



Title	分裂アクティブネットを用いた色情報による領域抽出
Author(s)	吉野, 和芳; 真木, みお; 川嶋, 稔夫; 青木, 由直
Citation	北海道大學工學部研究報告, 167, 87-96
Issue Date	1994-01-14
Doc URL	http://hdl.handle.net/2115/42391
Type	bulletin (article)
File Information	167_87-96.pdf



[Instructions for use](#)

分裂アクティブネットを用いた色情報による領域抽出

吉野 和芳 真木 みお
川嶋 稔夫 青木 由直

(平成 5 年 8 月 31 日受理)

Color-Based Region Extraction Using Splittable Active-Net

Kazuyoshi YOSHINO, Mio MAKI,
Toshio KAWASHIMA and Yoshinao AOKI

(Received August 31, 1993)

Abstract

In this paper, we propose color-based region extraction algorithms using splittable active-net. In the algorithms, we use *histogram backprojection function*, which maps the color of a image pixel to the ratio histogram of the target, as the image energy of splittable active-net. The convergency of splittable active-net is improved by determining its initial position by rough estimation of the target. Additionally, we extend the histogram backprojection function to team identification tasks to apply the algorithms to multiple-team-multiple-target images. The validity of our approach is shown by the experimental results for soccer T. V. programs.

1. はじめに

コンピュータビジョンの研究における目標は、次の 2 つに大別されるという考え方がある。1 つは、「興味のある対象物が、どこにあるのか?」という疑問を解決することであり、もう 1 つは、「そこに、何があるのか?」という疑問を解決することである。前者は、主として興味の対象となる目立つ性質がどこにあるかを見つけることに主眼があり、後者は、特定の位置にある対象物の形状や性質から、それが何であるかを同定することにある。Ballard は、従来の研究の反省から、この 2 つのタスクを分けて考えることで、コンピュータビジョンの手法が単純化されると述べている¹⁾。

ところで、従来のビジョンの研究では、画像全体の構造化、例えば、セグメンテーションが用いられてきた。この手法には、ターゲットの領域では、濃淡値や色などの特徴がほぼ均一であり、それ以外の領域との境界部分では、その特徴が急激に変化するという仮定に基づいたエッジ抽出法や領域分割法などがある。

エッジ抽出法は、エッジの連結関係などからターゲットの形状を抽出する手法である。しかし、雑音やテクスチャによって、エッジ点の欠落や無いはずのエッジが現われることがあり、安定し

て行なうことは困難である。そこで Kass らは、輪郭モデルを用いて、エッジ情報にモデルの滑らかさなどの制約条件を付加することによって、正則化問題として扱い、形状抽出を行なう SNAKES を提案し、安定化を図った²⁾。このアプローチは、形状の正確、安定な抽出に目的があり、後者のコンピュータビジョンに近い。

一方、領域分割法では、近傍領域を大局的に評価するため、雑音やテクスチャなどのような局所的な濃度値の変化に対してロバストであり、エッジ抽出法に比べ、有効であるが、背景や他のターゲットとの濃度変化になだらかな部分があった場合、それらが統合されてしまうことがある。この問題を解決したのが、坂上らが提案したアクティブネット³⁾である。この手法は、SNAKES におけるモデルを2次元の網のモデルへと拡張することにより、領域の情報を用いることができ、SNAKES に比べ、安定した形状抽出が可能である。アクティブネットのアプローチでは、興味の対象を表す指標が与えられるならば、これをエネルギー関数として組み込むことで、前者のビジョンと後者のビジョンを継ぎ目なく実現することができ、Ballard の指摘した問題点を回避できると考えられ、興味深い。

我々が先に提案した分裂アクティブネット⁴⁾では、上述したアクティブネットの収束過程において、網を構成している格子点間のリンクを切断することにより、網のモデルの構造を変形させることができ、ターゲットの複雑な形状や同一画像内から同時に複数のターゲット領域の抽出が可能である。従来のモノクロ画像に対するアクティブネットでは、興味の対象を表す指標が制限されるため、抽出できるターゲットが、限定されてしまうが、カラー画像を用いれば、それらの問題を解決できる。例えば、ターゲットが単色の物体である場合は、カラー画像でも抽出可能である。しかし実世界では、むしろ単色の物体の方がまれで、ほとんどの物体は、模様があたり、複数の色で構成されている。そのため、ターゲット中に含まれている色が、背景領域に含まれていることも多い。このため、背景領域もターゲットと見なされることがあり、安定してターゲット領域を抽出することは、不可能である。

Swain¹⁾らは、ターゲット画像、及び入力画像のヒストグラムを用いて、それらのヒストグラム値の比をとることにより、ターゲットか否かという確率を求め、その確率からターゲットの特徴となる色を検出する Histogram Backprojection を提案し、その手法により検出された色を手がかりとしてターゲットの位置探索を行なっている。この手法では、ターゲットの持つ色が背景領域にも多く含まれている場合、その色は、ターゲットとしての確率が低くなり、出力結果は、ターゲットとしての確率の高い部分が、疎らに散らばった状態で得られる。そのため、ターゲットの位置は大まかに検出可能であるが、ターゲットの形状を抽出することは困難である。

そこで本稿では、分裂アクティブネットと Histogram Backprojection との長所を生かし、それらを併用することにより、ターゲットの色情報を指標とし、ターゲットの位置及びその形状を抽出する手法を提案する。更に、複数のグループに分けられる対象物が混在する画像において、特定のグループの対象物だけを抽出する方法についても検討する。

以下、2章では、分裂アクティブネットの手法について簡単に論じ、3章では、Histogram Backprojection の手法、及び、その評価関数を分裂アクティブネットへ導入する方法について述べる。更に、領域抽出の高速化を図るため、分裂アクティブネットの初期位置決定法についても説明する。また、4章では、複数グループの解析として、サッカー中継を例にとり、特定のチームの選手領域だけを抽出する手法について論ずる。5章では、本手法を用いた実験を行なうことによって、本手法の有用性を示す。

2. 分裂アクティブネットの手法

分裂アクティブネットは、格子状の網のモデルを用いて画像内の領域抽出を行なうアクティブネットの手法において、その網を構成している格子点間のリンクを収束過程で切断することにより、複雑な形状の物体や複数の物体の領域を抽出可能とした手法である。その領域抽出は、アクティブネットと同様に、次のようなエネルギー関数を最小化することによって行なわれる。

$$E_{net} = \int_0^1 \int_0^1 (E_{int}(\mathbf{v}(p, q)) + E_{image}(\mathbf{v}(p, q)) + E_{con}(\mathbf{v}(p, q))) dpdq \quad (1)$$

ここで $\mathbf{v}(p, q) = (x(p, q), y(p, q))$ は、網を構成する格子点の位置を示す。また、積分式内の第1項は、網を収縮させ、かつ滑らかに保とうとするエネルギー（内部歪みエネルギー）であり、次のように定義される。

$$E_{int} = (\alpha(|\mathbf{v}_p|^2 + |\mathbf{v}_q|^2) + \beta(|\mathbf{v}_{pp}|^2 + 2|\mathbf{v}_{pq}|^2 + |\mathbf{v}_{qq}|^2)) / 2 \quad (2)$$

ただし添え字は、その文字についての偏微分を表し、 α 、 β は、それぞれ収縮性、滑らかさに対する重み係数である。第2項は、モデルを画像中のある領域へと引きつける力（画像の適合性エネルギー）であり、第3項は外部からの強制力に対応するエネルギーである。

これまで、モノクロ画像において、濃淡値により画像の適合性エネルギーを定義していたが、カラー画像では異なる色のものでも、モノクロ画像に変換すると、同じ濃淡値を持つことがあり、濃淡値だけによるターゲットの特定は、困難である。例えば、図1(a)の画像において、カラー画像では、それぞれの本が、図に示しているような色をしているため、それぞれを区別することができるが、モノクロ画像で表示すると区別がつかなくなってしまう。

そこで本稿では、ターゲットの持つ色情報を用いて、ターゲットの領域抽出を行うことを考える。しかし、実世界では、物体の持つ色が、背景にも含まれていることが多く、また、1つの物体にも複数の色があるため、特定の色だけでターゲットを表現することは困難である。そこで我々は、ターゲットに含まれている色の中からターゲットの特徴となる色を検出する Histogram Back-projection に注目し、その評価関数を分裂アクティブネットの画像の適合性エネルギーとして導入することによってカラー画像における領域抽出を行う。

3. Histogram Backprojection の導入

3.1 Histogram Backprojection

Histogram Backprojection は、ターゲットに含まれている色の成分から、そのターゲットの特徴となる色を検出し、その色を手がかりとして入力画像内からターゲットの存在する位置を探索する手法であり、次のような処理で行なわれる。

ターゲットの画像、及び入力画像の多次元ヒストグラムをそれぞれ T_i 、 I_i とする。ここで、ターゲット画像とは、検出対象だけを含む画像のサンプルを指し、入力画像とは、ターゲット画像と同じ対象を一部に含む画像を指す。ただし、入力画像は、検出対象の姿勢などが、変化していることもある。このとき、ターゲットの画像に含まれる色が、入力画像内にどれだけの割合で含まれているかという比を次式の評価関数から求める。

$$R_i = \min\left(\frac{T_i}{I_i}, 1\right) \quad (3)$$

次に、上式で求められたそれぞれの色における比 R_i を入力画像に逆投影することによって、出

力画像が得られる。

この処理の結果を図1に示す。ここで(c)の出力画像は、(2)式で求められた比を正規化し、白黒反転させた画像で、黒い色ほどターゲットであるという確率が高い。

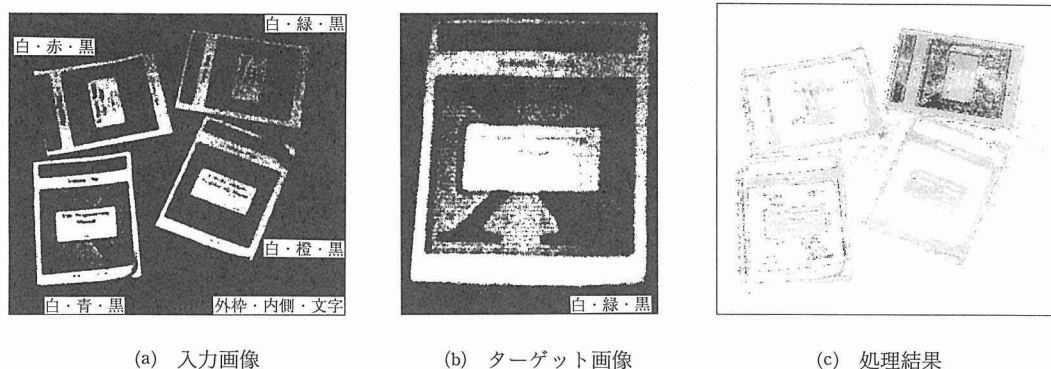


図1 Histogram Backprojection の処理結果

3.2 ターゲット画像の作成

図1(c)では、ターゲット（ここでは、本）領域だけでなく、背景領域においても黒っぽい部分があることから、その領域もターゲットの一部として見なされてしまう。そのため、ターゲットの限定が困難となる。これは、ターゲット画像の切り出し方に依存して生ずると考えられる。

ターゲット画像は、ターゲットが映っている画像からターゲットの部分の切り出すことによって作成される。そのため、ターゲット画像内にその画像における背景の一部も入ってしまい、入力画像内では背景となる部分もターゲットと見なされて、大きな値をとってしまう。この問題を回避するため、ターゲット画像内で背景と思われる画素は、あらかじめ取り除く必要がある。

ターゲット画像は、ターゲットをその画像の中心付近になるように作成してあるものとするれば、ターゲット画像における上下左右2 pixel 幅の枠に相当する画素を背景であると仮定し、それらの画素のR, G, B値と同じ値を持つ画素をターゲット画像から取り除いた。この処理を行った Histogram Backprojection の処理結果を図2に示す。このとき用いたターゲット画像及び入力画像は、図1(a)と同様の画像である。

図1(c)では、背景領域内でターゲットと見なされている部分（グレーになっている部分）が、右上の本の周囲にあったが、その部分が、

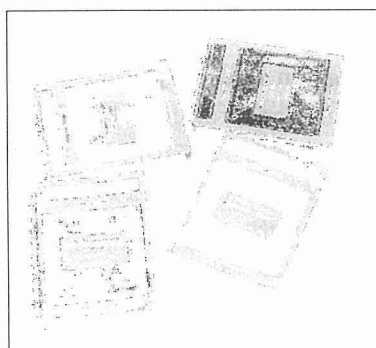


図2 背景除去後の処理結果

図2では、区別されている(図中では、白になっている)。このように、ターゲット画像作成時に、あらかじめターゲット画像から背景と思われる色を取り除くことによって、ターゲットの形状を

より正確に抽出することができる。

3.3 Histogram Backprojection を利用した分裂アクティブネット

Histogram Backprojection の分裂アクティブネットへの導入は、評価関数 ((2)式) を分裂アクティブネットにおける画像の適合性エネルギーとして、次式のように定義することによって行なう。

$$E_{image} = wR_i(x, y) \quad (4)$$

ここで w は、画像の適合性エネルギーの重み係数を表す定数であり、内部格子点(網の内側の格子点)では、 $w < 0$ とし、最外郭格子点(網のもっとも外側の格子点)では、境界条件として $w > 0$ とする。

このエネルギーによって、網はターゲットである確率の高い領域へと引きつけられ、ターゲット領域の抽出が可能となる。

3.4 分裂アクティブネットの初期位置決定

分裂アクティブネットによる領域抽出は、「画像内のどこかにターゲットがある」ということから、初期位置として画像全体にネットを張り、上述したエネルギー関数を反復計算によって最小化することにより行われる。そのため、画像が大きく、複雑になると計算コストも高くなる。しかし、本手法においては、Histogram Backprojection により、ある程度ターゲットの存在する位置がわかることから、分裂アクティブネットの初期位置をターゲット付近に決定し、その位置から対象物の輪郭に収束させてやることで、計算コストを削減し、高速化を図っている。

その初期位置は、

$$G_\sigma * R_i > Threshold. \quad (5)$$

を満たす領域を覆う位置とする。

ここで左辺は、Histogram Backprojection の評価関数 R_i の分布と分散 σ のガウスフィルター G_σ との畳み込みを表す。

上述したように、分裂アクティブネットにおける画像の適合性エネルギーとして、Histogram Backprojection の評価関数から定義し、分裂アクティブネットの初期位置を与えることによって領域抽出した結果を図 3 に示す。ただし、この図(a)は、(5)式のしきい値を0.87とした場合の結果であり、(b)は、0.92とした結果である。また、網は、図中に黒い線で描いてある。

この結果のように、(5)式のしきい値を緩くすると、目的のターゲット以外の領域も抽出され、逆にきつくすると、ターゲットの一部分しか抽出できない。したがって、ターゲット領域全体を抽出するためには、分裂アクティブネットの初期位置が、ターゲット領域だけを覆い、かつターゲット領域全体を覆うようなしきい値を設定しなければならない。

3.5 膨張力を持つ分裂アクティブネット

前節では、分裂アクティブネットの初期位置を決定するための問題点を挙げたが、本節では、その問題点を違った観点、つまり、分裂アクティブネットの収束過程から考えてみる。

分裂アクティブネットの動作は、画像の適合性エネルギーがない(値が0)場合、内部歪みエネルギーによって、網は小さくなる方向へと移動し、画像内にターゲット領域がある場合は、画

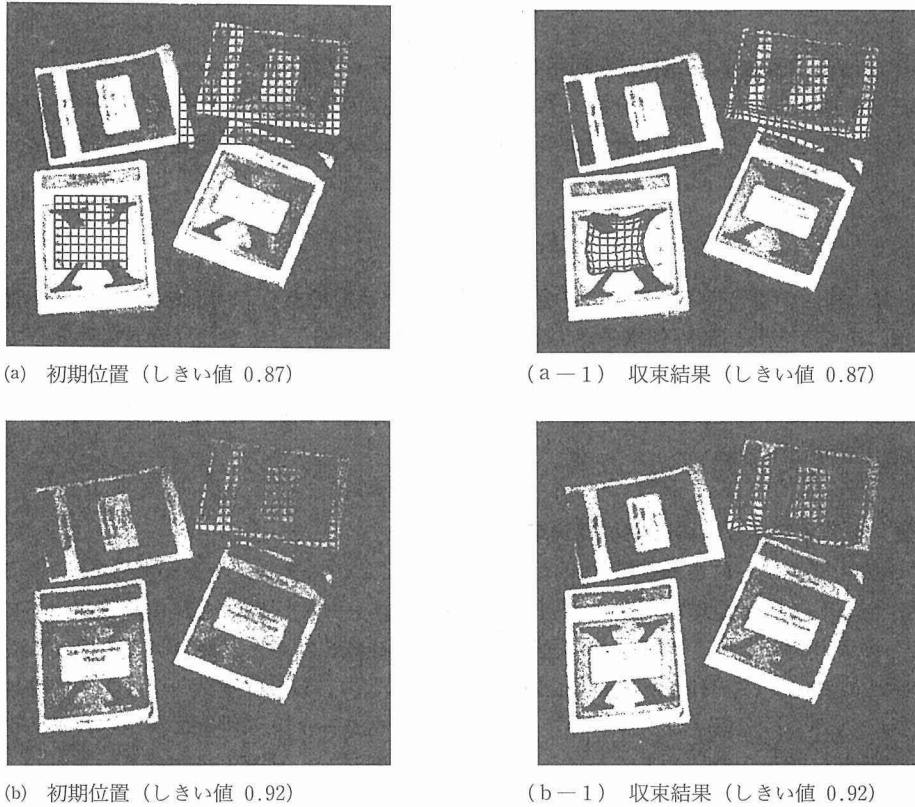


図3 分裂アクティブネットの収束結果 I

像の適合性エネルギーが得られ、その領域へと移動する。そして、最外郭格子点がターゲットのエッジ点上に来たとき、外向きの力が作用することによって、網の収束が停止し、領域抽出が行われる。そのため、最外郭格子点がエッジ点を越え、ターゲット領域内に入った場合、新たなエッジ点に来るまでは、網は、内部歪みエネルギーの作用で、小さくなる方向へと進んでいく。したがって、(5)式のしきい値を高くした場合、最外郭格子点は、初期位置の状態ですでにターゲット領域の内部にあるため、図3のような結果となってしまふ。そこで今回は、最外郭格子点がターゲットの内部にある場合、その最外郭格子点を外側に移動させることによって、越えてしまったターゲットのエッジ点まで戻している。

はじめに、最外郭格子点の位置が、ターゲットの内部か外部かという判断は、最外郭格子点の位置における画像の適合性エネルギーの値によって可能である。つまり、画像の適合性エネルギーは、(4)式で定義したように、Histogram Backprojectionの評価関数となっているので、その値が0より大きければ、ターゲットの内部であるという判断ができる。このとき、評価関数の値を持つ画素が、雑音であることも考えられるが、雑音の場合は、その画素の近傍に値を持つ画素がないため、滑らかさの制約条件によって、取り除かれる。

次に、最外郭格子点を外側へ移動させる方法について考えと、上述したように、分裂アクティブネットの最外郭格子点は、ターゲットのエッジ上に来たとき、画像の適合性エネルギーによって、外側への力が作用する。ここでは、そのことを利用し、全ての最外郭格子点が、その位置に

よらず、エッジ上にあるとし、画像の適合性エネルギーを与える。つまり、最外郭格子点の位置以外の周囲は、全て背景領域と仮定し、画像の適合性エネルギーを求める。このエネルギーによって、最外郭格子点は、ターゲット内部にあるとき外側へと移動し、背景領域にある場合は、従来通り、内側へと進んでいく。

以上のような操作によって、膨張力を与えた分裂アクティブネットによる収束結果を図4に示す。このとき、分裂アクティブネットの初期位置は、図3(a)と同様の位置である。

この結果のように、分裂アクティブネットに膨張力を与えることによって、ターゲットの内部に入ってしまった最外郭格子点を、外側へと移動させることが可能となった。これにより、適切なしきい値を設定する必要はなく、ある程度の高いしきい値を与えるだけで、ターゲットの領域抽出が可能である。

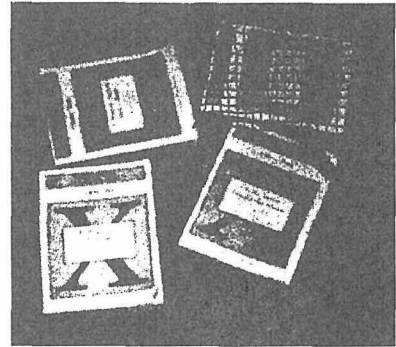


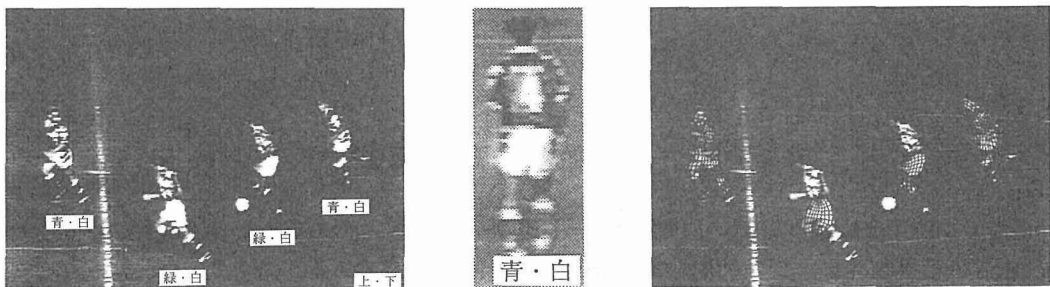
図4 膨張力を持つ
分裂アクティブネットの収束結果

4. 複数グループの解析

ここまでは、入力画像内に対象物が単一の場合について述べてきた。この章では、サッカー中継のシーンのように複数の対象物（選手）があり、その対象物が、複数のグループ（チーム）を成し、同一画像内に混在する場合、それらのグループを排他的に扱い、それぞれの対象物の領域を抽出することが必要となる。そこで、ここでは、サッカー中継のシーンを例に挙げ、本手法を用いて、それぞれのグループ分け（チーム分け）を行い、更に、そのグループ（チーム）に属する個々の領域（選手領域）を抽出する方法について述べる。

(3)式に示したエネルギーを導入した分裂アクティブネットを用いて、図5(a)の画像における選手（ユニフォームが青・白）領域の抽出を行った結果を同図(c)に示す。

この結果は、ターゲットとなるチームの選手以外に、相手チームの選手の領域までもが抽出された例である。このような結果となった原因、及び、その解決法について以下で述べる。



(a) 入力画像 (b) ターゲット画像 (c) 収束結果

図5 サッカー中継シーンにおける分裂アクティブネットの収束結果

4.1 評価関数

はじめに、図5(a)の入力画像における Histogram Backprojection の評価関数の値を正規化し、白黒反転させた画像を図6に示す。

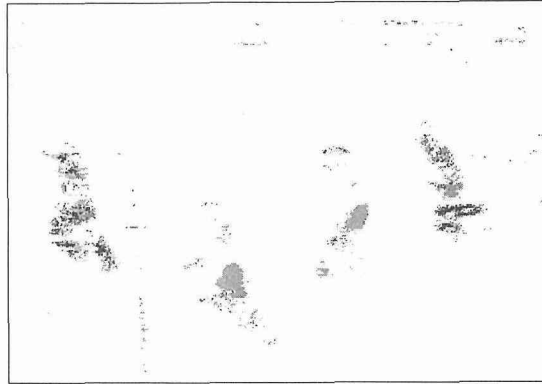


図6 Histogram Backprojection の処理結果

この例のように、ターゲットとなるチームの選手が複数いて、更に、ターゲットとなるチームのユニフォームの色と相手チームのユニフォームの色との間に同じ色がある場合、背景領域と両チームの選手領域との区別はできるが、どの選手がどのチームに属しているかということまでは、区別しがたい。これは、図6の結果から分かるように、ターゲットのチームの選手領域の評価値と相手チームの選手領域における評価値が、ほぼ同じ値になったためである。この原因としては、同一画像内に同じチームの選手が多くなったことにより、入力画像では、ターゲット画像内にある色の成分と同じ成分を持つ画素が増え、その色の評価値が下がったためと考えられる。この問題を回避するためには、Histogram Backprojectin の評価関数を変える必要がある。

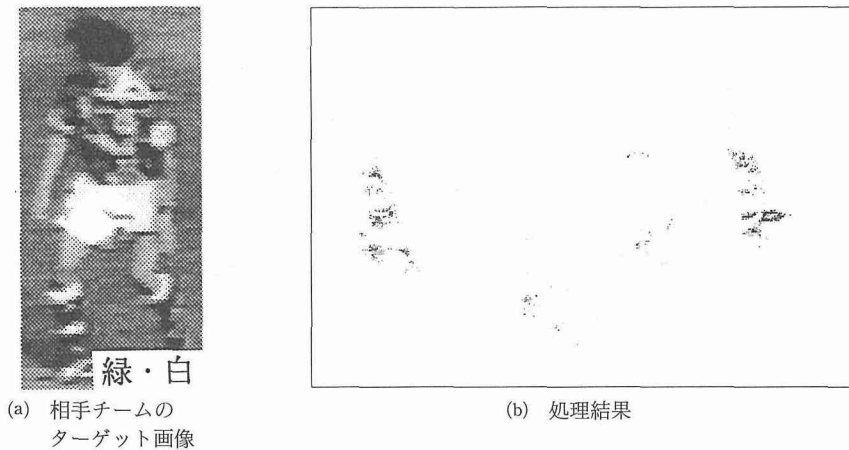
サッカー中継の画像においては、ターゲットとなるチームの選手領域を除いた領域内で、ターゲットと同じ色になる可能性があるのは、相手チームのユニフォームや選手の髪や皮膚など、ほとんどが相手チームの選手に関係しているものと考えることができる。そこで、ターゲット画像からあらかじめ相手チームの選手と共通した色を取り除くことを行い、(2)式で示した Histogram Backprojection の評価関数を次のように変更する。

$$R_i^* = \begin{cases} \min\left(\frac{T_i^A - T_i^B}{I_i}, 1\right) & , T_i^A \geq T_i^B \\ 0 & , T_i^A < T_i^B \end{cases} \quad (6)$$

ここで T_i^A 、 T_i^B は、それぞれAチーム、及びBチームにおけるターゲット画像のヒストグラムである。

図7に、相手チームのターゲット画像、及び、上述の評価関数によって求めた Histogram Backprojection の結果を示す。ここで用いた入力画像は、図5と同様である。

この結果からわかるように、この評価関数によって、ターゲットとなるチームの選手領域における評価関数の値をほぼ変えることなく、相手チームの選手領域内に現れていた中間的な評価値を持つ部分を除去することが可能となった。

図7 評価関数 R_i^* による処理結果

5. 実 験

Histogram Backprojection の評価関数を画像の適合性エネルギーとして導入した分裂アクティブネットを用いて、サッカー中継の1シーンにおける選手領域の抽出を次の手順で行なう。

1. 分裂アクティブネットの初期位置を

$$G_\sigma * R_i^* > Threshold. \quad (7)$$

で求める。

2. 分裂アクティブネットの画像の適合性エネルギーを

$$E_{image} = wR_i(x, y) \quad (8)$$

として、領域抽出を行なう。

この方法によって、選手領域の抽出を行なった結果を図8に示す。ただし、ここで扱った画像は、R, G, B, それぞれ256階調で、画像サイズは、入力画像が590×410, ターゲット画像が65×168とし、図5(a), (b)と同様の画像である。また、相手チームのターゲット画像は、図7(a)と同様の画像で、画像サイズは、85×180である。

分裂アクティブネットの初期位置は、4章で述べた手法によって、ターゲット画像の選手と同じチームの選手上にしかなく、収束した結果は、分裂アクティブネットに膨張力を加えたため、初期位置では選手領域の一部しか覆っていないが、最終的に選手領域全体を抽出している。しかし、初期位置を求める際、(7)式のしきい値によって、同一の領域であるのに、いくつか分割されてしまった場合などの問題点が残っている。

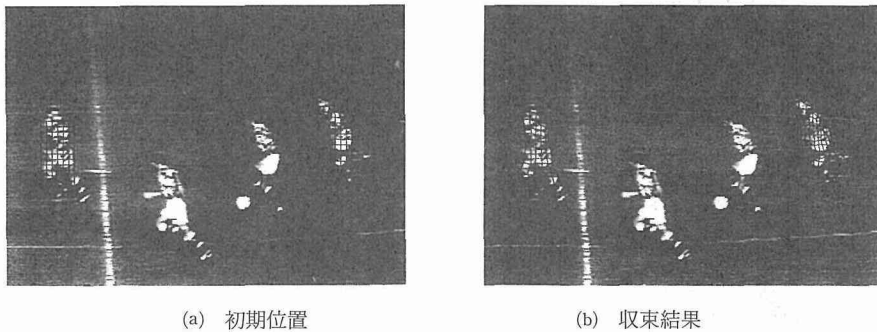


図8 サッカー中継シーンにおける分裂アクティブネットの収束結果II

6. おわりに

本稿では、Histogram Backprojection を分裂アクティブネットにおける画像の適合性エネルギーとして導入することによって、カラー画像内からターゲットの領域を抽出する手法を提案した。また、Histogram Backprojection の評価関数の値から、しきい値をとることによって、分裂アクティブネットの初期位置を決定し、領域抽出の高速化を図った。更に、サッカー中継シーンのように複数のグループが混在する画像において、特定のグループに属する選手領域の抽出を行なうため、Histogram Backprojectin の評価関数を変更し、選手領域の抽出を行った。

本稿では、Histogram Backprojection の評価関数は、ターゲット画像内の色 (R, G, B 値) だけをキーとしているため、照明などの影響によって、入力画像内のターゲットの色が変化した場合、ターゲットとして見なされないことがある。したがって、色の恒常性を実現し、ロバストな評価関数を定義する必要がある。また、分裂アクティブネットの初期位置や格子点間の間隔についても、更に検討が必要であると考えられる。

参考文献

- 1) M. Kass, A. Witkin and D. Terzopoulos: Snakes: Active Contour Models, International Journal of Computer Vision, Vol. 1, pp. 321-331, (1988).
- 2) 坂上勝彦, 山本和彦: 動的な網のモデル Active Net とその領域抽出への応用, テレビジョン学会誌, Vol. 45, No. 10, pp. 1155-1163, (1991).
- 3) 吉野和芳, 守田 了, 川嶋稔夫, 青木由直: 分裂アクティブネットによる複雑物体の抽出, 1993年電信学春期全大論文集, D-595, (1993).
- 4) Michael J. Swain and Dana H. Ballard: Color Indexing, IJCV, Vol. 7, No. 1, pp. 11-32, (1991).