にまで及んで、生成行動が不安定になり収束速度に影響がでると考えられる。空間周波数処理によって高い周波数に含まれる情報を初期段階では提示しないことで、低い周波数に関する情報が重点的に評価され、高い周波数に含まれる情報が初期の検索行動に影響を与えなくなると、生成効率を向上させることができるものと考えられる。笑顔や幸福といった言葉で表現される顔の表情認識について、その顔画像に含まれている低周波の情報に関与しているといわれている[4]。そこで今回はこのシステムを用いて笑顔をユーザーに生成してもらった実験を行った。

## 2 実験

空間周波数処理には、FFT フィルタを用いる。生成初期段階では、ローパスフィルタリングを施し、細かいパーツや輪郭線などの高周波成分の情報がカットされている。一度に 12 枚の顔画像がディスプレイに提示されユーザーはそのすべてを評定する。その評定値をもとにシステムは新たな顔画像を生成してディスプレイに表示する。この過程を繰り返して自分の満足の行く画像が表示されたら生成を終了とする。

実験で提示される顔画像は 256pixel 四方で 4 枚×3 段で表示される。システムでは顔画像を形状とテクスチャに分離して処理しており、変化させた形状にテクスチャを張りつけることで異なる表情の顔画像を表現している。顔画像サンプルに対して PCA を行うことにより固有顔が求まり、各固有顔を線形結合することによって任意の顔画像が得られる。このとき線形結合時の重み係数を変化させることにより、様々な顔画像を生成することができる。今回のシステムではテクスチャは一定で、重み係数を制御することで形状変化させる。実験は二つありフィルタの有無による収束の状況とフィルタ有りの条件でも顔の生成が利用者の意図どおりに進むことをそれぞれ確認する。

第一実験では、FFT フィルタを用いる課題と FFT フィルタを用いない課題を順に行った。FFT フィルタを用いる課題では、ユーザーが表情の判断が難しくなったら、手動でカット周波数を上昇させて、画像をある程度クリアにして評価できるようにした。自動的なカットオフ周波数の上昇は行っていない。カットオフ周波数は一定の値になるまで何回でも上昇させることができる。FFT フィルタを用いない課題では、実験の最初から最後まで同様のクリアな画像を提示した。被験者は大学生 7 人で行った。また実験終了後に最終出力画像についての評価について簡単なアンケートを行った。

第二実験では、第一実験と同様の方法で、33Hz/顔幅以上になったら FFT フィルタをはずして、それ以降はクリアな画像を評価させながら生成を行った。即ちフィルタが外れるまでは徐々にぼやけ方が小さくなる。最終出力画像が決定できた段階で終了とした。被験者は第一実験とは異なる大学生 11 人で行った。

それぞれの実験で、生成システムが顔画像の形状を計算するために用いている重み係数を顔の特徴とした。

各選択機会ごとの提示されている 12 枚の画像についての重み係数の分散値を収束の指標として記録した。

### 3 実験結果と考察

第一実験では、出来あがった画像の満足度は、7 段階評価の被験者全員の平均値で、フィルタありの場合が 2.4、なしの場合が 5.6 となり、フィルタなしで生成を行ったほうが高くなっている。終了直前の分散の平均は、フィルタありで 9.87、なしで 8.79 であった。

第二実験ではフィルタが解除になる段階で分散値の平均は 8.31 となった(図 1)。フィルタ解除後に分散値の合計が上昇するような場合があるが、数値としては大きな値ではなく、主に特定のパーツの再生成によってパーツごとの調整を行っているためであることが、分析の結果わかった。主に目、眉の動きが大きい。フィルタ解除以降の収束もフィルタありで生成を行った時と収束の仕方が大きく変化することはなく、終了時の分散の平均は 6.70 となった。フィルタありからなしにきりかえても、分散の大きな変動はない。

第一実験の結果から、最終出力画像の決定に必要な収束は、分散値で見るとフィルタありとなしの条件で大きな差が現れないことがわかった。また画像の最終的な判断のためには、詳細な情報が必要であることがいえる。また第二実験でのフィルタが解除になる段階と、第一実験でフィルタなしで生成を行った場合とを比べても、同等な値まで画像を収束させることが可能であるといえる。これらのことから、生成初期段階に FFT フィルタを用いて生成を行っても、大まかな表情の生成は可能であるといえる。生成回数から見ると、第一実験でのフィルタなしで 37.6 回、ありで 24.2 回となりフィルタありのほうが、終了までの生成回数は早くなっている(図 2)。第二実験では回数は多くなったが、フィルタ解除時までの回数は少なくなり、終了時の分散値も小さくなったため、回数の増加は表情の改良に当てられたとみることができる。

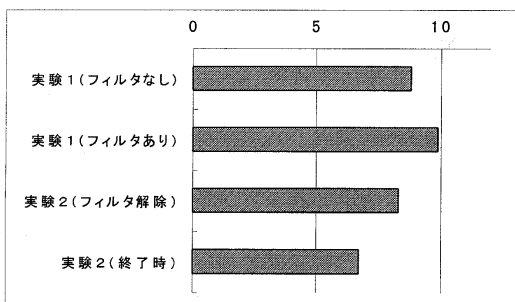


図 1. 分散値の比較(平均値)

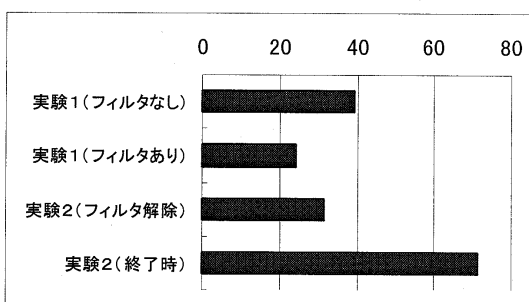


図 2. 生成回数の比較(平均値, 単位 回)

### 4 結論

低周波では、ユーザーは全体の大まかな評価を行うが、それでもユーザーの望む顔画像にある程度まで収束させることができる。つまり高周波に含まれる情報に左右されることがなく、収束速度を速めることができる。フィルタありでは生成回数が少なく、高周波成分ありになってからの調節を一連のプロセスに組み込むことで全体の生成効率を向上させることがわかった。

#### 参考文献

- [1] 尾田政臣, “人間のイメージ形成過程の特性を利用した画像検索システム”, 情処学論, vol135, no7, pp. 1449~1456, July 1994
  - [2] 磯野勝宣, 尾田政臣, 赤松茂, “遺伝的アルゴリズムを用いた顔画像生成システム——感性的イメージの生成——”, 電子情報通信学会論文誌 D-II Vol. J82-D-II No. 3 pp. 483-493 1999年3月
  - [3] 山本光穂, 磯野勝宣, 赤松茂, “遺伝的アルゴリズムを用いた顔表情生成システム” (信学技報, HIP98-52, 1999-3)
  - [4] 桐田隆博, 遠藤光男, 赤松茂, 深町英夫 “表情認識と空間周波数情報”, (ATR HIP Res. Labs., 1994 Vol2 of 2 P570)
- なお本研究のデータの一部は、科学技術庁の平成 11 年度科学技術振興調整費による「視覚におけるニューロインフォマティクスに関する研究」の一環として得られたものである。

#### Application of human property of spatial frequency for synthesizing facial expressions

Yoshioka FUMIKI \*, Oda MASAOMI +, Isono KATSUNORI #, Akamatsu SHIGERU #

\* Ritsumeikan Univ. college of letters Institute

+ Ritsumeikan Univ. college of science and engineering / ATR Human information processing laboratory

# ATR Human information processing laboratory