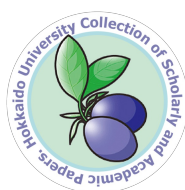




HOKKAIDO UNIVERSITY

Title	アクティブネットを用いたサッカー映像におけるパス可能領域の推定
Author(s)	高橋, 翔; 今, 宏史; 長谷山, 美紀
Citation	電子情報通信学会論文誌 D, J92-D(4), 501-510
Issue Date	2009
Doc URL	https://hdl.handle.net/2115/71961
Rights	copyright©2009 IEICE
Type	journal article
File Information	IEICE_J92-D(4)_501-510.pdf



アクティブネットを用いたサッカー映像におけるパス可能領域の推定

高橋 翔[†] 今 宏史[†] 長谷山美紀[†]

Active Net-Based Non-interception Region Estimation in Soccer Videos

Sho TAKAHASHI[†], Hirofumi KON[†], and Miki HASEYAMA[†]

あらまし 本論文では、チームスポーツ映像からアクティブネットを用いてパス可能領域を推定する手法を提案する。チームスポーツ映像の一つであるサッカー映像の意味内容解析を行うために重要なサッカーの戦術は、選手の移動とボール運びによって表現されるため、ボール運びを実現するパスを分析することは重要である。一般にパスコースはボール保持者と味方チームの選手へとつながる緩やかな曲線で表される。提案手法では、新たなエネルギーの定義とパス可能領域を推定するための画像生成により、アクティブネットを用いて前述の曲線が存在する領域を抽出する。また、パス可能領域は守備の選手から離れるほど、パスが成功する可能性が高いという特徴をもつ。提案手法では、格子点の密度に着眼することで、パスが成功する可能性をパス可能領域の推定と同時に得る。更に、アクティブネットの収束結果は多少の選手位置の誤差を許容するため、選手の動きを用いた従来手法における、選手位置の誤差の影響を受けやすいという問題点を解決することが可能である。したがって、提案手法はカメラワークが存在し、高精度な選手位置の推定が困難であるテレビ映像に対しても、高精度にパス可能領域の推定が可能である。

キーワード スポーツ映像, サッカー映像, 意味解析, パス, アクティブネット

1. ま え が き

近年、デジタル放送の開始やインターネットによる映像配信の増加により、デジタル映像を視聴する機会が増加している。デジタル映像では、映像内容を表現する情報をメタデータとして映像に付加することで、映像検索や映像の編集及び内容の理解を補助することが可能となる。しかし、配信されるデジタル映像のすべてに手動でメタデータを付与することは困難である。

このような背景のもと、自動で映像にメタデータを付与するための研究がドラマやスポーツ、ニュースなど様々な種類のデジタル映像に対して行われてきた[1], [2]。研究が盛んな分野の一つにスポーツ映像解析があり、これまで、画像処理や音響信号処理によりスポーツ映像から重要なシーンを検出する手法が数多く提案されている[3]~[8]。これらの手法により、スポーツにおいて重要なシーンと考えられる映像を抜き

出したハイライト映像を作成することや、特定のシーンを検索することが可能となる。しかし、視聴者が望む重要なシーンは、個人の嗜好に依存して選択されるため、真に望む映像を提供するためには映像の意味内容を解析する必要がある[9]。ここで、スポーツ映像における意味内容の解析とは、試合内容を戦術的かつ技術的な観点から解析することである。このような解析のためには、これまでの画像処理や音響信号処理によって得られる結果に加え、各スポーツの特徴を用いて選手のプレーやチームの戦術にまで踏み込んだ検討を行う必要がある。

現在、サッカー映像における意味内容の解析を実現するために、サッカー映像から自動でカメラパラメータを推定する手法[10], [11]や、選手とボールの位置を検出する手法[12], [13]が提案されている。これらの手法により、フィールドにおける選手のおよその位置を得ることが可能となる。一方、このような手法により算出された選手の位置から戦術を解析する手法が提案されている[14], [15]。例えば、文献[14]では、サッカーに限らず、二つのチームに分かれた選手がフィールド上を自由に移動する競技（以降、チームスポーツ）において、ボロノイ領域を拡張した優勢領域を用いて、

[†] 北海道大学大学院情報科学研究科, 札幌市
Graduate School of Information Science and Technology,
Hokkaido University, Kita-14 Nishi-9, Kita-ku, Sapporo-shi,
060-0814 Japan

チームワークの評価を行う。

ところで、サッカーにおける戦術とは、選手の移動とボールの運び方により、実現される。したがって、後者のボール運びを実現するパスについて分析することもサッカーの戦術を知るために重要と考えられる。パスの分析手法に関する先行研究として文献[15]がある。文献[15]では、文献[14]を発展させ、優勢領域の大きさや形状を分析することにより、パスの成否判定を行う。また、この手法は各選手にかかるプレッシャー強度の定量化も行い、戦術的な観点から試合内容の解析を実現している。しかしながらこの手法では、パスの成否を判定するだけでパスが可能な領域を通過するパスが成功する可能性（以降、パスコースの優劣）を推定することができない。また、事前に選手の動きを正確に推定する必要があり、テレビ放送などのサッカー映像に対して適用することは困難である。

そこで、本論文では、選手の位置関係からパスが成功する可能性が高い領域（以降、パス可能領域）を推定し、その領域を通過するパスコースの優劣を可視化する手法を提案する。一般にパスコースは、ボール保持者から味方チームの選手へとつながる緩やかな曲線で表される。また、パスコースの優劣、すなわち、パスが成功する可能性は、パスコースと守備の選手の距離により決まる。以上の性質に基づき、我々はアクティブネットを用いることでパス可能領域の推定手法を導出する。アクティブネットは内部ひずみエネルギーと適合性エネルギーの和を最小化することにより得られる格子点の密度に基づき、領域を抽出するアルゴリズムである。提案手法は、本論文の問題に適した新たなエネルギー関数の定義とパス可能領域を推定する画像生成により、アクティブネットの最終結果における格子点の密度からパス可能領域を推定する。また、パス可能領域の推定だけでなく、格子点の密度に注目することで、その領域を通過するパスコースの優劣についても推定が可能となる。

2. 前 処 理

パス可能領域は選手の位置関係によって決定されるため、まず、サッカー映像からフィールドでの選手の位置を推定する必要がある。サッカー映像にはテレビ放送のようにカメラがパンやズームなどのカメラワークを行う映像とカメラワークを行わない固定カメラの映像がある。どちらの映像であっても、カメラの設置位置や焦点距離などのカメラパラメータを推定し、そ

の後に映像から選手領域を検出することで、フィールドにおける選手位置を推定することが可能となる。

現在、これらを実現する手法が種々提案されている[10]～[13]。本論文ではパス可能領域の推定を行う手法を提案することを目的としている。このため、提案手法を適用する際の前処理として選手位置の推定を行う必要がある。例えば、選手位置の推定は、文献[10]を用いカメラパラメータを推定し、文献[12]などの選手領域を抽出する手法を用いることで可能となる。しかしながら、撮影環境や選手のオクルージョン等の影響により、カメラパラメータの推定や選手領域の検出の際に誤差が生じ、推定される選手の位置は正確な位置であるとは限らない。そのため、本論文で提案されるパス可能領域の推定法は前処理として行われる選手位置の推定誤差に対してロバストである必要がある。次章で述べる提案手法は、選手の位置関係を用いてパス可能領域を推定するため、厳密に選手位置を推定する必要はない。そのため、選手位置の推定誤差に大きな影響を受けずにパス可能領域の推定が実現される。

3. アクティブネットによるパス可能領域の推定

チームスポーツ映像の試合内容を解析する際に、戦術的な観点からの解析は非常に重要である。チームスポーツの戦術を考えた場合に、パスは重要なプレーである。特にサッカーでは、パスの組合せやパスを行うための選手の移動によって成立する戦術が多い。例えば、ワンツーパスはボール保持者（A 選手）が同チームの選手（B 選手）に対してパスを行い、その直後に前方の守備の選手が存在しない領域に移動し、B 選手が再び A 選手に対してパスを行う戦術である。このようなパスにかかわる戦術を解析するために、パス可能領域の推定が必要となる。加えて、パス可能領域の推定は選手のプレーを評価するためにも利用可能である。サッカーではパス可能領域は複数存在する場合が多く、複数のパス可能領域から選手は一つのパス、あるいは他のプレーを選択する。そのときに、複数のパス可能領域が分かっているならば、選手が選択したパスあるいはプレーの評価が実現される。

そこで、本文では、前章で得られる選手位置からパスコースを推定する手法を提案する。まず、パスがもつ特徴について説明し（3.1）、パス可能領域の推定にアクティブネットを適用するために選手位置を用いて画像を生成する方法について述べる（3.2）。そして、

アクティブネットを実際に用いてパス可能領域を推定する手法を示す (3.3, 3.4)。なお、提案手法で用いるアクティブネットの詳細については、付録に示し、本章で用いるパラメータについては付録に従うものとする。

3.1 選手へのパスとスペースへのパス

パスは一般にボール保持者から選手、あるいは選手が存在しない領域（以降、スペースと呼ぶ）に対して行われる。本文では、以降で、前者を選手へのパス、後者をスペースへのパスと区別する。パスコースはボール保持者とパスの対象、すなわち選手あるいはスペースとを結ぶ曲線であり、その曲線は守備の選手からの距離が遠い領域を通過するという特徴をもつ。ただし、パスコースが曲線となるのは、パスを行った際にボールに回転がかかり、パスコースが直線ではなくわずかに左右に変化をする場合である。したがって、パス可能領域を推定するためには、ボール保持者と選手あるいはスペースを結び、守備の選手からの距離が遠い位置を通過する曲線が存在する領域を抽出すればよい。

上述の特徴をもつ領域を抽出するために提案手法では、アクティブネットを用いる。アクティブネットは、内部ひずみエネルギーと画像の適合性エネルギーの和を最小化する形状へと変形を繰り返すことで領域抽出を可能とする。アクティブネットを適用することで、格子点は抽出対象のエッジに引き寄せられ、領域抽出が実現する。このとき、内部ひずみエネルギーにより格子点が平行性を有するため、抽出された領域のエッジには急激な曲率の変化が生じない。したがって、エネルギーの収束により得られた最終結果における格子点の密度が高い領域は、緩やかな曲率を示し、それをボールの軌跡と仮定し、曲線が存在する領域をパス可能領域として推定することが可能である。

3.2 画像の生成

一般に、アクティブネットを用いて画像内の領域抽出を行う場合、対象領域の特徴からその領域において低いエネルギーとなる画像の適合性エネルギーを定義する。これにより、アクティブネットの格子点は対象領域に引き付けられ、領域抽出が実現される。一方、サッカー映像より得られる画像（以降、サッカー画像）に対して直接アクティブネットを適用しても、パス可能領域を推定することは不可能である。アクティブネットを適用してパス可能領域を推定するためには、パス可能領域に格子点が引き付けられる画像の生成とエネルギーの定義が必要である。そこで、提案手法で

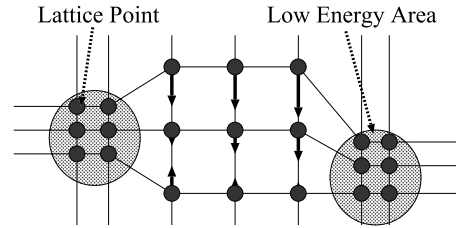


図 1 内部ひずみエネルギーによる平行性の保存
Fig. 1 Smoothing by internal energy.

は、階調値によって表現される式 (A-11) を画像の適合性エネルギーとして、選手の位置に基づいてパス可能領域を推定するための画像を生成する。式 (A-11) を見て分かるように、 $\omega > 0$ とすることでアクティブネットの格子点は画像の階調値が低い領域に引き付けられる。したがって、守備の選手が存在する領域を高い階調値で、パスの対象となる攻撃の選手及びスペースを低い階調値で示す画像にアクティブネットを適用すれば、格子点は守備の選手が存在する領域で密度が低くなり、攻撃の選手及びスペースが存在する領域で密度が高くなる。このとき、アクティブネットが内部ひずみエネルギーによって平行性を保つ力の影響を受けるため、攻撃の選手及びスペースが存在する領域の間でも格子点の密度は高くなる。例えば、図 1 に示す位置に格子点が存在する場合、中央の九つの格子点は矢印で示される位置に移動する。したがって、パスコースが存在する領域で格子点の密度が高くなる。

このように、提案手法では、選手の位置に基づいて画像を生成し、アクティブネットによるパス可能領域の推定を実現する。ただし、パスにはボール保持者から選手へ行うパスの他にスペースへのパスがある。提案手法では、どちらの場合も選手の位置に基づいて画像を生成することで、パス可能領域の推定を実現する。生成される画像は、選手へのパス、スペースへのパスのどちらであっても攻撃の選手がパスを受けることが可能な位置でエネルギーが低くなければならない、すなわちその位置で階調値が低い画像でなければならない。具体的には、選手へのパスの場合は、選手の位置に応じて階調値を定め、スペースへのパスの場合は選手の位置からスペースを推定し、それに従って階調値を定める。以下に、処理の詳細を示す。

3.2.1 選手へのパス

選手へのパスは、ボール保持者から攻撃の選手に対して行われるため、攻撃の選手の位置で低い階調値

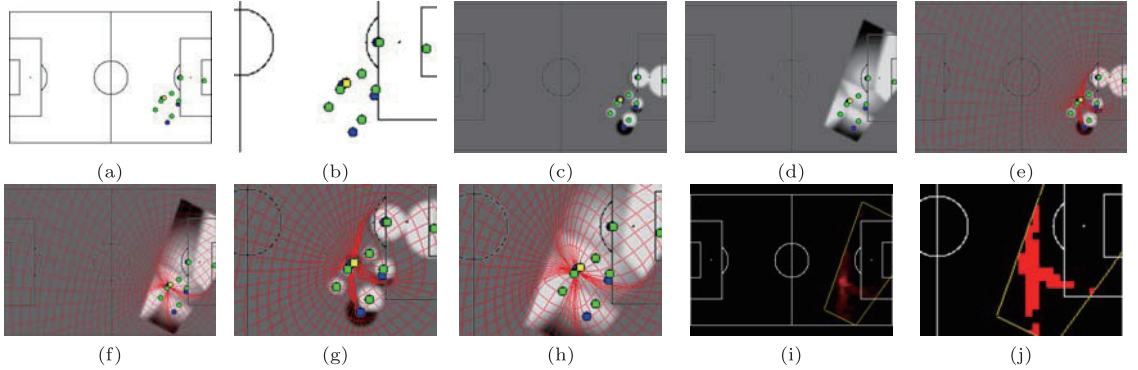


図2 パス可能領域の推定 (a) 入力画像, (b) (a) の拡大画像, (c) 選手へのパスのパス可能領域の推定のための生成画像, (d) スペースへのパス可能領域の推定のための生成画像, (e) (c) に対するアクティブネット適用結果, (f) (d) に対するアクティブネット適用結果, (g) (e) の拡大画像, (h) (f) の拡大画像, (i) パス可能領域の推定結果の可視化 (j) (i) の拡大画像

Fig. 2 Estimation of non-interception region: (a) input image, (b) Enlarged portion image of (a). (c) generated image for estimating non-interception region to players, (d) generated image for estimating non-interception region to spaces, (e) the result of applying Active-Net to (c), (f) the result of applying Active-Net to (d), (g) Enlarged portion image of (e), (h) Enlarged portion image of (f), (i) the result of estimation of non-interception region, (j) Enlarged portion image of (i).

を、守備の選手の位置で高い階調値をもつ画像を生成する。このとき、パスが到達するまでの間に選手が移動することが可能であるため、選手位置に対して階調値によって与えられるエネルギーを円形の形状とする。パスが到達するまでの時間はボール保持者との距離によって異なるため、円の半径 r は、ボール保持者との距離 d_1 [m] を用いて次式とする。

$$r = \frac{d_1 v_h}{v_b} \quad (1)$$

ここで、 v_h 、 v_b はそれぞれ、選手及びパスの速さを表すパラメータであり、これにより、ボールが到達するまでの選手の移動を表現する。このようにして、提案手法では、選手へのパス可能領域を可視化するために、画像を生成する。

例えば、図 2 (a) のサッカー画像から生成される画像を図 2 (c) に示す。ただし、図において、青が攻撃チームの選手、緑が守備チームの選手、黄がボールを表す。また、攻撃の選手の位置に対して最小階調値 l_{min} 、守備の選手の位置に対して最大階調値 l_{max} を割り当て、背景を $\frac{l_{max}+l_{min}}{2}$ とした。更に、サッカーにおいて、守備は攻撃を制限する働きであるため、守備の選手が攻撃の選手を制限するものとして守備選手の円を優先する。

3.2.2 スペースへのパス

スペースへのパスはスペースに対してパスが行われ、そのスペースへ攻撃の選手が移動して、パスを受けることで成立する。そのため、攻撃の選手が到達可能なスペースである度合を定義し、これを選手へのパスにおける選手位置とみなすことで、選手へのパスと同様にパス可能領域の推定が可能となる。提案手法では、まず守備の選手の位置からフィールド上の各点におけるスペースらしさ（以降、スペース度）を定義し、その点にパスよりも早く攻撃の選手が到達する可能性（以降、到達可能性）を推定することで、スペースへのパスの対象となるスペースを推定し、その結果に基づいて画像を生成する。まず、スペースは守備の選手が存在しない領域であるため、各点において最も距離が近い守備の選手の位置との距離 d_2 [m] を用いて、次式でスペース度 s を算出する。

$$s = \lambda d_2 \quad (2)$$

ここで、 λ は、距離に応じてスペース度を制御するパラメータである。これにより、守備の選手との距離が離れている点ほどスペース度が高くなる。次に、到達可能性を算出する。到達可能性は守備の選手及び攻撃の選手がスペースに到達するまでの時刻で決まる。到達するまでの時刻は距離で決まるため、提案手法では、

スペース度が定義された各点と攻撃の選手及び守備の選手の距離 d_a [m], d_d [m] を用いて次式により到達可能性を定義する．

$$p = \begin{cases} \gamma(d_d - d_a) + 0.5 & (0 \leq p \leq 1) \\ 1 & (p > 1) \\ 0 & (p < 0) \end{cases} \quad (3)$$

上式により，攻撃の選手が守備の選手よりも早く到達可能な場合に $p > 0.5$ となり，それ以外で $p < 0.5$ となる．到達可能性 p とスペース度 s の積を画像の最大階調値から減算した値を階調値とした画像を生成することで，スペースへのパスが成功する可能性が高いスペースの領域で低い階調値となり，スペースへのパスのパス可能領域を推定することが可能となる．

以上により生成される画像を図 2(d) に示す．図 2(d) において，フィールド中央よりもボール保持選手側のタッチラインに近い領域（以降，ニア）に存在するスペースとフィールド中央よりもボール保持選手と反対側のタッチラインに近い領域（以降，ファー）に存在するスペースが低い階調値で示されており，パスが成功する可能性が高いスペースの階調値が低く表示されることが確認できる．ここで，図 2(d) の台形領域はサッカーフィールドの撮像領域である．

3.3 アクティブネット

本節では，3.2.1 及び 3.2.2 で生成した画像に適用するパス可能領域を推定するためのアクティブネットの設計法について説明する．アクティブネットの性質は，抽出対象となる物体の性質から初期形状，内部ひずみエネルギー，画像の適合性エネルギーを定義することで決まる．まず，アクティブネットの初期形状について述べる．初期形状は抽出対象であるパス可能領域となる領域の形状によって決まる．パスはボール保持者から同チームの選手に対して行われるため，パス可能領域はボール保持者から放射状に存在する．そのため，アクティブネットの初期形状は放射状が最適である．初期形状を放射状にすることで，アクティブネットの形状がパス可能領域に対して平行となる．これにより，ボール保持者とパスの対象の間，すなわちパス可能領域に格子点が引き付けられる．次に，内部ひずみエネルギーを定義する．内部ひずみエネルギーは式 (A-5) により，平行性を保つ力と収縮力をアクティブネットに与える．平行性を保つ力は，パスコースが急激に変化しない特徴を満たすパス可能領域に格子点が引き付けられる力となるため重要である．一方，

収縮性により，アクティブネットが画像を包含しない形状に変化することで，パス可能領域の抽出精度が劣化するため，収縮性は不要である．そのため，式 (A-5) の α , β は，それぞれ 0.0, 1.0 とする．また，アクティブネットが変形後も画像全体を包含することを可能とするため，最外殻格子を固定する．最後に，画像の適合性エネルギーについては，前節で示したとおりである．

このようにして，定義されたアクティブネットを前節で生成した画像に適用すると，図 2(e), (f) となる．図 2(e), (f) とその拡大図，図 2(g), (h) において，パスが可能であると考えられるニアとファーに存在するスペースや，ファーに存在する選手，右サイドの選手へのパス可能領域にアクティブネットの格子点が集中していることが確認できる．

3.4 パス可能領域の推定

本節では，アクティブネットの格子点の密度に着目し，パス可能領域を通過するパスが成功する可能性（以降，パスコース可能性と呼ぶ）を推定する．提案手法ではパスコース可能性をフィールドの小領域ごとに算出する．そのため，まず，サッカーフィールドを $n \times m$ 等分し，小領域 D_{ij} ($i = 1, 2, \dots, n, j = 1, 2, \dots, m$) を得る．このとき，分割数を細かくすれば，パスのずれを許容する幅が小さくなり，大きくすれば，パスのずれを許容する幅が大きくなる．パス可能領域は，アクティブネットにおいてボール保持者から放射状に存在する曲線によって表現され，その曲線の単位面積当りの本数に従って，パスコース可能性が決まる．そのため，分割された領域に対してパスコース可能性を推定するためには，各領域を通過する曲線の本数を調べればよい．そこで，各領域 D_{ij} に対してパスコース可能性 F_{ij} を以下で定義する．

$$F_{ij} = \sum_{k=0}^{K-1} T_{ij}(N_k) \quad (4)$$

ここで， N_k は，3.3 で設定したアクティブネットの初期形状において，中心から放射状に延びる各直線 ($k = 0, \dots, K-1$: K は直線の総数) に対応している．また， $T_{ij}(N_k)$ は各直線がアクティブネットの収束後に，領域 D_{ij} を通過する場合に 1 となりそれ以外は 0 となる関数である．したがって，式 (4) で定義される F_{ij} は，領域 D_{ij} を通過する曲線の総数を示し，その値が大きい場合，パスコース可能性が高いと判断できる．得られたパスコース可能性 F_{ij} を次式で表示

可能な階調値 R_{ij} に割り当てることで図 2 (i) となり、フィールド上に、各パス可能領域を通過するパスの成功率を可視化することができる。

$$R_{ij} = R_{max} \frac{F_{ij}}{F_{max}} \quad (5)$$

ここで、 R_{ij} は割り当てる階調値、 R_{max} は画像の赤色の最大階調値、 F_{max} は全小領域における F_{ij} の最大値である。また、図 2 (i) の台形領域はサッカーフィールドの撮像領域である。更に、図 2 (j) にその台形領域を含む部分の拡大図を示す。拡大図では図 2 (i) の R_{ij} に対し、しきい値処理を施すことによって 2 値化を行い、得られたパス可能性が高い領域のみが示されている。この領域は、ボール保持選手の位置を中心とする放射状に延びた形状を示しており、放射状に延びた領域の各々が、パス可能性が高い上位のパスコースに対応する。

4. 実 験

本章では、実際にテレビで放映されたサッカー映像に対して提案手法を適用した結果を示し、その有効性を確認する。実験対象として、実際に行われた試合のサッカー映像においてパスが行われた場面を用いた。パスが行われる直前のフレームに対して提案手法を適用した結果を図 3～図 6 に示す。各図の説明は次のとおりである。処理対象とした映像においてパスが行われる直前の選手の位置を図 3 に示す。ここで、図 3 はボール保持選手から A4 の選手に対してパスが行われる場面である。図 3 に示す選手位置から、**3.2.1** 及び **3.2.2** に示した手法により選手へのパス及びスペースへのパスのパス可能領域を推定するための画像が生成される。生成された各々の画像を図 4 (a), (b) に示す。これらの画像に対して、初期形状を図 4 (c), (d) とし

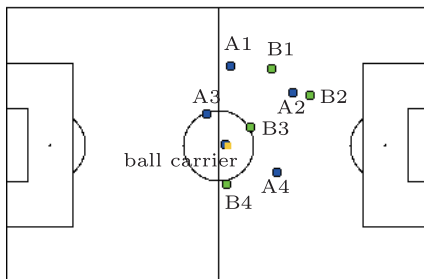


図 3 入力画像の選手位置

Fig. 3 Player positions of the input image.

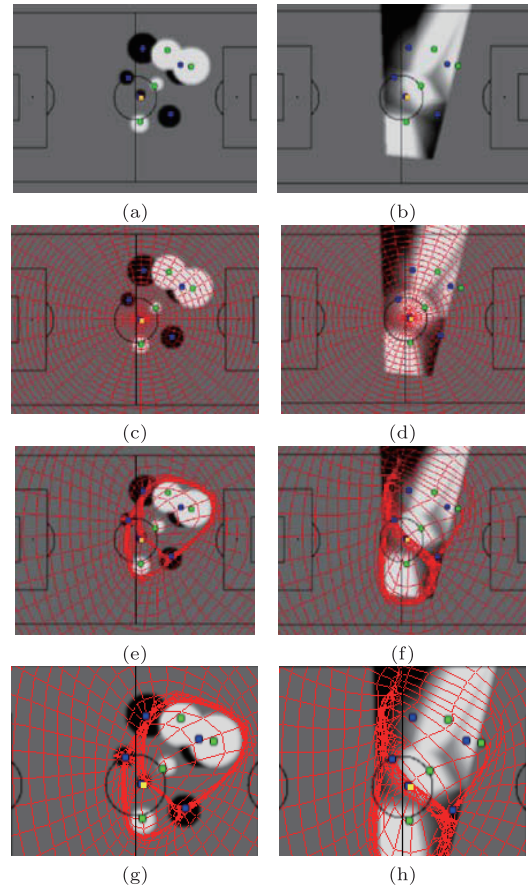


図 4 アクティブネット適用：(a) (b) 生成画像，(c) (d) 初期形状，(e) (f) 収束後のアクティブネット，(g) (h) (e) の拡大図，(h) (f) の拡大図

Fig. 4 Applying Active Net: (a) (b) generated image, (c) (d) Active Net before deformation, (e) (f) Active Net after convergence, (g) Enlarged portion image of (e), (h) Enlarged portion image of (f).

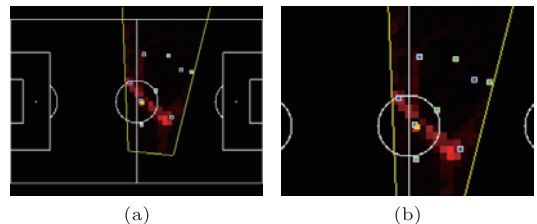


図 5 パス可能領域 (a) パス可能領域の推定結果，(b) (a) の拡大図

Fig. 5 The non-interception region: (a) The result of estimation of non-interception region, (b) Enlarged portion image of (a).

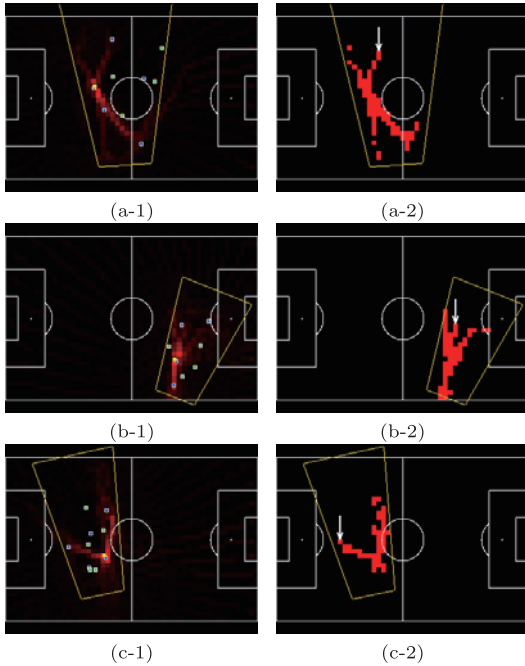


図 6 パス可能領域の推定結果 (a-1) 左へのパス, (a-2) (a-1) にしきい値処理を施した結果, (b-1) 斜め前へのパス, (b-2) (b-1) にしきい値処理を施した結果, (c-1) 前へ向かうパス, (c-2) (c-1) にしきい値処理を施した結果

Fig. 6 The results of pass source estimation: (a-1) left side pass, (a-2) The result of thresholding to (a-1), (b-1) diagonal pass, (b-2) The result of thresholding to (b-1), (c-1) forward pass, (c-2) The result of thresholding to (c-1).

て、3.3で提案したアクティブネットを適用する。ただし、ボール保持者はボールの位置に最も近い選手とした。また、アクティブネットの格子点数を 52×36 とし、他のパラメータは $\lambda = 15$, $\gamma = 0.2$, $\omega = 5.0$ に設定した。更に、 v_h と v_b はそれぞれ 24 km/h, 30 km/h であると仮定する。3.3で定義した内部ひずみエネルギー及び画像の適合性エネルギーの和を最小化することで、図 4(e), (f), (g), (h) の結果を得る。このとき、反復回数は十分に収束させるために 3000 回とした。最後にアクティブネットの格子点からパス可能領域を推定した結果を図 5 に示す。この図 5 が提案手法による最終出力となる。以下で実験結果について考察する。

図 3 を見て分かるように、実験に用いた場面では、ボール保持者を除いた攻撃の選手が 4 人 (図 3 A1, A2, A3, A4) と守備の選手 (図 3 B1, B2, B3, B4)

が 4 人存在している。ここで、ボール保持者からのパスを考えると選手 A2 とボール保持者との間には守備の選手 B3 が存在し、選手 A2 へのパスは成功する可能性が低いことは明らかである。一方、選手 A1, A3, A4 とボール保持者の間には守備の選手が存在していないため、パスが成功する可能性が高い。図 5 に示されるパス可能領域の推定結果では、図 3 に示すボール保持者と選手 A4 の間には赤成分の階調値が高い領域が存在し、すなわち、これが成功する可能性が高いパス可能領域を表している。一方、図 3 に示すボール保持選手と選手 A2 の間に対応する図 5 の部分には赤成分の階調値が低い領域が存在し、パスが成功する可能性が低く示されている。今、図 3 においてボール保持選手から選手 A4, A3, A1 の各々が存在する三つの異なるパス可能領域に注目する。A4 が存在するパス可能領域と A3 が存在するパス可能領域は A1 が存在するパス可能領域に比べ、ボール保持者からの距離が近いことや守備選手との距離が離れていることから、パスが成功する可能性が高いことが認識できる。この事実の上に図 5 の結果を考察すると A1 が存在するパス可能領域が他の二つのパス可能領域に比べ、赤色成分の強度が低く示されていることが分かる。結果として、提案手法により、正しくパス可能領域を示していることが確認できた。以上、図 3～図 5 を用いて、提案手法の有効性を確認したが、一つの場面の例では十分に確認できないことから、図 6 (a-1)～(c-1) に、異なる三つの場面に対して提案手法を適用した結果を示す。併せて図 6 (a-2)～(c-2) に R_{ij} に 2 値化処理を施すことによって得られたパスの可能性が高い領域を示す。更に、図 6 (a-2)～(c-2) には実際にボール保持者からパスが行われた選手の位置を白矢印で示した。すべての場面において、実際に行われたパスが、提案手法により得られた図 6 (a-2)～(c-2) における赤色領域に含まれていることが確認できる。

次に、提案手法により推定されたパス可能領域の精度について更に検証し、その結果を表 1 に示す。表 1 では、実際に行われた試合 1 試合のサッカー映像においてパスが行われた全場面に提案手法を適用し、全映像時間を 4 分割し、各時間における正解率とパス相手が撮像領域内に存在する場合の正解率を併せて示している。なお、正解率は、正解数と全パス数との比で定義し、実際に行われたパスが提案手法により推定されたパス領域に存在した際に正解とした。ただし、推定されたパス領域とは R_{ij} にしきい値処理を施し、パス

表 1 提案手法におけるパス領域推定結果と実際に行われたパスとの比較

Table 1 The comparison of the result of the non-interception region estimation with the pass.

試合経過	1st	2nd	3rd	4th
実際のパスが推定されたパス領域上に存在した数	38	28	37	22
実際に行われたパスの全数	42	35	44	26
正解率	90%	80%	84%	85%
パス相手が撮像領域内に存在する場合の正解率	90%	82%	88%	85%

の可能性が高い上位のパスコースのみを含む領域に限定することで得られる。具体的には、 R_{ij} にしきい値処理を施した後に細線化の処理を施し、1 試合分の映像全体において、得られた曲線の本数が最大で 5 本となるようにしきい値を設定した。表 1 より提案手法によるパス領域推定の正解率は 80%以上であることが確認できる。20%の失敗は、実験対象の映像においてロビングなどのロングパスが行われており、パスを受ける選手が撮像領域内に存在しない場合や守備選手の移動方向とは逆の領域に向かうパスが行われている場合であった。前者の問題については、カメラワーク推定や選手の位置や動きの推定手法を用い、撮像領域外の選手位置を推定するなどして解決する必要がある。なお、フィールド全体が見渡せるような広角映像により試合が撮影された場合にはこのような問題は生じない。また、後者の問題については、提案手法がサッカー映像の一つのフレームのみを用いパス可能領域を算出する手法となっているため、選手の移動方向が考慮されていないことが原因である。このため、選手の動きに関する特徴を用いた、より高精度な手法について検討する必要がある。

5. む す び

本論文では、事前に得られた選手位置からパス可能領域を推定するための画像を生成し、これにアクティブネットを適用することで、パス可能領域とその優劣を推定する手法を提案した。また、実験により、提案手法が選手の位置関係のみを用いるため、テレビ放送の映像から得たサッカー画像においてもパス可能領域の推定が可能であることを示した。提案手法を用いてパス可能領域を推定すれば、戦術を解析する研究への発展が期待される。また、今後の発展によって、視聴者が効果的なパスの場面や、守備が効果的に行われている場面などを自動選択して視聴することが可能となる。しかしながら、提案手法は選手の頭上を越えるパスを考慮しておらず、今後は、このようなパスについても検討を行う必要がある。ところで、サッカーでは

ゴールすることが目的であり、ゴールへ向かうパスの重要性は高くなる。よって、今後は、選手とゴールとの位置関係を考慮した推定手法についても検討を行っていく予定である。これにより、選手が行うパスの予測や実際に行ったパスの評価が可能になると考えられる。

謝辞 本研究の一部は、日本学術振興会科学研究費補助金基盤研究 (c) (課題番号: 18500068) により行われた。ここに記して謝意を表す。

文 献

- [1] 長坂晃朗, 田中 譲, “カラービデオ映像における自動索引付け法と物体探索法,” 情処学論, vol.33, no.4, pp.543–550, April 1992.
- [2] 井手一郎, 山本晃司, 浜田玲子, 田中英彦, “ショット分類に基づく映像への自動的索引付け手法,” 信学論 (D-II), vol.J82-D-II, no.10, pp.1543–1551, Oct. 1999.
- [3] 牛尼剛聡, 広部一弥, 渡邊豊英, “利用者の視点に基づくシーン検索のためのイベント-アクティビティモデル,” 信学論 (D-I), vol.J82-D-I, no.1, pp.256–267, Jan. 1999.
- [4] 新田直子, 馬場口登, “放送型スポーツ映像の意味内容獲得のためのストーリー分割法,” 信学論 (D-II), vol.J86-D-II, no.8, pp.1222–1233, Aug. 2003.
- [5] 宮森 恒, “映像と音響情報の協調による内容検索のためのテニス動作自動注釈付け,” 信学論 (D-II), vol.J86-D-II, no.4, pp.511–524, April 2003.
- [6] A. Ekin, A.M. Tekalp, and R. Mehrotra, “Automatic soccer video analysis and summarization,” IEEE Trans. Image Process., vol.12, no.7, pp.786–807, July 2003.
- [7] B. Li, H. Pan, and I. Sezan, “A general framework for sports video summarization with its application to soccer,” IEEE Int. Conf. Acoustics, Speech and Signal Processing, vol.3, pp.169–172, 2003.
- [8] R. Leonardi, P. Migliorati, and M. Prandini, “Semantic indexing of soccer audio-visual sequences: A multimodal approach based on controlled markov chains,” IEEE Trans. Circuits Syst. Video Technol., vol.14, no.5, pp.634–643, May 2004.
- [9] 長谷山美紀, “ユーザが望む映像を提供するために—画像認識とクラスタリングそして意味理解への発展,” 映像学技報, vol.29, no.47, pp.49–52, 2005.
- [10] T. Watanabe, M. Haseyama, and H. Kitajima, “A soccer field tracking method with wire frame model from TV Images,” IEEE Int. Conf. Image Processing,

- pp.1633–1636, 2004.
- [11] D. Farin, S. Krabbe, P.H.N. de With, and W. Effelsberg, “Robust camera calibration for sport videos using court models,” Proc. SPIE Storage and Retrieval Methods and Applications for Multimedia 2004, vol.5307, pp.80–91, Jan. 2004.
- [12] Y. Ohno, J. Miura, I. Ide, S. Sakai, and H. Tanaka, “Tracking players and estimation of the 3D position of a ball in soccer games,” Proc. 15th Int. Conf. on Pattern Recognition, vol.1, pp.145–148, Sept. 2000.
- [13] M. Xu, J. Orwell, and G. Jones, “Tracking football players with multiple cameras,” Proc. ICIP 2004, pp.2909–2912, Oct. 2004.
- [14] 瀧 剛志, 長谷川純一, “チームスポーツにおける集団行動解析のための特徴量とその応用,” 信学論 (D-II), vol.J81-D-II, no.8, pp.1802–1811, Aug. 1998.
- [15] 藤村 光, 杉原厚吉, “優勢領域に基づいたスポーツチームワークの定量的評価,” 信学論 (D-II), vol.J87-D-II, no.3, pp.818–828, March 2004.

付 録

アクティブネット

本章では、提案手法で用いるアクティブネットについて説明する。アクティブネットはエネルギー最小化原理に基づいた、動的な網のモデルであり、網に対して定義されたエネルギーが極小状態になるように、変形を繰り返すことにより、領域の抽出を行う。

網のモデルは二次元の点列のパラメータ表現として表し、図 A・1 に示すとおりに網は格子点、

$$\mathbf{v}(p, q) = (x(p, q), y(p, q)) \quad (\text{A} \cdot 1)$$

からなり、各々の格子点は隣り合う 4 近傍の点 $\mathbf{v}(p, q-l)$, $\mathbf{v}(p-k, q)$, $\mathbf{v}(p+k, q)$, $\mathbf{v}(p, q+l)$ と結合されて網を形成している。ここで、 p, q は網のパラメータ表現のための二つのパラメータであり、 $0 \leq p \leq 1$, $0 \leq q \leq 1$ である。また、 x, y はそれぞれ画像の水平、垂直方向の座標を表し、 $k = 1/(x \text{ 方向の格子点数}-1)$, $l = 1/(y \text{ 方向の格子点数}-1)$ である。

この網は、網を収縮させ形状を滑らかに保とうとする力に対応する網自身の内部ひずみエネルギー $E_{internal}^*$ と網を画像内の特徴的な領域に引き付ける力に対応する網と画像の適合性エネルギー E_{image}^* を備える。したがって、網のエネルギー汎関数 E_{net}^* は両者の和として以下の式で定義される。

$$E_{net}^* = E_{internal}^* + E_{image}^* \quad (\text{A} \cdot 2)$$

ただし、

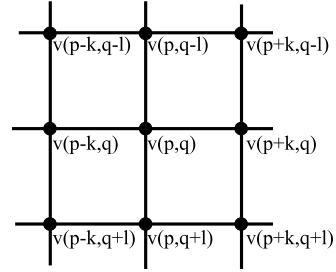


図 A・1 格子点による網の形成

Fig. A・1 Generate of the net by lattice points.

$$E_{internal}^* = \int_0^1 \int_0^1 E_{internal}(\mathbf{v}(p, q)) dp dq \quad (\text{A} \cdot 3)$$

$$E_{image}^* = \int_0^1 \int_0^1 E_{image}(\mathbf{v}(p, q)) dp dq \quad (\text{A} \cdot 4)$$

これにより、内部ひずみエネルギー $E_{internal}^*$ と画像の適合性エネルギー E_{image}^* を適切に定義することで、安定な領域抽出を可能とする。内部ひずみエネルギー及び画像の適合性エネルギーの定義を以下で説明する。

まず、内部ひずみエネルギーは網を収縮させ、かつ滑らかに保とうとするエネルギーであるため、次式で定義する。

$$\begin{aligned} E_{internal}(\mathbf{v}(p, q)) &= \{\alpha(|\mathbf{v}_p|^2 + |\mathbf{v}_q|^2) + \beta(|\mathbf{v}_{pp}|^2 + |\mathbf{v}_{pq}|^2 \\ &\quad + |\mathbf{v}_{qq}|^2)\} \end{aligned} \quad (\text{A} \cdot 5)$$

ただし、

$$\mathbf{v}_p = \frac{\partial \mathbf{v}}{\partial p} = \left(\frac{\partial x}{\partial p}, \frac{\partial y}{\partial p} \right) \quad (\text{A} \cdot 6)$$

$$\mathbf{v}_q = \frac{\partial \mathbf{v}}{\partial q} = \left(\frac{\partial x}{\partial q}, \frac{\partial y}{\partial q} \right) \quad (\text{A} \cdot 7)$$

$$\mathbf{v}_{pp} = \frac{\partial^2 \mathbf{v}}{\partial p^2} = \left(\frac{\partial^2 x}{\partial p^2}, \frac{\partial^2 y}{\partial p^2} \right) \quad (\text{A} \cdot 8)$$

$$\mathbf{v}_{pq} = \frac{\partial^2 \mathbf{v}}{\partial p \partial q} = \left(\frac{\partial^2 x}{\partial p \partial q}, \frac{\partial^2 y}{\partial p \partial q} \right) \quad (\text{A} \cdot 9)$$

$$\mathbf{v}_{qq} = \frac{\partial^2 \mathbf{v}}{\partial q^2} = \left(\frac{\partial^2 x}{\partial q^2}, \frac{\partial^2 y}{\partial q^2} \right) \quad (\text{A} \cdot 10)$$

式 (A・5) は二つの項から定義されており、第 1 項は網を収縮する力を表し、第 2 項は網の平行性と直線性を保つ力を表している。また、 α, β はそれぞれ第 1 項と第 2 項とを考慮する割合を決める重み係数である。

次に、画像の適合性エネルギーを定義する。網は適

合性エネルギーを最小化するための変形を繰り返すことで、画像内の特徴的な領域に向かって収縮する。そのため、目的とする対象領域の特徴に依存すべきであり、注目する領域を顕著に表す画像の適合性エネルギーを定義することで、領域を抽出することが可能となる。ここでは、最も簡単な例として、次式に示す画像の適合性エネルギー E_{image}^* を示す。

$$E_{image}(\mathbf{v}(p, q)) = \omega I(p, q) \quad (\text{A} \cdot 11)$$

ここで、 $I(p, q)$ は、格子点 $\mathbf{v} (= (p, q))$ における画像の階調値を表す。また、 ω は重み係数であり、 $\omega > 0$ の場合、画像の適合性エネルギー E_{image}^* は網を階調値の低い方へ動かす力となり、その大きさにより階調値の影響力を制御することができる。

以上で定義されたエネルギーを用いて、網のエネルギーが最小となるように変形を繰り返すことで、アクティブネットを用いた画像内の領域抽出が実現される。

(平成 20 年 7 月 10 日受付, 11 月 17 日再受付)



高橋 翔 (学生員)

2008 北大・工・情報工学卒。現在、同大学院情報科学研究科修士課程在学中。映像意味理解に関する研究に従事。



今 宏史 (正員)

2005 北大・工・情報工学卒。2007 同大学院情報科学研究科修士課程了。現在、KDDI (株) に勤務。



長谷山美紀 (正員)

1986 北大・工・電子卒。1988 同大学院工学研究科修士課程了。1989 同大学電子科学研究科助手。1994 同大学工学研究科助教授。2006 同大学院情報科学研究科教授、現在に至る。ディジタル信号処理及び画像処理の研究に従事。