

状況理解と映像評価を考慮した講義の知的自動撮影

村上 昌史 大西 正輝 福永 邦雄
大阪府立大学 大学院工学研究科

本稿では、講義を人手を介さずに自動的に撮影する手法として、カメラマンのようにシーンの状況を理解して撮影領域を決定し、スイッチャーのように複数の映像の中から最も効果的にシーン状況を表現する映像を選択する知的自動撮影手法を提案する。まず、講義の状況を理解するために、固定カメラによって撮影した講義映像から講義者と黒板の板書に関する情報を抽出し、それらの情報を用いて講義者の行動認識を行う。次に、講義者の行動に基づいて各カメラにおいて撮影領域を決定し、複数のカメラ位置から映像を取得する。最後に、得られた複数の映像をそれぞれ評価することにより、現在の講義状況を最も効果的に表現している映像を選択する。模擬的な講義の自動撮影を行い、本手法の有効性を確認した。

Computer-controlled Camera Work at Lecture Scene Considering Situation Understanding and Evaluation of Video Images

Masashi Murakami Masaki Onishi Kunio Fukunaga
Graduate School of Engineering, Osaka Prefecture University

In this paper, we propose a computer-controlled intelligent camera work system which shoots object scenes as cameramen and select a best image from plural video images as a switcher. We apply this system to a shooting of a lecture scene. In the first, to understand the situation of lecture, this system estimates actions of a lecture person on the basis of extracted each movement of a lecture person and written characters on a blackboard. In the next, each camera fixes a shooting region based on the actions of a lecture person. In the last, this system selects the best image from plural images by evaluating each image. And we tried an experiment of shooting of an imitational lecture and have confirmed the effectiveness of our approach.

1 はじめに

デジタル多チャンネル時代を迎え、映像生成の省力化の要求が高まってきており、カメラマンが撮影するような自然な映像を自動的に生成する研究が行われている。これらの技術は、スタジオ番組やスポーツ番組などの撮影だけではなく、災害地のような危険な場所やカメラマンが撮影できないよ

うな位置からの映像生成が可能になるなど応用範囲は広い。そこで、プロのカメラマンのような撮影知識を計算機に組み込み、自動的に最適な映像を取得することを、知的自動撮影と呼ぶことにし、本稿ではその手法について検討する。

知的自動撮影システムは、シナリオの有無により設計法が大きく異なるものと考えられる。ドラマや能などのように、あらかじめシナリオが設定さ

れているシーンを撮影対象とする場合には、前もって計算機にシナリオを与えておくことができるため、シナリオに同期したカメラワークを実現する手法が考えられている [1, 2, 3]。一方、スポーツや講義などのように決まったシナリオがないシーンを撮影対象とする場合には、撮影領域を決定する処理に加えてシナリオに相当するイベントをシーン中から自動的に抽出し、この結果をもとにカメラワークを決める必要がある。

これまでも、抽出したイベントをトリガーとして撮影領域を決定する自動撮影システムが数多く提案されている。スポーツをカメラマンなど人手を介さずに自動的に撮影するシステムとしては、選手の動きをトリガーとしてアメリカンフットボールを撮影する手法 [4] や体操の床演技を撮影する手法 [5]、サッカーを撮影する際にボールの軌跡から視点を決定する手法 [6] などが報告されている。また、講義を自動的に撮影するシステムとしては、講義者の位置や動きをトリガーとして撮影する手法 [7, 8, 9]、ならびに指示や例示動作をトリガーとして手元作業を中心としたプレゼンテーションを撮影する手法 [10] などが報告されている。知的自動撮影システムでは各カメラにおける撮影領域の決定法と複数台のカメラから得られる映像を切り替えるスイッチング法が問題となるが、これまでに提案されてきた手法の多くは、撮影領域の決定法かスイッチング法のどちらか一方に焦点を当てた研究が多く、両方を同時に扱った研究は少ない。

通常、シナリオのないシーンを撮影対象とした番組制作においては、複数のカメラマンがシーンの状況を理解して各々の位置から最も適したアングルで撮影を行い、スイッチャーが得られた複数の映像の中からそのシーンを表現するのに最も適した映像を選択することが多い。本稿では、撮影対象として講義シーンを取り上げ、シーン中の状況理解により決定した領域を撮影した後、得られた複数の映像の中からより良い映像を選択する知的自動撮影システムを提案する。また、実際に知的自動撮影システムを構成し、実験により本手法の有効性を確認する。

2 講義映像におけるイベント抽出

スポーツや講義などを撮影する場合には、シーン中に起こっている状況から、撮影領域を決定す

るトリガーを抽出する必要がある。本稿では撮影領域を決定するためのトリガーをイベントと呼ぶ。講義の撮影を想定したカメラワークを設定する上では、講義者の行動が重要なトリガーとなる場合が多い。ここでは、固定カメラで撮影した講義映像から講義者や黒板の板書についての情報を抽出して、イベントとして設定した講義者の行動を認識する。

2.1 講義者と板書の分離

講義映像に現れる対象物として、講義者と板書文字の2つが考えられる。講義者や板書文字に関する情報を抽出するためには、講義映像からこれら2つの対象物を分離して抽出する必要がある。本研究では講義者は動いており板書文字は静止していることに注目し、時空間画像の断面図にエッジ処理を施すことで講義者と板書文字を分離して抽出する [11]。そして時空間画像の断面図から抽出した動物体によるエッジを動エッジ、静止する物体によるエッジを静止エッジと呼ぶ。

図1(a)の入力画像から、動エッジ(黒色)と静止エッジ(灰色)を抽出した例を図1(b)に示す。動エッジを生成する講義者と静止エッジを生成する板書文字が分離できていることが確認できる。

2.2 講義者情報の抽出

講義映像において動エッジを生成するものは講義者であると考えられる。そこで、入力画像から得られた動エッジを用いて講義者に関する情報を抽出する。イベントとして設定した講義者の行動認識を行うための情報として、講義者の位置や顔の向き、さらには手の位置を抽出する。

2.2.1 頭部の位置推定

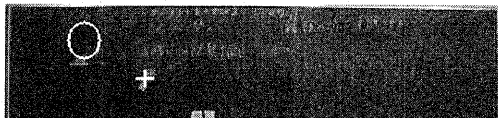
動エッジを用いて講義者の頭部の位置を推定する。まず、得られたすべての動エッジ点の重心から講義者頭部の概略的な位置を調べる。次に、頭部は楕円形であると仮定して、動エッジの重心付近で楕円に近い分布をしている動エッジを探索し、得られた楕円領域を講義者の頭部とする。ただし、ここで用いる楕円の横径と縦径の比は1:1.2とした。一方、講義者が動かなかった場合には動エッジがほとんど現れないため、動エッジ数がしきい値以下



(a) 入力画像



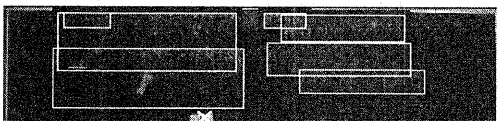
(b) 動エッジ (黒) と静止エッジ (灰) 画像



(c) 講義者情報の抽出



(d) 肌の抽出



(e) 板書情報の抽出

図 1: 状況理解結果

になれば講義者は動いていないものと考え、前フレームでの講義者の頭部の位置推定結果を用いる。

以上の処理によって講義者の頭部領域を抽出した例を図 1 (c) に示す。図中の楕円が抽出された頭部を表しており、楕円の中心を講義者の位置とする。

2.2.2 顔の向き推定

入力画像の各画素の色情報を調べることで人物の肌を抽出し、頭部抽出結果を用いて講義者の顔の向きを推定する。

人物の肌色を抽出するための色情報として明るさの変化の影響を比較的受けにくい、CIE1976UCS 表色系を用いる。人物の肌を含む複数の画像から顔や手などの肌領域の画素を N 個抽出し、それら

の画素の画素値を CIE1976UCS 色度図上に投影した $C_i = (u_i, v_i)$ ($i = 1, 2, \dots, N$) に対し、人物の肌色の平均値 μ_S および共分散行列 Σ_S を次式で求める。

$$\mu_S = \frac{1}{N} \sum_{i=1}^N C_i \quad (1)$$

$$\Sigma_S = \frac{1}{N-1} \sum_{i=1}^N (C_i - \mu_S)^T (C_i - \mu_S) \quad (2)$$

上式で求めた平均値 μ_S と共分散行列 Σ_S から、入力画像中の各画素 x に対する画素値 C_x が肌色にどれくらい近いかを表す値 $P_S(C_x)$ を次式で求める。

$$P_S(C_x) = \frac{1}{2\pi |\Sigma_S|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (C_x - \mu_S) \Sigma_S^{-1} (C_x - \mu_S)^T \right\} \quad (3)$$

すべての画素において (3) 式から求めた値をしきい値処理により 2 値化して肌色素を抽出する。抽出した肌色素を図 1 (d) に示す。黒い画素が肌色領域を示している。

次に、講義者の頭部 (2.2.1 で求めた楕円) 領域に含まれている肌色素の分布から顔の向きを推定する。ここでは概略的な向きを調べることを目的としているため、楕円を左側と右側の 2 つに分けてそれぞれの領域に含まれる肌色素の数をかぞえ、しきい値処理によりそれらの大小関係を調べることで、顔の向きを「正面」「右」「左」「後ろ」の 4 つに分類する。例えば、肌色素が楕円の左半分によく含まれているが、右半分にはあまり含まれていない場合には、講義者の顔の向きは「左」とであると推定することにする。

2.2.3 手の位置推定

次に、2.2.1 で得られた頭部領域と 2.2.2 で得られた肌色抽出結果を用いて、講義者の手の位置を推定する。

講義映像に現れる肌色領域は、講義者の顔領域と手領域であると考えられるため、抽出した肌色素のうち、頭部領域に含まれない肌色素が手領域を形成していると考えられる。そこで、頭部領域を除くすべての肌色素を位置情報によってクラスタリングした後に、各クラスタの重心を求

表 1: 行動推定規則

イベント (講義者の行動)	条件				
	POSI	FACE	HAND	CHAR	ERAS
1. “板書している”	停止	後ろ	動作	増加	*
2. “黒板を消している”	停止	後ろ	動作	減少	なし
3. “板書について説明している”	停止	正面以外	動作	変化なし	*
4. “学生に対して説明している”	停止	正面	動作	変化なし	*
5. “左に移動している”	左	*	*	変化なし	*
6. “右に移動している”	右	*	*	変化なし	*

(* : don't care)

めれば手の位置を推定することができる。ここでは K -平均アルゴリズムを用いて肌色画素のクラスタリングを行う。求めるクラスタ数は両手を想定した 2 としてクラスタリングを行い、最終的に決定した 2 つのクラスタ中心の座標間距離がしきい値未満であった場合には手領域が 1 つしか映っていないものとして、2 つのクラスタの平均座標を手の位置とする。また 2 つのクラスタ中心間の距離がしきい値以上であった場合には、手領域が 2 つ映っているものと考え、それぞれのクラスタ中心の座標を手の位置とする。

手の抽出結果を図 1 (c) に +印で示す。入力画像中に映し出されている片手のみが抽出できている。

2.3 板書情報の抽出

2.3.1 板書文字量の推定

設定したイベントを抽出するためには、講義者に関する情報だけでは十分とはいえない。例えば、“板書している”という行動と“黒板を消している”という行動は、どちらも黒板の方を向いて手を動作させる行動であるため、講義者に関する情報のみからこれらの行動を区別するのは難しい。そこで、“板書している”ときには板書文字量が増加し、“黒板を消している”ときには板書文字量が減少することに注目して、板書文字量の増減から、これらを区別する。講義シーンを考えた場合には、板書文字は黒板中の静止エッジで構成されていると考えられるため静止エッジのドット数を調べ、その数を板書文字量とする。

2.3.2 黒板消しの位置推定

“黒板を消している”という行動を認識するためには、板書文字量が減少したことを確認すればよいが、板書文字が講義者の陰に隠れてしまったた

めに、板書文字量が減少してしまうことも考えられるため、板書文字量の変化のみから“黒板を消している”ことを認識すると誤認識に結びつくことが多い。そこで、“黒板を消している”場合には講義者が黒板消しを使用することから、黒板消しの色情報を用いて黒板下部のルール上における黒板消しの位置を推定し、黒板消しがルール上に存在するか否かを確認する。

まず、2.2.2 で肌色画素を抽出した方法と同様に、黒板消しを含む複数の画像から黒板消しを含む画素を抽出する。次に抽出した画素の画素値から黒板消しの色の平均値 μ_E および共分散行列 Σ_E を求め、これらの値を用いて入力画像中における黒板下部のルール上の画素 x の画素値 C_x が黒板消しの色にどの程度類似するかを表す値 $P_E(C_x)$ を次式より求める。

$$P_E(C_x) = \frac{1}{2\pi |\Sigma_E|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(C_x - \mu_E)\Sigma_E^{-1}(C_x - \mu_E)^T\right\} \quad (4)$$

次に、求めた値をしきい値処理により 2 値化することで得られた黒板消し領域を形成する全画素の重心を求め、その重心の座標値をルール上における黒板消しの位置とする。一方、黒板下部のルール上に黒板消し領域が抽出されなかった場合には、黒板消しがルール上に存在しないものとする。黒板消し位置の抽出結果を図 1 (e) に ×印で示す。

2.4 講義者の行動認識

講義者、板書に関する情報をもとに講義者の行動認識を行い、その結果をイベントとする。まず、現在のフレームで得られた情報と、過去のフレームで得られた情報から、講義者の位置変化 (POSI)、顔の向き (FACE)、手の位置変化 (HAND)、板書文字量の変化 (CHAR)、黒板消しの有無 (ERAS)

表 2: 撮影規則

イベント	撮影要求
1.“板書している”	講義者と講義者の手が含まれる 板書ブロック
2.“黒板を消している”	黒板全体
3.“板書について説明している”	講義者と講義者の手が含まれる 板書ブロック
4.“学生に対して説明している”	講義者のみ (中心となるように)
5.“左に移動している”	講義者のみ (左側に空きを作る)
6.“右に移動している”	講義者のみ (右側に空きを作る)

について調べる。そして、POSI, FACE, HAND, CHAR, ERAS を用いて経験的に作成した行動推定規則から講義者の行動を推定する。講義者の行動とその条件を表す行動推定規則を表 1 に示す。ただし、表中の * は don't care を表す。

例えば、講義者の位置が変化せず、黒板の方を向いて手を動かし、板書文字量が増加していれば、“板書している”と推定され、講義者の位置が変化せず、黒板の方を向いて手を動かし、板書文字が減少して黒板消しがルール上になれば、“黒板を消している”と推定されることになる。

3 状況理解による撮影領域の決定

講義中に受講者が注目している領域は講義者の行動によって大きく左右される。例えば、講義者が“板書について説明している”場合に受講者が注目する領域は、講義者と講義者が説明している板書になる場合が多い。このとき、受講者が注目する板書は講義者が直接説明している板書文字のみとは限らず、その板書文字を理解するための重要な情報を含む周囲の板書文字であることも考えられる。この場合に撮影領域に含むべき板書の範囲は講義者の手付近にある板書文字だけではなく、その板書文字と意味的につながりを持つ板書領域（以下、板書ブロックと呼ぶ）であると考えられる。そこで、文献 [11] で報告されている手法を用いて板書ブロックを抽出する。板書ブロックの抽出例を図 1 (e) に示す。図中の黒線で囲まれた領域が板書ブロックを表している。板書ブロックの抽出結果と講義者の手の位置から、現在説明に使用している板書ブロックを特定することができるため、講義者が“板書について説明している”ときには、撮影すべき領域を決定することができる。

以上のことを考慮して、講義を想定した知的自

表 3: 映像評価規則

イベント	評価法
1.“板書している”	講義者が使用している 板書がよく見える
2.“黒板を消している”	黒板と正対している
3.“板書について説明している”	講義者が使用している 板書がよく見える
4.“学生に対して説明している”	講義者と目線が合っている
5.“左に移動している”	講義者とカメラの位置が近い
6.“右に移動している”	講義者とカメラの位置が近い

動撮影システムの撮影規則を作成した。撮影規則は、イベントとイベントをトリガーとして発生する撮影要求を対応させたものである。講義を撮影する場合に、撮影領域を決定するためには講義者の行動が重要になると考えられるため、講義者の行動をイベントとして、撮影規則を表 2 のように決める。

例えば、講義者が“板書している”ならば、講義者と講義者が板書している板書ブロック全体を撮影領域とする。一方、講義者が“黒板を消している”時には、特に受講者が注目している領域はないと考えられるので、黒板全体を撮影領域としている。

また、撮影領域を決定する手法として、撮影要求に合致した撮影要求領域の加重平均を求めることによって撮影領域を決定するデジタルカメラワークを用いた手法 [5] を用いている。デジタルカメラワークは、撮影された映像の一部分を拡大表示することで、疑似的なカメラの動きを実現しているため、解像度が粗くなるという問題が残るが、機械的なカメラワークではなく自然なカメラワークを実現することが可能な手法である。

4 講義映像における映像評価

次にスイッチング規則について考える。複数のカメラを用いて撮影を行った場合にも、視聴者が同時に見ることのできる映像は一つであるため、複数のカメラから得られた映像の中から、最適な映像一つを選択するスイッチングが必要になる。スイッチャーは、各カメラのカメラワークを指示すると共に、それぞれの映像を見比べて、今の状況から考えてどの映像を視聴者が要求しているかを判断し、映像を切り替えていることが多い。ここでは、このような実際の撮影現場で取られている手法を計算機で実現する手法を試みる。

例えば、講義シーンを撮影対象とする場合に、カメラを選択する基準としては、板書が見やすいことや、講義者と目線が合っていることなどが考えられる。板書が見やすいように撮影すべき講義状況は、講義者が“板書している”場合や“板書について説明している”場合であり、講義者と目線が合うように撮影するのは、講義者が“学生に対して説明している”場合になることが多い。また、講義者が移動している場合には、その移動方向に合わせて撮影カメラを変えていくことなども考えられる。

このように、設定したイベントがトリガーとなって映像の評価規準も変化するものと考えられる。そこで、複数のカメラから得られた映像に対してシーン上で起きているイベントにあわせた映像評価法を設計しておき、それぞれの映像からその評価値が最も高い映像を選択する。

講義を対象としたスイッチングのための映像評価規則を表3に示す。イベントに対応した評価値が最も高くなる映像を選択する。ここでは、撮影するためのカメラは3台とし、図2のように正面・左・右にカメラが黒板中央から等距離になるように設置した。各カメラから得られた映像を入力として、それぞれ講義者の行動推定を行い、その結果に基づいて撮影領域を決定し、それぞれのカメラで映像を得る。そして、得られた映像に対する評価値を求め、最も評価値が高い映像を選択する。

評価値は0~1の値を持つように設定した。例えば講義者が“板書している”，または“板書について説明している”場合には、使用している板書ブロックと講義者の重なりが少なく、かつ離れ過ぎない

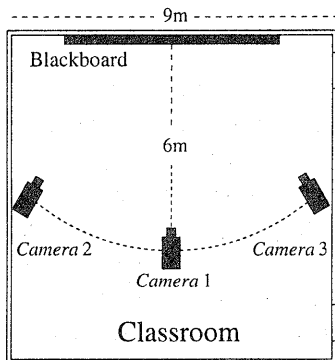


図2: 撮影カメラの配置

いほど、使用している板書ブロックは受講者にとって見えやすく映る。そこで板書ブロックの横端に講義者位置が重なる時に最大値1をとるような正規分布関数を考え、評価値とする。講義者が“学生に対して説明している”場合には、映像を見ている受講者と講義者の目線が合うように講義者の顔がより正面を向いている映像が得られた時に評価値が高くなるようにする。ここでは、頭部の肌色画素の数を頭部の画素数で割ったものを評価値としている。また、講義者が左または右に移動している時は、図2において講義者が左付近にいる場合は Camera 2, 中央付近にいる場合は Camera 1, 右付近にいる場合は Camera 3 で撮影した映像を選ぶ。講義者が“黒板を消している”場合には、黒板全体が撮影されるように規則を作っているため、Camera 1 で撮影した映像を選択する。

5 実験および考察

5.1 実験

本手法を用いて実際に講義の知的自動撮影システムを構築した。画像の処理領域は黒板全面が映る大きさとして 320 × 80 画素とした。また、左右に設置したカメラ Camera 2, Camera 3 から得られる映像は、黒板に対して正対しておらず、板書ブロックが矩形で表せないため、黒板が長方形となるようにアフィン変換を施して処理を行った。また、3台の計算機を用いて同時にイベント抽出などの処理を行っているため、計算機によってイベント抽出や撮影領域の決定結果に違いを生じることがある。その場合には多数決によって最も多くの計算機に支持されている結果を採用した。また、すべての計算機において結果が異なる場合には中心に設置した Camera 1 の映像から算出した結果を用いた。処理速度は毎秒3フレーム程度である。

実験によって得られた講義撮影映像を図3に示す。図3の映像は、左の列から順に講義者が“板書している”，“板書について説明している”，“学生に対して説明している”，“黒板を消している”場合に各カメラで撮影された映像を表している。また、映像の下にある数字はそれぞれの映像の評価値を示しており、各列で黒い太枠で囲まれている映像は、それぞれの時刻で最も評価値が高い映像を表している。その結果得られた映像を図3の最下段



図 3: 講義映像生成の例

に示している。

講義者が“板書している”場合や、“板書について説明している”場合には、使用している板書ブロックが講義者によって遮られることなく最も見やすいと思われる映像が選択できていることがわかる。また、“学生に対して説明している”場合には講義者の顔の向きがより正面となり目線があっている Camera 3 の映像が選択できていることが確認できる。

5.2 考察

提案した知的自動撮影システムについて考察する。

(1) 講義の知的自動撮影

講義を撮影対象として、シーンの状況を理解することで撮影領域の決定を行い、異なる

位置に配置した複数のカメラから得られた映像を評価することでスイッチングを行う知的自動撮影システムを構築した。構築した知的自動撮影システムを用いて講義を撮影することで、板書などが見え易いカメラ位置から講義に適した映像を撮影することができる。

(2) 受講者の満足度を満たす映像生成への応用

本研究では講義撮影すべての自動化を目的としている。しかし、双方向テレビなどを用いた遠隔講義への利用を考えた場合には、受講者の要求に合わせた映像生成の必要性もでてくると考えられる。受講者の要求を満たす映像を生成するためには、本手法で自動化している部分の一部である撮影要求の決定部分に受講者の要求を調停するなどの手法で容易に応用できると考えられる。

(3) 撮影規則・映像評価規則の作成

今回用いた撮影規則と映像評価規則は、講義者の行動の中から主な6種類をイベントとして経験的に作成している。しかし実際にカメラマンが撮影したり、スイッチャーが映像切り替えを行う際には、さらに多くの講義者の行動を理解していると考えられる。認識対象とする講義者の行動(イベント)を増やすことができれば、さらに効果的な映像生成が可能になると思われる。

(4) スイッチングの問題

本手法では映像を生成する際に各カメラで撮影された映像を評価し、最も評価値が高いものをその時刻において最適な映像として選択した。しかし実験で得られた講義映像では、不必要なスイッチングが行われることもあり、視聴者にとって見づらくなると考えられる箇所があった。この問題は、現在の各映像の評価値だけではなく過去における評価値も考慮することや、ショットの最短持続時間を設定することなどで解決できると考えている。

(5) 生成映像の評価

実験で生成した講義映像が、受講者にとってどの程度満足度の高いものであるかを評価する必要がある。評価方法としては、作成した撮影規則をどの程度忠実に再現できているかを調べる方法や、本手法で生成した映像と人手によって生成した映像を比較し、受講者がどのような相違を感じるかを調べる方法などが考えられるが、今後の課題とする。

(6) 他シーンへの適用

今回は講義を撮影対象として知的自動撮影システムを構築した。しかし講義に限らず、撮影したいシーン中のイベントをそのシーン映像から抽出することができれば、スポーツなどの様々なシーンに本手法を適用することができる。

6 まとめ

本稿では、シーンの状況理解に基づいて自動的に撮影し、撮影された複数の映像を評価すること

によって適切な映像を選択して映像生成を行う手法を提案した。また、これらの手法を講義撮影に適用することについて述べた。

今後の課題として、本手法で生成した映像を視認評価することなどが挙げられる。

参考文献

- [1] 灰塚凡樹, 井上誠吾, “カメラワーク生成に関する一考察,” 信学技報, PRMU96-201, pp.67-74, Mar. 1997.
- [2] 田中 彰, 東海彰吾, 松山隆司, “イベント駆動型カメラワークによる動的シーンの効果的映像化,” 情処研報 CVIM-121-10, pp.73-80, Mar. 2000.
- [3] Claudio.S.Pinhanez and Aaron.F.Bobick, “Intelligent Studios: Using Computer Vision to Control TV Cameras,” Proc. of IJCAI'95 Workshop on Entertainment and AI/Alife, pp.69-76, Aug. 1995.
- [4] 井口泰典, 土居元紀, 眞鍋佳嗣, 千原國宏, “アクティブカメラの協調によるスポーツ映像撮影システム,” 信学技報 PRMU2000-46, pp.37-44, Jul. 2000.
- [5] 大西正輝, 泉 正夫, 福永邦雄, “デジタルカメラワークを用いた自動映像生成,” 画像の認識・理解シンポジウム (MIRU2000), pp.I-331-I-336, Jul. 2000.
- [6] 松本圭介, 須藤 智, 斉藤英雄, 小沢慎治, “サッカーシーンにおけるボール追跡に基づく最適視点決定システム,” 信学技報 PRMU2000-27, pp.29-36, Jun. 2000.
- [7] 亀田能成, 石塚健太郎, 美濃導彦, “状況理解に基づく遠隔講義のための実時間映像化法,” 情処研報 CVIM-121-11, pp.81-88, Mar. 2000.
- [8] 大西正輝, 泉 正夫, 福永邦雄, “情報発生量の分布に基づく遠隔講義撮影の自動化,” 信学論 (D-II), vol.J82-D-II, no.10, pp.1590-1597, Oct. 1999.
- [9] 大野直樹, 先山卓朗, 椋木雅之, 池田克夫, “遠隔講義における講義状況に応じた送信映像選択,” 第5回知能情報メディアシンポジウム論文集, pp.31-38, Dec. 1999.
- [10] 尾関基行, 中村裕一, 大田友一, “プレゼンテーションの知的撮影システム-手元作業を対象とした適応的カメラワーク-,” 信学技報 PRMU2000-104, pp.31-38, Nov. 2000.
- [11] 大西正輝, 泉 正夫, 福永邦雄, “講義映像における板書領域のブロック分割とその応用,” 信学論 (D-I), vol.J83-D-I, no.11, pp.1187-1195, Nov. 2000.