

色情報を用いた領域抽出法

中 須 英 輔
NHK 放送技術研究所

画面の合成技術はテレビ番組の制作上重要な役割を果たしているが、対象物を抜き出すために青色などの一様な色の背景が必要であったり、繁雑な手作業で画像の切り出しを行ったりしている。このため、どのようなシーンの画像に対しても有効で、効率的な画像の切り出し法の開発が望まれている。

ここで提案する方法では、まず照明条件などによって変わる「明るさ」と物質固有の色相や彩度に関係する「色」の情報をそれぞれ考慮して、繰り返ししきい値処理により領域分割を行った。さらに領域分割して得られた境界から弛緩法を用いたエッジの強調で望ましい輪郭線を求めた。

この方法は特定の絵柄やシーンの画像に限定されず、多くの画像に対して効率的に領域の抽出が行えるものである。

A Region Extraction Method using Color Information

Eisuke NAKASU

NHK Science and Technical Research Laboratories
1-10-11, Kinuta, Setagaya-ku, Tokyo, 157, Japan

This paper describes a region extraction method for montage.

Some montage techniques have been used at TV program production, but these methods require the background with a particular color or the troublesome operation. A practical and flexible way to extract a certain object from the natural scene is desired.

We propose a region extraction method using the color information that is the most significant factor when we understand the natural scene. This method consists of two procedures: (1) to segment the scene into regions based on three histograms of color feature, considering the chromaticity and luminance; (2) to confirm the correct contour among boundaries of the regions by relaxation method.

Experimental results indicate that this method is effective to obtain the contours of many kinds of objects.

1. はじめに

テレビ番組の映像制作ではコンピュータによって生成したCG(Computer Graphics)画像やDVE(Digital Video Effect)などのデジタル映像特殊効果装置で加工した映像が日常的に用いられ、様々な演出効果が得られるようになってきた。

テレビ画像の一部を自由に抜き出して他の画像にはめ込むことができれば、数台のカメラで撮った映像を1枚の画像に再構成したり、画像データベースを有効に活用でき、番組制作の上でたいへん大きな効果が期待できる。現在、この合成手法として背景の青色などの特定の色で分離するクロマキー、実写画像を見て人間が分離のためのキー図形を入力するビデオマット¹⁾などが実用化されている。

しかし、これらの方法は特殊な色の背景を必要としたり、画像の切り出しのために複雑な手作業を必要としたりしており、どのような絵柄やシーンの画像に対しても有効で効率的な切り出し法の開発が望まれている。

対象が限定されない複雑なシーンの自然画像を扱う場合、画像の持つ様々な情報を有効に利用する必要がある。このうち色の情報は物体を見分ける上で最も大きな役割を果たしており、この色の情報を使って画像を領域に分割する方法を検討した。

ここでは、色の情報を照明条件などで変わる「明るさ」と物質固有の色相や彩度に関係する「色」にわけて扱い、さらに画像処理の画像の分割の代表的な手法であるエッジ検出法と領域分割法を相補的に使ったカラー画像の特定領域抽出手法について検討したので報告する。

2. カラー画像の分割

2-1 画像の分割

画像の分割の手法は、大別すると次の2つの方法に分けられる。

- (1) 対象物の境界線を抽出する エッジ検出
- (2) 対象物内の均一な性質に注目して部分画像に分割する 領域分割

エッジ検出は異なる対象物間の境界部分では濃度や色、テクスチャーなどの特徴が急激に変化するという性質から、微分オペレータなどによって境界線を求める方法である。この方法は局所的な情報を良く示すが、微分オペレータによるエッジ検出だけでは雑音などによってあいまいな部分が生じやすい。このため、検出結果を2値化すると線がとぎれたり、余計な枝がでたりして、望ましい閉じた曲線を得ることは難しい。

領域分割は1つの対象物に対応した領域内では特徴がほとんど変化しないという仮定のもとに、特徴空間でクラスターを求める方法である。幾つかの特徴量について画像のヒストグラムを計算し、その分布状態から同じ特徴を持つ領域に分割する手法が代表的である。この方法で抽出した領域の境界線は必ず閉じており、複雑なシーンの解析に多く用いられる。しかし、大きな領域の特徴が把握しやすい反面、小さな領域の特徴が埋もれてしまったり、あいまいな特徴の部分が抽出できなかったり、過剰に分割したりすることがある。

このようにそれぞれの方法は特有の長所短所を持っており、互いに相補的に用いることが必要である。

2-2 色の情報の扱い

色の情報を数値的に扱う場合、人間の色に対する視覚上の性質とカラー画像の3次元色空間上の分布について考慮しなければならない。

我々がカラー画像を見て、そのシーンを認識する際、“赤い花”、“青い服”といったように色によって物体を識別することが多い。「色」は人間が画像を認識するうえで最も大きな役割を果たしているといえるだろう。

色の性質は人間の視覚上、「明度」、「彩度」、「色相」の3属性に分けられる。このうち、「明度」即ち「明るさ」は物体の表面の状態(滑らかさなど)や形状、あるいは物体と光源、視点の相対的な位置関係によって変化する。一般に「色」と呼ばれているのは色の3属性のうち「彩度」と

「色相」に関係した量である。この「色」は物質によって固有の性質としてとらえることができる。

次に、この「明るさ」と「色」が R, G, B 値にどのように表われるか色空間上の分布から見てみる。

任意の色 (F) は図 1 に示されるように赤 (R), 緑 (G), 青 (B) からなる 3 次元の色空間上の点で表現される。

$$(F) = R(R) + G(G) + B(B) \quad (1)$$

$$0 \leq R \leq 1, 0 \leq G \leq 1, 0 \leq B \leq 1$$

I T E 標準画像の一部 (330×330 画素)

[写真 1] の 3 次元色空間上の分布を 2 次元平面上に投影した 2 次元ヒストグラムを図 2 に示す。ここで、頻度は各点の明るさで表す。R, G, B の値は各 8 bit (0~255) である。

図 3 に R, G, B それぞれの 1 次元ヒストグラムを示す。

視覚上同じ「色」の領域は、 $R=G=B$ の無彩色 (黒 - 白) の軸を中心にバナナ状に分布している。「明るさ」はこの分布を無彩色軸に投影した値と考えることができる。図 2 にみられるように、R, G, B 相互には強い相関が認められ、R, G, B それぞれの画像を独立に処理することは適当でない。また、分布の形状が複雑であることから、簡単にクラスター化することも難しいことが解る。

2-3 カラー画像の分割

色の違いをどのように定義するかはカラー画像の微分をどうとらえるか、あるいはどんな集合を近い色の領域としてクラスター化するかなど、大変重要な問題である。

3 次元色空間上の 2 つの色 $F1(r1, g1, b1)$ $F2(r2, g2, b2)$ の距離 D は一般的に次のように定義することが多い。

$$D = \sqrt{(r1-r2)^2 + (g1-g2)^2 + (b1-b2)^2} \quad (2)$$

この距離は 3 次元 (R, G, B) 空間上の 2 点間の距離を表しているに過ぎず、人間が感じる色の違いや色空間上の分布の形状などを考慮してはいない。

3 次元色空間上の分布を 1 次元に投影して、そのヒストグラムから繰り返ししきい値を定めて領域分割する方法は Ohlander²⁾ らによって提案されており、カラー画像の

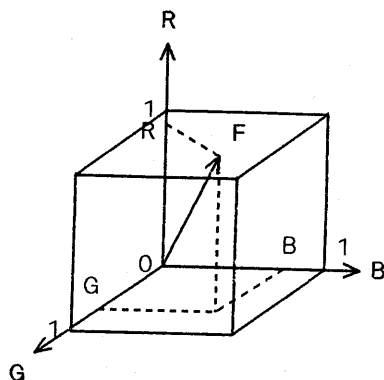


図 1 (R, G, B) 色空間



写真 1 実験に用いた画像

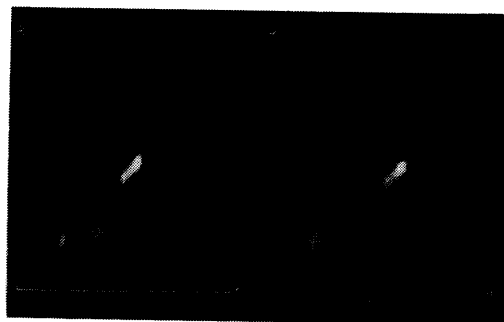


図 2 (R, G, B) 色空間上の分布

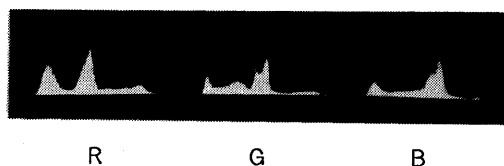


図 3 R, G, B 1 次元ヒストグラム

領域分割の有力な手段となっている。しかし、この方法は投影する軸の選択が大きな問題となり、処理を行うまえに相関の強い(R, G, B)座標系から相関が小さく視覚的な性質のはっきりした座標系へ変換する必要がある。この場合、人間の視覚に必ずしも厳密に一致した座標系を選ぶ必要はなく、ある意味では特徴を強調して扱ったほうが良いときもある。また、離散的なR, G, B値から(L, a, b)表色系や(明度, 彩度, 色相)などの座標系へ非線形な変換を行うと、明度の低いところで大きな誤差が生じたり、ヒストグラムにクシ状の見掛けのピークが現れたりする。

このような理由から、大田ら³⁾は次の3軸

$$X1 = (R + G + B) / 3 \quad (3)$$

$$X2 = R - B \quad (4)$$

$$X3 = 2G - (R + B) \quad (5)$$

が繰り返ししきい値処理法に相当であると報告している。この(X1, X2, X3)座標系はK. L. (Karhunen-Loeve)変換の結果から互いの軸が無相関になるように選ばれ、X1, X2, X3の順に分散が小さくなっている。

写真1の画像の(X1, X2, X3)座標系での分布と1次元ヒストグラムを図5, 6に示す。

3. 色情報を使った領域抽出法

3-1 色の違い

図4の座標系でX1軸は明るさを表す。X1軸に直交するX2, X3軸は彩度と色相に関係し、彩度S, 色相Hはそれぞれ図7に示すように

$$S = \sqrt{X2^2 + X3^2} \quad (6)$$

$$H = \arctan(X3/X2) \quad (7)$$

にほぼ対応する。

ここでは色の違いを2-2で述べた「明るさ」の差と「色」の差に分けて考える。

F1(X11, X21, X31)とF2(X12, X22, X32)の2点の「明るさ」の差ΔLは

$$\Delta L = |X11 - X12| \quad (8)$$

で表される。

一方、「色」の差ΔCは

$$\Delta C = \sqrt{(X21 - X22)^2 + (X31 - X32)^2} \quad (9)$$

で定義する。

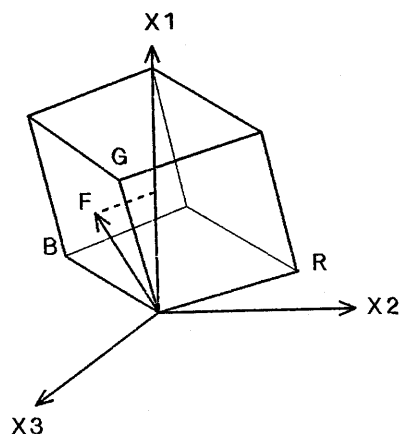


図4 (X1, X2, X3)色空間

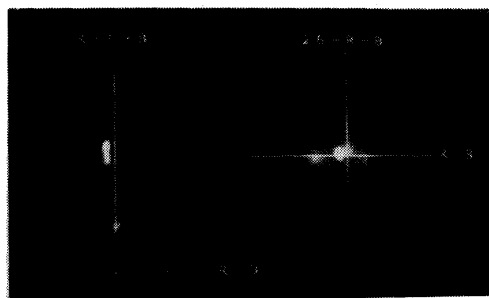


図5 (X1, X2, X3)色空間上の分布

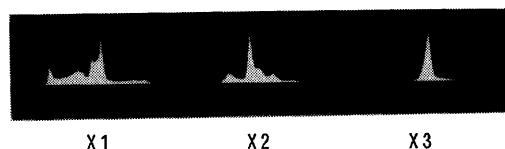


図6 X1, X2, X3 1次元ヒストグラム

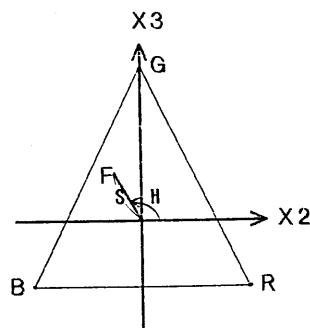


図7 X2-X3平面と彩度, 色相の関係

2つの色の間の距離は「明るさ」と「色」の差のうち大きいものを考えることとする。

$$D = \text{Max}(\Delta L, \Delta C) \quad (10)$$

画像の微分値を計算する場合、たとえば3×3などの大きさの微分オペレータを掛けて求めることが一般的である。この場合のX、Y方向の出力を ΔxX 、 ΔyX としたとき、 ΔL 、 ΔC はそれぞれ

$$\Delta L = \sqrt{(\Delta xX1)^2 + (\Delta yX1)^2} \quad (11)$$

$$\Delta C = \sqrt{\frac{(\Delta xX2)^2 + (\Delta yX2)^2}{*} + \frac{(\Delta xX3)^2 + (\Delta yX3)^2}{*}} \quad (12)$$

で計算される。

3-2 繰り返ししきい値処理による領域分割

大田らは $X1$ 、 $X2$ 、 $X3$ の各値の色彩的な意味を考慮せず、領域分割においてそれぞれを同等に扱っている。しかし、「明るさ」を表す $X1$ と物質に固有な性質の「色」を表す色差 $X2$ 、 $X3$ とを同等に扱って分割することは好ましくない。

人間が同じ「色」の領域と認める部分でも、実際には明るさが変化している場合が多く、色($X2$ 、 $X3$)で分割した後に、さらに明るさ($X1$)を使って分割すると過剰に分割する傾向がある。このため、彩度がある程度高く、色のはっきりしている場合には $X1$ によるしきい値処理は行わないこととした。本分割法は図8に示す方法で分割を行い、各領域にラベル付けをする。

(1) 色差 $X2$ 、 $X3$ のヒストグラムから最も顕著なピーク(ヒストグラムの山)を探し、その両側の極小となる値(谷の値)をしきい値としてこのピークに属する画素を抽出する。即ち、 X におけるヒストグラムの値を $H(X)$ 、 $X2$ および $X3$ の最大となる点を $X2M$ 、 $X3M$ この両側の極小になる点を $X2L$ 、 $X2R$ 、 $X3L$ 、 $X3R$ としたとき

$$(H(X2L) + H(X2R)) / H(X2M) \quad (13)$$

$$(H(X3L) + H(X3R)) / H(X3M) \quad (14)$$

の値の小さい色差について極小となる点でしきい値処理する。

(2) 次に(1)で抽出した画素についてもう1つの色差のヒストグラムを計算し、同様に最も顕著な単峰のピークに属する画素を取り出す。

(3) (1)、(2)で決定した $X2$ 、 $X3$ の各ピークが原点を含んで括がっている場合、即ち

$$X2L \leq 0 \leq X2R$$

$$X3L \leq 0 \leq X3R$$

のときは彩度が低く「色」による分割ができないので、明るさ $X1$ で同様のしきい値処理を行う。

以上の操作で抽出した領域にラベル付けをし、この領域を除いた残りの部分について(1)～(3)の処理を繰り返し行う。抽出した領域の面積が一定の値(IAREA)以下になったら分割を終了する。

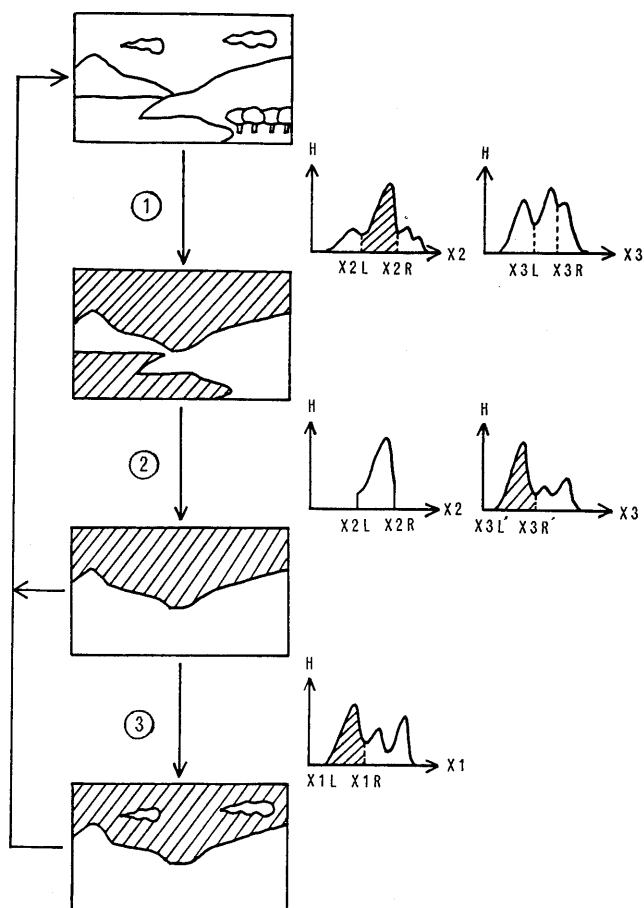


図8 ヒストグラムを用いた領域分割

3-3 処理手順

本研究では、画面の中から例えば、そこに登場する「人間」だけを抜き出すような手法を検討した。人間は顔や頭髮、衣類、手足などの異なった性質の領域から構成され、我々は画像を見てそれぞれの領域を統合して「人間」という一つの対象物を認識することができるが、現在の画像処理技術でその様な把握はなかなか難しい。

このため、ある程度対話的に人間がプランを指示してやる必要がある。しかし、操作が繁雑になることを避けるために、与えるプランは最小限にすべきである。本手法では操作者が抜き出したい対象の輪郭線近傍を最初に指示し、このガイド内を処理することとした。

以下に、図9に示す処理手順について具体的に述べる。

最初に操作者が画面を見て、抽出したい対象物のおおよその輪郭をガイドとして与える。

[手順1]

R, G, B 画面を (3)~(5) 式を用いて、 X_1, X_2, X_3 画面に変換する。

[手順2]

3-2で示したヒストグラムを用いた繰り返ししきい値処理によって領域の分割を行う。

[手順3]

分割できなかった画素について周辺画素との比較をし、領域統合を行う。

こうして分割した領域の境界と統合できなかった部分を輪郭候補とする。

[手順4]

輪郭候補点について、 X_1, X_2, X_3 の値から微分値を求め、弛緩法を用いたエッジの強調を行う。

[手順5]

4の結果を2値化、細線化後、枝の除去、とぎれた部分の回復を行い、輪郭線の整形をする。

[手順6]

得られた輪郭線の内部の領域を抽出する。

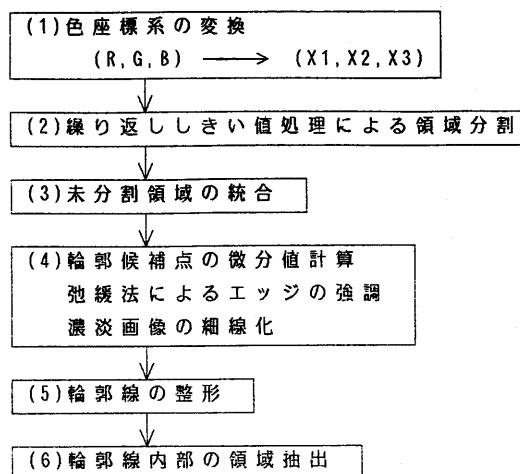


図9 処理手順

いろいろな性質の部分から構成され、複雑な形状をしている対象物を抽出するためには領域分割とエッジ検出をうまく使いわける必要がある。

ここではまず、3-2で述べた方法で大域的に色による領域分割を行う。

領域分割の結果、分割できなかったあいまいな色の画素は領域と領域の境界部に多く存在する。そこで、未分割の画素について、その8近傍の画素のうち最も近い色の画素がすでに領域に分割されていれば、注目画素をその領域に統合する。

このように領域の分割と未分割画素の統合を行うが、ここでは各領域の画像上の位置関係を考慮していないので離れたところに小さな領域が散在したり、隣接する似た色や明るさの領域を別の領域として分割したりすることがある。このうち、小さな領域の境界線は削除することとした。

近傍画素とあまり特徴が変わらない見掛上の輪郭線については、3-1で述べた方法で微分値を求めた後、弛緩法を用いたエッジの強調⁴⁾により、確かな輪郭線だけを取り出した。

このエッジの強調では注目する画素と周囲の近傍画素とのエッジの大きさと方向を比較して、反復演算によって局所的な矛盾を取り除き、輪郭線としてつじつまの合う部分を強めてゆく。

初期エッジ確率は X_1, X_2, X_3 画面にSobel

オペレータを掛け(10)～(12)式で計算した。
初期エッジの方向D0は

$$D0 = \arctan(\Delta y X1 / \Delta x X1) \quad (15)$$

とした。

この初期値から繰返し確率更新を行うが、その結果、エッジが太くなっていくことがある。このため、近傍画素で既にエッジと認められる画素が検出されたところでは更新手順をものはや行わないように、エッジ確率をあらわす画面を濃淡画像として扱って細線化を行う。この濃淡画像の細線化⁵⁾では、尾根線近くの単純傾斜面上の点を削除する。この操作を確率更新の度に行うことにより、更新のためのパラメータの選定が容易になり、大局的にあいまいな所についてのみ確率更新を行うので収束がはやい。

4. 実験結果と検討

4-1 繰返ししきい値処理による領域分割

(R, G, B), (X1, X2, X3)座標系において3軸を同等に扱って領域分割した結果をそれぞれ図10, 11に、3-2で述べた手法により領域分割した結果を図12に示す。ここでは分割した領域の境界と未分割の部分を白で表している。

表1 分割法による分割領域数と
未分割面積(比)の比較

分割法	領域数	未分割面積
(R, G, B)	100	19025 (17.5%)
(X1, X2, X3)	63	3527 (3.2%)
本方式	44	354 (0.3%)



図10 (R, G, B)分割結果



図11 (X1, X2, X3) 分割結果



図12 本手法による分割結果

分割の終了条件は抽出面積が5画素以下になったときとした。

それぞれの場合の分割数と分割できなかった部分の面積を表1に示す。

本方式では過剰な分割が避けられ、分割できなかった面積も最も小さい。

(R, G, B)座標系による分割は人間の感じる色に対応して行っていないため、背景や髪の毛のように平坦な色の部分を除いて分割がうまく行っていない。(X1, X2, X3)の各軸を同等に扱った分割では領域内で明るさに変化のある顔や花びらのような対象物で適切な分割が行われていない。本手法は各軸と色の3属性との対応を考慮して分割しているので、このような顔や花びらについても比較的うまく分割できており、人間の認識する領域に良く合致している。しかし、なだらかに明るさの変化している背景を2つの領域として分割している、光沢のある胸パッチや植物の葉の分割が過剰にされているなど、この処理だけでは十分な結果は得られない。

4-2 領域抽出例

図13のガイド内で領域分割して得られた輪郭候補を図14に示す。ここで、分割数は24、未分割点は223画素であった。このうち連結した画素数が20以下のものは削除している。3回の繰返し領域統合により未分割点は98になった。

図15に輪郭候補点の微分結果、図16に弛緩法で2回エッジを強調した結果を示す。

さらに、図16の結果を2値化して輪郭線を整形した結果を図17に、輪郭内の対象物を抽出した結果を図18に示す。

このように人間という対象がほぼうまく取り出せた。女性の肩の部分について輪郭線がギザギザしているところがあるが、これは境界部があいまいで近くに輪郭線と間違えやすい線があったために生じている。このほか手元で一部輪郭を誤っているが、これは輪郭線の整形上の問題である。

5. まとめ

色の情報を照明条件などで変化する「明るさ」と彩度、色相に関係する物質に固有な「色」に分けて考えることにより、色の違いによる領域分割が従来より適切に行えた。

ここで提案した特定対象物の抽出法は最初に大雑把な輪郭線をガイドとして与えてやるだけで望ましい輪郭を求めることができるので、画像の切り出し作業の手間が大巾に軽減されるであろう。さらに輪郭線の図形的な情報や色の変化の情報を用いて、より適切な輪郭線の決定が行えるように改良をはかってゆきたい。

今後、一定時間間隔で連続したフレームの動画像について、過去のフレームの輪郭線をガイドとして使うことにより画面上で動く対象物の輪郭を追跡する方法について検討していくつもりである。

参考文献

- 1) A. Iwata, Y. Monjo, T. Niikura, H. Tamura: A new method of video synthesis developed by NHK, SMPTE Journal, July, 1986, 702-704
- 2) R. Ohlander, K. E. Price, R. Reddy: Picture segmentation using a recursive region slitting method, CG&IP, 8, 1978, 313-333
- 3) Y. Ohta, T. Kanade, T. Sakai: Color information for region segmentation, CG&IP, 13, 1980, 222-241
- 4) B. J. Schachter, A. Lev, S. W. Zucker, A. Rosenfeld: An application of relaxation methods to edge reinforcement, IEEE Trans., SMC-7, 11, 1977, 813-816
- 5) 横井、鳥脇、福村: 濃淡画像における特徴点の抽出と細線化の一手法、信学会(D), 58-D, 1975, 601-608



図13 指定ガイド

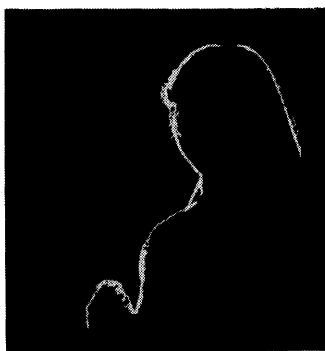


図14 領域分割結果

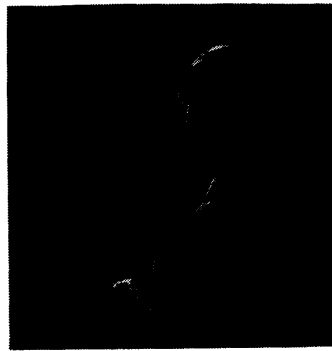


図15 微分結果



図16 弛緩法によるエッジ強調

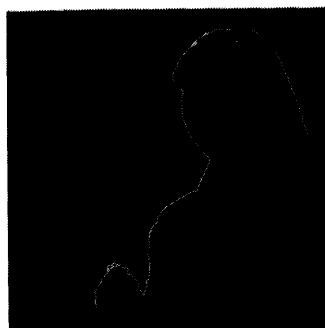


図17 整形後の輪郭線

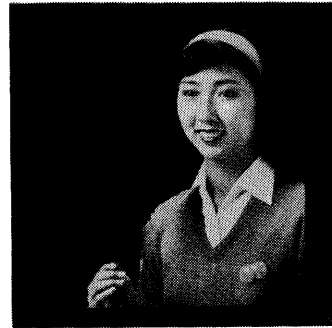


図18 領域抽出結果