

目領域内の濃度値変動に基づく瞬き検出とその応用

大森匡彦† 村上和人‡

† 愛知県立大学大学院情報科学研究科 ‡ 愛知県立大学情報科学部

〒480-1198 愛知県愛知郡長久手町熊張

E-mail: im021008@cis.aichi-pu.ac.jp, murakami@ist.aichi-pu.ac.jp

顔画像を用いた個人認証や個人性特徴抽出では、正面顔で無表情な顔画像が使われることが多く、瞬きをしていない顔画像が必要である。我々は、瞬きと目領域内の平均濃度値の時間変動との間に相関関係があることを実験的に確認してきた。本論文では、2値化手法と楕円当てはめによって顔領域を検出し、射影処理とエッジ検出手法を用いて目領域内のエッジ点数を計測する方法について説明する。また、これを一連の動画画像に適用し、それらの中から、瞬きをしていない顔画像の自動抽出法について述べる。

A Simple Method to Extract the Best Shot from Image Sequences Based on the Variance of Edges

Masahiko Omori†, Kazuhito Murakami‡

† Graduate School of Information Science and Technology, Aichi Prefectural University,

‡ Faculty of Information Science and Technology, Aichi Prefectural University
Kumabari, Nagakute-cho, Aichi, 480-1198 Japan

E-mail: im021008@cis.aichi-pu.ac.jp, murakami@ist.aichi-pu.ac.jp

In the fields of face recognition or identification system, facial images with normal expression are used as input face data. And, also in the research areas of personal feature extraction, winking (close-eye) images have not been used as the experiment data. In this paper, we propose a simple but robust method composed of edge extraction, binarization and ellipse fitting in order to select the best shot or reject inconvenient face images from image sequences. In this paper, we show experimentally the effectivity and robustness of the proposed method.

1. はじめに

顔画像処理はセキュリティやコンピュータを用いたヒューマンインタフェースを始めとして、広範な分野で用いられ始めている。顔画像を用いた個人認証では、正面顔で無表情な顔画像が使われることが多い[1]。また、個人性特徴抽出や表情抽出などの場

合も、瞬きをしていない顔画像が実験データ、あるいは教師データとして使われている[2]。顔の向く方向や顔の傾きに関しては、ある程度ロバスト性を高めることも可能であるが、瞬きや表情などのように認識対象（一般の利用者や、特定のユーザ、など）に対して強要することが困難な要素もある。顔画像

処理の実利用を想定すると、ある瞬間の1枚の静止画像を用いて認識するよりも、一連の動画像中から目的に合致した顔画像を用いて認識した方が好ましいことは容易に推察できる。

これらの観点から筆者らは、動画像中から目的に合致した顔画像を抽出する(いわゆるベストショット抽出)あるいは、目的に合致しない顔画像を排除するシステムの構築に取り組んでいる。予備実験により瞬きと目領域内の平均濃度値の時間変動との間には相関があることを確認した。本稿ではこの特性を利用して、一連の動画像から瞬きをしていない画像データを抽出する顔画像処理手法と基礎実験結果を中心に述べる[3]、[4]。

以下、2.にて目領域の濃度値に基づいた瞬き検出法の基本的な考え方について説明した後、3.にてエッジ検出法、2値化手法、楕円当てはめを用いた画像中からの顔領域と目領域の抽出法を、4.にて瞬き検出法を提案する。そして、5.にてエッジ点の時間変動を用いたベストショット抽出方法を、実験結果とともに示し、提案手法の有効性について考察する。

2. 目領域の濃度値に基づいた瞬き検出法

本論文で提案する顔画像抽出法は、一連の動画像からの「よい顔」の抽出を目的としている。「よい顔」とは、例えば集合写真の場合は、全員が目を開いている画像の抽出であり(たぶん、そうであると思われる)記念撮影などでは全員が微笑んでいる画像の抽出である。これらの抽出対象は目的によって変化するが、基本的には数秒間動画像を撮影し、それらの中からベストショットフレーム画像の抽出を行う、つまり、最もよい画像を選択することになる。その基本手順は、

- Step1. 顔画像の入力をする。
- Step2. 顔領域の抽出をする
- Step3. 顔部品(目、口)の抽出を行う。
- Step4. 目の形状変化による瞬き検出を行う。
- Step5. 口の形状変化を検出をする。
- Step6. ベストショット抽出を行う。

となる。

ところで、1枚の静止画像から目が開いている、あるいは、目が閉じている状態を判断する方法としては、上下の瞼の間の距離を測定する方法[5]、Hough変換などによる瞳検出法を応用し、瞳が検出

できなければ目を瞑っていると判断する方法[6]、などがある。これらは静止画像から瞬き検出を試みている例である。本論文で提案する手法は、目を含んだ部分領域に対して、エッジ処理を行い、その変動をもとに瞬き検出を行う方法である。つまり図1に示すように瞳に対して、エッジ検出処理を行うと上瞼と下瞼、瞳と白目部分とのエッジが検出される。瞬き時には上下2本の瞼のエッジは一つになり、目領域内のエッジ点は約半分になると期待できる。このアイデアを利用して瞬き検出を行うものである。

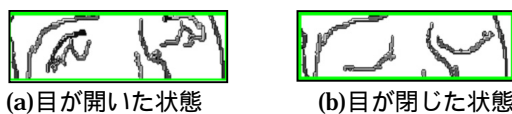


図1 目の近傍領域のエッジ画像

3. 顔領域抽出法と目領域抽出法

3.1 顔領域抽出法

入力画像からの顔領域抽出処理は画像中の顔領域の大きさに依存するため、次のように分類して行う。

(1)被写体が一人の場合:

顔領域抽出手順を次に示す。

<顔領域抽出アルゴリズム(1)>

- Step0 濃淡画像を入力する。
- Step1 2値化閾値Tの決定。
- Step2 2値化処理により部分領域の抽出。
- Step3 部分領域から半径r以上の円が含まれている領域を見つけ、それらの円形領域の論理和をマスク画像とする。
- Step4 原画像にマスク画像を当てはめ、顔領域と背景を分割し、顔領域だけを残す。

まず、Step0では濃淡画像を入力する。この際、顔領域全てが含まれていることが条件である。そして、Step1で原画像を閾値Tで2値化し、顔領域を含んだ2値連結図形を得る。次に、Step3でこの2値パターンに対して、半径rの円を内包する領域の論理和をとってマスク画像(パターン)を作成する。この処理により細かい出っ張りなどを取り除くことができる。Step4でこのマスク画像(パターン)を原画像に当てはめ、顔抽出画像とする。背景領域と顔領域を分割し、顔領域だけを残す。

(2)被写体が多数の場合:

次に、写真に複数人が写っているケースを考える。まず、カラー画像を入力し、色情報を用いて肌領域

の抽出を行う。そして、肌領域ごとに少し大きめに矩形で切り出し、顔領域の面積に応じて、次の(a),(b)のいずれかの処置を行う。(a)面積が小さい場合は、矩形内の顔領域に対して楕円当てはめを行う。(b)面積が大きい場合は、画像一人一人に対して、被写体が一人の場合と同様の操作を行う。アルゴリズムを次に示す。

<顔領域抽出アルゴリズム(2)>

- Step0 カラー画像を入力する。
- Step1 色情報を用いて肌領域の抽出を行う。
- Step2 肌領域ごとに少し大きめに矩形で切り出す。
- Step3 顔領域の面積に応じて以下の Step3-(a) あるいは Step3-(b) の処理を行う。

(顔領域が小さい場合)

- Step3-(a) 矩形内の肌領域に対して楕円当てはめを行う。

(顔領域が大きい場合)

- Step3-(b) 矩形に対して一人のときと同様の操作を行う。

3.2 目領域抽出法

画像に含まれる人数にかかわらず、目領域は顔画像の特定の部分である。目領域の抽出法を以下に示す。

まず、<顔領域抽出アルゴリズム>で得られた顔領域に対して重心Gを求める。次に、領域の横サイズを計算してWとする。また、次に縦のサイズを計算してHとする。Gを中心として縦 $H \times 0.2$ と横 $W \times 0.8$ の大きさの矩形領域を目領域とする。アルゴリ

ズムを次に示す。

<目領域抽出アルゴリズム>

- Step0 顔領域抽出アルゴリズムで抽出した顔領域の重心Gを求める。
- Step1 顔領域の横サイズを計算してWとする。
- Step2 顔領域の縦サイズを計算してHとする。
- Step3 Gを中心として縦 $H \times 20\%$ 、横 $W \times 80\%$ の矩形領域を目領域とする。

4.濃度値変動に基づく瞬き検出法

4.1 瞬き検出アルゴリズム(1)

次に、瞬き検出の方法を示す。まず、<目領域抽出アルゴリズム>で求められた目領域に対して、Deriche & Shen のオペレータを用いてエッジ検出処理を行い、エッジ画像に対して2値化の処置を行って黒と白の領域に分ける。そして、目領域内の黒画素(画素値=0)の画素数を数える。これをフレームごとに行う。瞬き検出のアルゴリズムを次のように示す。

<瞬き検出アルゴリズム(1)>

- Step0 顔領域抽出アルゴリズムで求められた目領域に対して、エッジ検出を行う。
- Step1 エッジ画像に2値化処理を行う。
- Step2 目領域内で黒画素(画素値=0)の画素数を数えて、横軸に時間軸をとってグラフを作成する。
- Step3 画素数の変化の時間的推移を見て、目のし

っかり開いているところを抽出する。

4.2 瞬き検出アルゴリズム(2)

3.1の(2)のアルゴリズムのStep3、すなわち顔領域の面積の小さな場合について考える。まず、当てはめた楕円に対して<目領域抽出アルゴリズム>を適用する。目領域内に対して閾値処理を行い、黒画素(画素値=0)の度数(面積)を計算する。これをフレームごとに繰り返す。このアルゴリズムを次に示す。

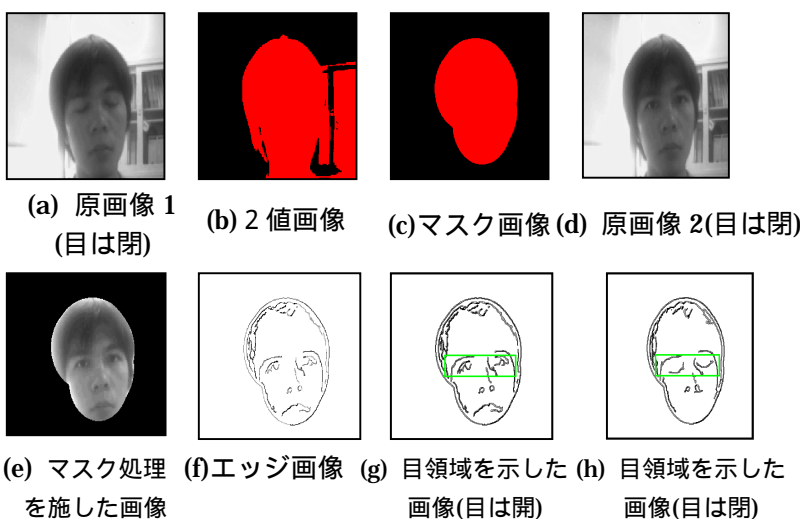


図2 エッジ点変動に基づく瞬き検出過程

<瞬き検出アルゴリズム(2)>

- Step1 楕円に<目領域抽出アルゴリズム>を適する。
- Step2 閾値 T で 2 値化を行う。
- Step3 目領域内で、黒画素の面積を計算する。
- Step4 黒画素の変化の時間的推移を見て、目のしっかり開いているところを抽出する。

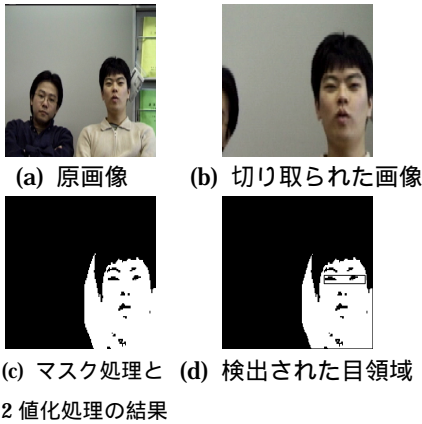


図 3 2 値化による黒画素面積変動に基づく瞬き検出過程

5. 実験方法

5.1 実験方法

まず、<顔領域抽出アルゴリズム>で求められた顔領域に対して、エッジ抽出処理を行い、次いで 2 値化処理を行う。その後、求められたエッジ画像に対して<目領域抽出アルゴリズム>によって目領域を設定する。そして<瞬き検出アルゴリズム>によって目領域内のエッジ点総数をフレームごとに数え（瞬き検出処理の過程でエッジがよく出ない画像の場合、目領域において濃淡画像に 2 値化処理を行い目の範囲の面積を調べる。）その値の変化を基にして、ベストショットフレームを検出する。実験で用いた動画のあるフレーム画像を図 2 (a), (b) に示す。フレーム間隔は 0.033 秒で、120 フレーム、約 4 秒間を撮影した。また、処理の途中経過の様子を図 2、図 3 に示す。図 2 はエッジ点変動に基づいた手法、図 3 は 2 値化による黒画素の面積変動に基づいた手法の処理過程を示したものである。実験環境は、CPU が Pentium 800Mhz の DOS/V 機を使用し、ソフトウェアは C++、一部に HALCON ソフトウェアを使用した。画像サイズは 320×240 のものを使用した。

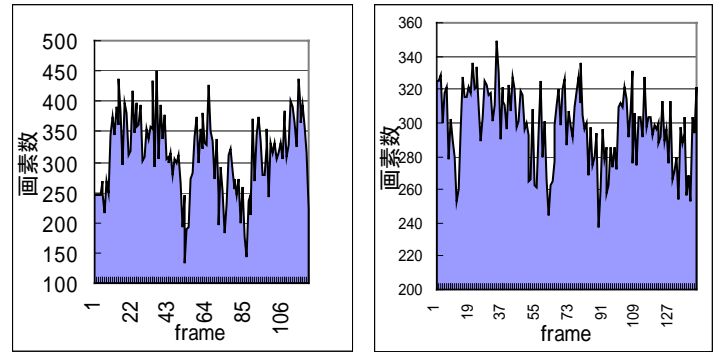


図 4 エッジ点変動の様子

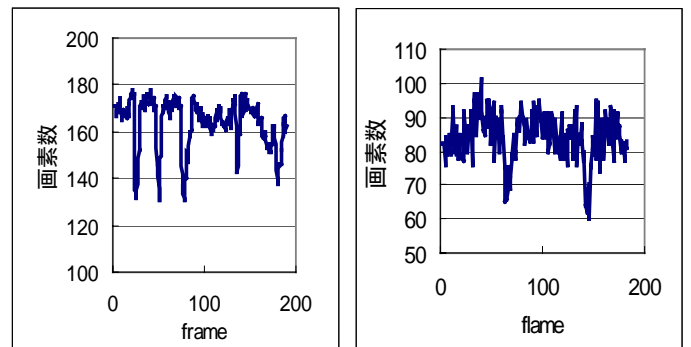


図 5 黒画素の面積変動の様子

5.2 実験結果

目領域内のエッジ点数の時間変化の様子をグラフ化した結果を図 4 に、出力例を図 6 に示す。また、2 値化による黒画素の面積変動をグラフ化した結果を図 5 に、出力例を図 7 に示す。フレームごとに目領域内のエッジ点総数が変動するのは、瞬き時には瞳や睫毛によるエッジが減少する(上下 2 本のラインが 1 本になる)ためと考えられる。検出されるエッジ点数はエッジ検出の手法によって異なる [3]。また、エッジ点数は顔領域の大きさや、画像の明るさによって出やすい場合と出にくい場合がある。顔領域の大きい画像に対しては、目領域内のエッジ点が明確に検出され、その変化の様子もつかみやすいが、顔領域の小さな画像はエッジが出にくい。このような画像に対しては瞳の部分は他の肌領域よりも黒いので、エッジ検出をするのではなく原画像を 2 値化閾値処理することにより、黒画素の面積変動として目の閉閉状態が検出できると考えられる。瞬き時には瞳による黒点、睫毛等による黒領域が減少すると考えられ、エッジ点変動と同様の変化が得られている。

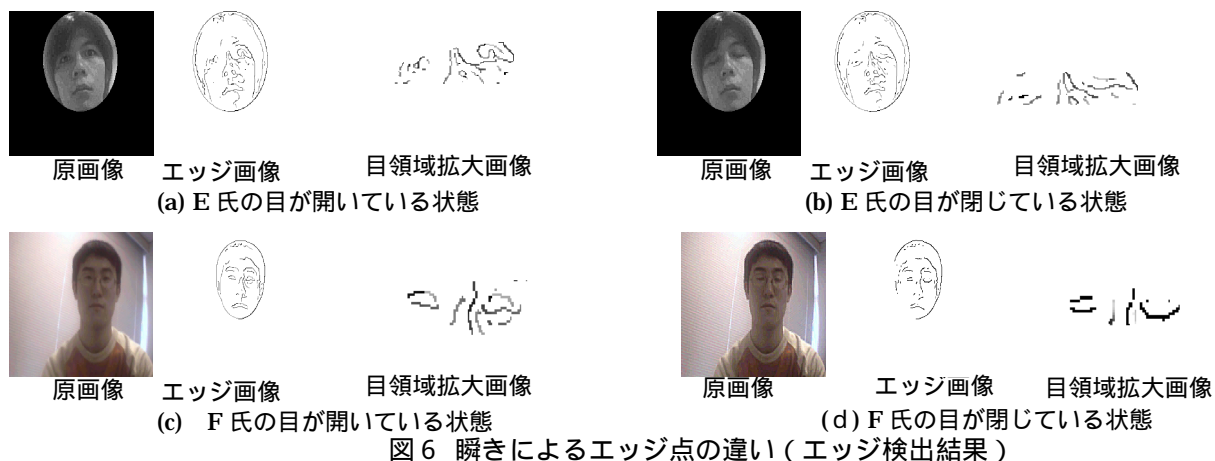


図6 瞬きによるエッジ点の違い(エッジ検出結果)

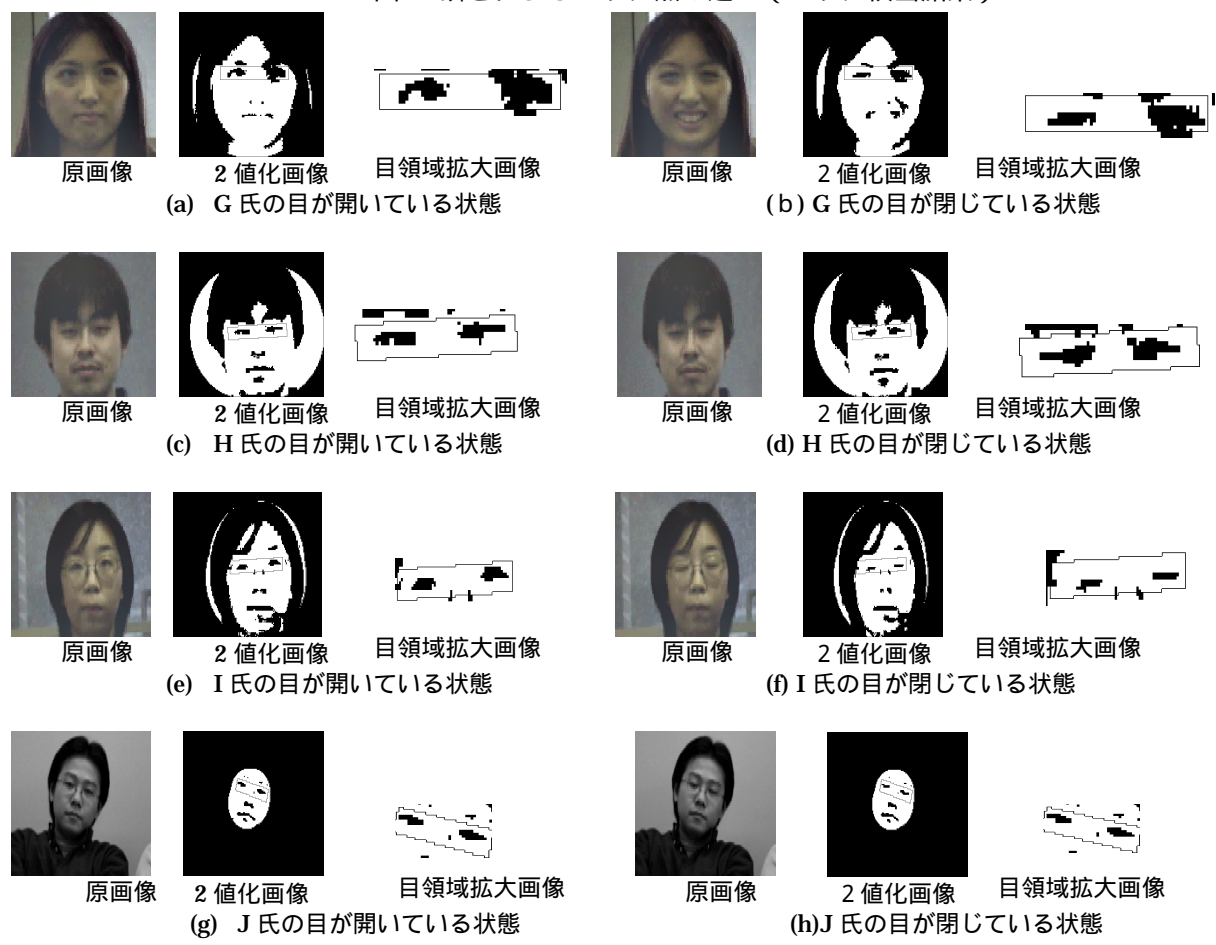


図7 瞬きによる濃度値の違い(2値化処理結果)

6. 考察

6.1 考察

図4,5の結果から分かるようにフレームごとのエッジ点変動をグラフ化したものの中にはかなりのノイズが含まれている。このままでは、局所的な谷を数多く拾ってしまう可能性があるため移動平均法に

より平滑化した。黒画素の面積変動についても同様の移動平均法で平滑化した。その結果を図8に示す。移動平均の幅を増やしていくと、グラフは滑らかになるがエッジ点変動の山や谷が見つけにくくなる。また、移動平均の幅が小さいとエッジ点変動が細かく起きるためベストショットフレームを見つけることが難しくなる。実験した結果、移動平均の幅が5

～7の時に最適な状態となることが分かった。この移動平均の幅は0.2秒～0.3秒に相当している。瞬きの周期は1～3秒であるから、平滑化によって瞬きを見逃すことはないと考えられる

移動平均幅5で平滑化したグラフを利用して平均と分散から閾値を定め、ベストショットフレームを抽出する。被写体が一人の場合は、閾値以上の値が数フレーム続く場所がベストショットフレームになりうる場所である(図9-(a))。また、被写体が複数人の場合は、一人一人のベストショットの場所を求め、全員が閾値以上となる場所を抽出する。図9-(b)は、2人が写っているケースについて模式的に表したグラフである。

の場所がベストショットフレーム候補となり、の場所は、ベストショット候補がなかった場合の第二候補である。xの場所は、どちらかが瞬きをしているフレームであり、ベストショットになりえない。原画像でに該当するフレームを確認したところ、確かに2人とも目を開いていたので本手法が有効に機能していると考えられる。

7. まとめ

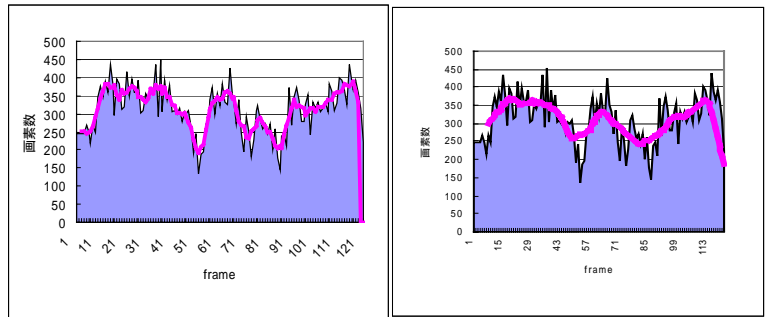
本稿では、一連の動画像から顔領域を抽出し、さらに目、口などの顔部品を抽出して、目的に合致したフレーム画像を抽出する基礎実験を行った。

エッジ点や、黒画素面積の時間変化を調べた結果、瞬きをすると目領域内の画素数が減少することがわかった。また、これにより目を閉じている画像を排除することができ、本研究の目標であるベストショットフレーム画像の抽出を行うことができた。背景画像中に、顔領域に似た濃度値を持つ領域が含まれていると、顔領域検出がうまくいかない場合が起こる。このような画像に対する顔領域の検出精度を上げることは今後、必要である。目領域内に髪や眼鏡がある場合、エッジ点変動、黒画素の面積変動ともに誤った判定をしてしまう可能性がある。目領域を矩形以外の形状でも設定できるように改善することも今後の課題である。

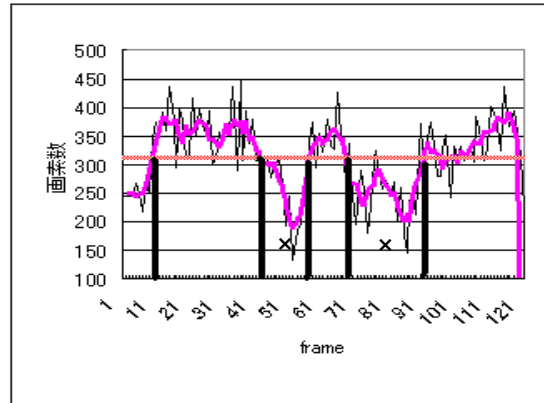
謝辞 本研究の一部は文部科学省科研費補助金、(財)立松財団研究助成金の支援による。記して感謝する。

文 献

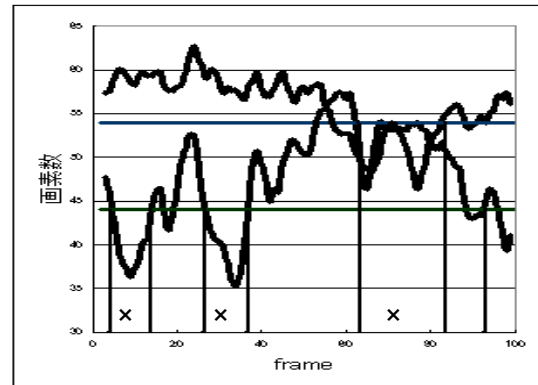
- [1] 顔論文特集 電子情報通信学会論文誌 D-II Vol.J80-D - ,No.8(1997).
- [2] 瀬川、遠藤、村上、鳥生、輿水：“似顔絵生成のための顔と瞳の輪郭抽出方式”、映像情報メディア学会誌、



(a)移動平均幅5フレームで平滑 (b)移動平均幅15フレームで平滑化
図8 移動平均法による平滑化結果



(a) 被写体が一人の場合



(b) 被写体が複数人(2人)の場合

図9 ベストショットフレーム抽出結果

Vol.51, No.11, pp.1935 - 1944(1997).

- [3] 大森匡彦、村上和人：“動画像中のベストショット抽出のための顔認識手法について”、電子情報通信学会技術研究報告 PRMU2001-117, pp.27-32(2001.11).
- [4] 大森匡彦、村上和人：“動画像中のベストショット抽出のための簡易顔認識手法” 電気関係学会東海支部連合大会講演論文集 No.776 , p388(2001.11).
- [5] 三田、伊藤、伊東、小塚、山本：“運転中のドライバの状態検知-意識低下と視線方向の検出-”、第3回動画像処理実利用化ワークショップ講演論文集、pp.102～105(2002)
- [6] 山口、浅山、戸田、富永、輿水：“動画像からの顔部品認識によるアイコンタクトカメラ顔映像生成”、第3回動画像処理実利用化ワークショップ講演論文集、pp.68～73(2002)