

# 撮影対象の運動履歴に基づく固定ショット切り替え式撮影法

亀田 能成<sup>†</sup> 新 康孝<sup>††</sup> 西口 敏司<sup>†††</sup> 美濃 導彦<sup>†</sup>

<sup>†</sup> 京都大学学術情報メディアセンター 〒606-8501 京都市吉田二本松町

<sup>††</sup> 京都大学大学院情報学研究科

<sup>†††</sup> 京都大学大学院法学研究科

E-mail: <sup>†</sup>{kameda,minoh}@media.kyoto-u.ac.jp, <sup>††</sup>{atarashi,nishigu}@mm.media.kyoto-u.ac.jp

あらまし 複数の首振りカメラを用いて、ある範囲を自由に動き回る撮影対象を、固定ショットのみで自動的に撮影する方法について提案する。固定ショット切り替え式撮影をすれば、視聴者の映像閲覧に対する心理的負担を軽減することができる。固定ショット撮影では、人物撮影が開始されてからは一切制御が不可能となるため、事前に、撮影対象の振舞とカメラ制御パラメータとの関係を把握することが重要である。本研究では、撮影対象の運動履歴に注目し、3階層からなる映像評価基準に従って運動履歴を分割していくことで固定ショット撮影ルールを生成する。

キーワード 固定ショット, 自動撮影, 運動履歴, カメラ制御

## A Filming Method with Fixed Shots Based on Object Trajectory Records

Yoshinari KAMEDA<sup>†</sup>, Yasutaka ATARASHI<sup>††</sup>, Satoshi NISHIGUCHI<sup>†††</sup>, and Michihiko MINOH<sup>†</sup>

<sup>†</sup> Academic Center of Computing and Multimedia Studies, Kyoto University Yoshidanihonmatsucho,  
Kyoto, 606-8501 Japan

<sup>††</sup> Graduate School of Informatics, Kyoto University

<sup>†††</sup> Graduate School of Law, Kyoto University

E-mail: <sup>†</sup>{kameda,minoh}@media.kyoto-u.ac.jp, <sup>††</sup>{atarashi,nishigu}@mm.media.kyoto-u.ac.jp

**Abstract** We propose a new camera control method that can image a free-moving person within a certain area by using multiple pan-tilt-zoom cameras. We compose video images only by camera-fixed shots in order to let the video images more comfortable to see. As a camera on filming fixed shots must not move, it is important to obtain a set of imaging rules that can arrange pan/tilt/zoom parameters of the cameras to frame target person before they start filming. We obtain the rule set by segmenting object trajectory that was recorded in advance.

**Key words** Fixed shot, Automatic filming, Object trajectory record, Camera control

### 1. はじめに

映像メディアコンテンツにおいて、大きな比率を占めるのが人物撮影である。特に大学においては講義等をしている人物の振舞や所作を映像化することができれば、価値あるコンテンツを多く作成することができ、社会的影響も大きい。

人物撮影において、その自然な所作を映像化するためには、その動ける範囲や動作内容を制限するべきではない。このことから、撮影対象が自由に動き回る状況下での自動撮影が近年望まれている。このとき、自動撮影の結果である映像は、視聴者である人間にとって見やすい映像でなくてはならない。

本研究では、映像の最小単位をショットと呼ぶ。ショットとは、ある1つの首振りカメラの撮影によって生成された映像単位であり、ショット内には不連続的な映像は存在しないものと

する。映像は複数のショットによって構成されるので、人間にとって見やすい映像の評価基準は、以下の3階層で表現できる。

- ショット評価
- ショット間評価
- ショット構成評価

ショット評価とは、ショット単独で評価される要因であり、画像上での対象の配置などの構図などが挙げられる。ショット間評価は、2つのショットの切り替えにおいて評価される要因であり、構図の差などがその例である。ショット構成評価は映像全体を通じて評価されるべき要因であり、撮影対象や構図の選択、切替のタイミングなどがこれに含まれる。

これまで提案されている複数カメラを用いた自動撮影の研究 [1] [2] [3] [4] [5] では、ショット評価が高くなるような撮影方法が工夫されていたり、ショット間評価について考慮した方法

が提案されている。これらの研究では、対象がどのような動きでどの範囲にいる場合にどう撮影するかという、状況と撮影のセグメント化とその組み合わせ方法はトップダウンで与えられている。そのため、新たな状況を想定する場合、システム構築者による撮影ルールの再構築が必要となり、撮影の汎用的な自動化は難しい。

それに対し、我々は、映像評価基準を、状況と撮影のセグメント化およびその組み合わせの決定に用いることで、より汎用的な自動撮影方法の実現を目指している。

本稿では、自由に動き回る単一人物の撮影を取り上げる。このような前提のもとでは、撮影内容としてはその対象を撮影し続けるだけでよく、望ましい構図がほぼ一定でよいと考えられる。これは、ショット構成が比較的単純になることを意味する。この場合、映像獲得のための映像評価の問題は、ショットとショット間の評価を中心に定式化できる。この定式化は、対象の軌跡と撮影カメラ群のカメラパラメータだけを用いて表現することが可能である。さらに、あらかじめ対象の運動傾向を獲得しておけば、事前に分析を行い映像評価がより高くなるように状況と撮影のセグメント化を行い、各状況に対する最適なカメラ制御方法を定めることもできるようになる。本稿ではこれを撮影ルールと呼ぶ。我々の提案する手法では、映像評価基準に従って撮影ルールを生成することができる。

一方、撮影ルールを生成する際、どのような撮影技法を基にするべきかという問題がある。撮影技法には大きく分けて、追跡撮影と固定撮影の2種類が存在する。対象が運動する場合、これまで、視覚サーボ [6] [7] による手法や予測を用いた速度制御による手法 [8] [9] などの追跡撮影手法が提案されてきている。しかし、映画などやテレビ番組においてみられるように、撮影は主に固定ショットの切り替えによって行われており、追跡撮影・流し撮りはそのショットを挿入することに意味がある場合のみに抑えられている [10]。我々は映像メディアの消費者が機械観測装置ではなく人間であることを想定しているため、固定撮影を前提とし、固定ショット切り替え式撮影によって映像を獲得することを考える。

複数のパン・チルト・ズームカメラを用いて固定ショット切り替え式撮影を実現する場合、対象の動きに合わせてカメラのパン・チルト・ズームをどのように設定し、カメラの切り替えをどのように行うかという固定ショット撮影ルールの決定が非常に重要となる。なぜなら、カメラの切り替えを行った後は、撮影に入った出力カメラのパン・チルト・ズーム値は対象の動きに関わらず変更することができないためである。従って、撮影に入る前に事前に適切なパン・チルト・ズーム値を設定しておき、カメラの切り替えに備えなければならない。これは、対象の運動に対する十分な知識なしには不可能である。本研究では、事前に獲得した対象の運動履歴を利用・分析することで、この問題を解決する。

以上を踏まえ、本研究では、事前に獲得した対象の運動履歴から、映像評価基準を満たす固定ショット撮影ルールを生成し、そのルールを用いて複数カメラの固定ショット切り替えによる撮影を行う手法を提案する。運動履歴と映像評価基準を導入することで、事前に自動的に固定ショット撮影ルールを生成する。撮影段階においては、観測によって得られた情報と生成済みの

固定ショット撮影ルールからカメラの制御と切り替えを行う。

本研究では、運動履歴は世界座標系での対象の位置・顔の向き・進行方向からなるものとする。また、固定ショット撮影ルールは、観測によって得られた情報から、各カメラのパン・チルト・ズーム値ならびに出力として選択するカメラ番号への写像として表現できる。

以下、2. では、映像評価基準の定義法について述べる。3. では固定ショット切り替え式撮影の枠組みについて述べると共に、映像評価基準から固定ショット撮影ルールを生成する手法について述べる。4. では本手法の有効性について実験を通して確認する。5. は本研究のまとめである。

## 2. 映像評価基準

1つの対象を撮影して得られる映像を規定する要因としては、前景と後景の2つが考えられるが、本稿では前景についてのみに考える。

撮影対象である1人の人物について、世界座標系でのその位置を  $X$ 、人物が向いている方向を  $F$ 、そのときの移動方向を  $V$  とし、これらをまとめて  $P = (X, F, V)$  で表現する。また、カメラ  $i$  のパラメータを  $C^{(i)}$  で表し、これにはパン・チルト・ズーム値が含まれる。撮影に供されているカメラを特に出力カメラと呼び、カメラ番号  $j$  で表す。

本研究では、ある時刻  $t$  におけるショット・ショット間・ショット構成の映像評価をそれぞれ関数形式で表す。

まず、ショットに関する評価関数  $E_f$  を

$$E_f(P, j, c^{(j)}) = \left( w_{iloc} \|x^{(j)} - \tilde{x}\|^2 + w_{isize} \|k^{(j)} - \tilde{k}\|^2 + w_{wdir} \|F^{(j)} - \tilde{F}\|^2 + w_{spd} \|V^{(j)} - \tilde{V}\|^2 \right)^{\frac{1}{2}} \quad (1)$$

と定義する。ただし、 $x^{(j)}, k^{(j)}$  はカメラ  $j$  の画像上での対象の位置と大きさを表し、 $F^{(j)}, V^{(j)}$  はカメラ  $j$  の三次元座標系での向きと運動方向を示す。 $\tilde{x}, \tilde{k}, \tilde{F}, \tilde{V}$  は与えられた目標値であり、物体が画像上でどのようにフレーミングされるかを示している。式 (1) は理想とするフレーミングからのずれを数値化していると言える。

次に、ショット間の切り替えに対する評価関数  $E_d$  を下のよう

$$E_d(P, j, c^{(j)}) = \left( w_{iloc} \|x_t^{(j)} - x_t^{(j-t-0)}\|^2 + w_{isize} \|k_t^{(j)} - k_t^{(j-t-0)}\|^2 + w_{wdir} \|F_t^{(j)} - F_t^{(j-t-0)}\|^2 + w_{spd} \|V_t^{(j)} - V_t^{(j-t-0)}\|^2 \right)^{\frac{1}{2}} \quad (2)$$

ここで、 $t-0$  は時刻  $t$  の直前の時刻を示す。本式は、出力カメラが切り替わる時刻ではフレーミングの相違に基づく値をペナルティとして与え、そうでない固定ショット区間では0となる。

最後に、ショット構成の評価関数  $E_s$  であるが、本研究では、ある1つの対象を撮影し続け、その撮影を通して望ましいフレーミングは一定であるとする。そのため、ここではショット

切替の回数が多いほど見にくい映像になることのみを考慮し、ショット切り替えの回数に基づく評価関数を定義する。

$$E_s(t) = \sum_u \delta(t - t_u) \quad (3)$$

ただし、 $\delta(\cdot)$  はディラックのデルタ関数であり、 $t_u$  は映像切替が生じた時刻を示す。

全体的な映像評価  $E$  は、式 (1), (2), (3) を撮影時間で積分して得られる。

$$E = \int (w_f E_f + w_d E_d + w_s E_s) dt \quad (4)$$

### 3. 固定ショット切り替え式撮影

#### 3.1 固定ショット切り替え式撮影の定義

本稿では、状況を規定するのは対象物体の観測データ  $P$  である。つまり、自動撮影とは、観測データ  $P$  を基に、出力カメラ  $j$  と各カメラのカメラ制御パラメータ  $c^{(j)}$  をまとめた制御ベクトル  $(j, c^{(1)}, \dots, c^{(N)})$  を実時間で決定していく処理と捉えることができる。ここで  $N$  はカメラの台数である。この制御ベクトルのうち、出力カメラに関する  $(j, c^{(j)})$  を特に出力映像制御パラメータと呼ぶ。

固定ショット切り替え式撮影とは、制御ベクトルの決定において、 $\{c^{(i)} | i \neq j\}$  のみを変化させてよい撮影方式である。さらに、カメラ制御パラメータ  $c^{(i)}$  の変更には一定時間必要であることを考慮すると、非出力カメラのカメラ制御パラメータの変更はそのカメラが出力カメラに切り替わる可能性が生ずる前に終了してはならない。

固定ショット切り替え式撮影のもう一つの特徴は、あるショットが継続される間は出力カメラが固定であるため、出力映像制御パラメータ空間が必ず離散的に分割されることである。つまり、離散化された出力映像制御パラメータ部分空間ごとに対応する観測データ空間を割り当て、かつ、非出力カメラの制御ベクトルのみの変更で1つの部分空間から別の部分空間への遷移を実現することが、本研究の目的となる。

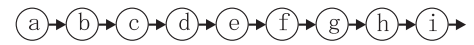
#### 3.2 運動履歴の出力映像制御パラメータによる構造化

観測データ空間は観測データの集合  $\{P\}$  から構成される。観測データの系列を運動履歴と呼ぶ。図 1(a) に例を示す。なお、観測データは一定時刻  $\Delta t$  毎に得られるものとする。前節で述べたように、最終的には観測データ空間は分割されるので、ここでは観測データ空間を離散的に表現することを考え、それをもとに観測データと出力映像制御パラメータとの関係を記述する。

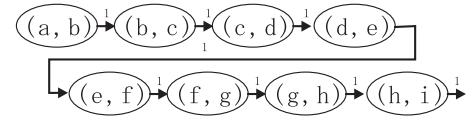
まず、物体の運動にはマルコフ性があると仮定し、各時刻  $t$  において、過去の  $d\Delta t$  間の観測をまとめてグラフで表現する。図 1(b) がその例である。図中の矢印は遷移を表し、その添字は遷移回数を示す。当然ここでは全ての遷移回数は 1 である。

ここで、観測履歴データの集合  $\{(a, b), (b, c)\}, \{(c, d), (d, e), (f, g)\}, \{(e, f), (g, h)\}, \{(h, i)\}$  の其々に、異なった出力映像制御パラメータが割り当てられたとする。同じ出力映像制御パラメータをもつノードを一つのノードに統合することを考えると、グラフは図 1(c) のように簡約化され、矢印の添字で表している遷移回数も変化する。

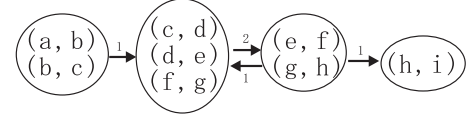
こうして得られた簡約運動履歴グラフのことを、出力映像グ



(a) 観測データの系列



(b) ノード統合前の運動履歴グラフ ( $d = 1$ )



(c) 出力映像グラフの生成

図 1 運動履歴からの出力映像グラフの生成

ラフと呼ぶ。出力映像グラフは、観測データと出力映像制御パラメータとの関係を記述している。しかしながら、これだけでは固定ショット切り替え式撮影の実現にはならない。なぜなら、ノードを統合する基準を与える問題と、固定ショット撮影であることを保証するように出力カメラを選択する問題が残されているからである。次節でこれらの問題を解決する方法を述べる。

#### 3.3 固定ショット撮影グラフの生成

固定ショット撮影ルール生成は、運動履歴と映像評価基準から出力映像制御パラメータを考慮しつつ出力映像グラフを生成する第一段階と、固定ショット撮影の定義に従って制御ベクトルを決定する第二段階とに分かれる。第二段階では、各ノードでは其々のカメラの制御パラメータが全て設定されることになり、かつ、この段階の処理では必要に応じてグラフは書き直される。このようにして最終的に得られるグラフのことを固定ショット撮影グラフと呼ぶ。これが固定ショット撮影ルールとなる。

第二段階で問題になるのは、固定ショット切り替え式撮影ゆえの出力カメラの選択に関する以下の2つの制約条件である。なお、ここで、出力映像グラフの段階で、ある二つのノード  $A, B$  が同じ出力映像制御パラメータ  $(j_A, c^{(j_A)}) = (j_B, c^{(j_B)})$  である場合はすでにノード統合されているものとする。

**遷移間カメラ選択条件** あるノード  $A$  に対して出力映像制御パラメータ  $(j_A, c^{(j_A)})$  が割り当てられており、ノード  $A$  から遷移するノード  $B$  に対して、出力映像制御パラメータ  $(j_B, c^{(j_B)})$  が割り当てられているとする。ここで、 $j_A = j_B$  かつ  $c_A^{(j_A)} = c_B^{(j_B)}$  の場合は存在せず、かつ一方で、 $j_A = j_B$  かつ  $c_A^{(j_A)} \neq c_B^{(j_B)}$  の場合は固定ショット撮影要件に反する。従って、 $j_A \neq j_B$  でなければならない。

**兄弟ノードカメラ選択条件** ノード  $A$  から遷移する  $B$  以外のノード  $C$  に対して、出力映像制御パラメータ  $(j_C, c^{(j_C)})$  が割り当てられているとすると、固定ショット撮影の定義から、ノード  $A$  の制御ベクトル  $(j_A, c_A^{(1)}, \dots, c_A^{(N)})$  について、 $c_A^{(j_B)} = c_B^{(j_B)}$  かつ  $c_A^{(j_C)} = c_C^{(j_C)}$  である。このとき、 $j_B = j_C$  かつ  $c_B^{(j_B)} = c_C^{(j_C)}$  の場合は出力映像グラフの定義から存在しない。 $j_B = j_C$  かつ  $c_B^{(j_B)} \neq c_C^{(j_C)}$  とすると、 $c_B^{(j_B)} = c_A^{(j_B)} = c_A^{(j_C)} = c_C^{(j_C)}$  となり矛盾する。従って、 $j_B \neq j_C$  でなければならない。

各ノードに映像評価値を最小にするカメラを割り当てることとこれらの2条件は必ずしも両立しないため、出力カメラの選

択について最適解が常に見つかる保証はない。そこで、出力カメラの決定を次のように実現する。まず全てのノードに対して出力カメラを決定していない状態で、映像評価基準を基に同一の出力映像制御パラメータを割り当ててよいと見なせるノードを統合し、その後、統合後のノードに対して出力カメラを割り当てる。

以上をまとめて、固定ショット撮影グラフの生成処理を手順として記述すると以下ようになる。

- (1) 映像評価基準に基づくノード統合
- (2) 固定ショット撮影実現のためのノード統合
- (3) 出力映像制御パラメータの決定
- (4) 非出力カメラの制御パラメータ決定

### 3.3.1 映像評価基準を基にしたノードの統合

映像評価基準を基にしたノードの統合を行う方法について述べる。あるノードの組について映像評価値の見積もりを行い、統合せず遷移を残した場合よりも、統合して1つのノードにする方が見積もりが小さくなる場合、このノードの組を統合する。ノード  $A$  とノード  $B$  が統合の候補であり、統合したと仮定した場合のノードを  $M$ 、 $A$  から  $B$  へのアークの重みを  $W_{AB}$  と表すことにする。運動履歴グラフの定義より、 $W_{AB}$  はノード  $A$  からノード  $B$  への遷移の回数を表す。

#### (i) ショット評価

観測履歴データ  $s$  があるノード  $Z$  に含まれることを  $s \in Z$  と表し、ノード  $Z$  内での観測履歴データ  $s$  の重みを  $W_Z^s$ 、観測履歴データ  $s$  のうち最新の観測データを  $P_s$  と表すことにする。

ノード  $A$  とノード  $B$  を統合した場合のショット評価  $E_f$  の変化値の見積もり  $\Delta E_f$  は、式 (1) を利用して、ノード  $M$  に対する評価見積もり値からノード  $A$  に対する評価見積もり値とノード  $B$  に対する評価見積もり値とを引いた値、すなわち、

$$\begin{aligned} \Delta E_f &= \min_j \min_{c^{(j)}} \sum_{s \in M} W_M^s E_f(P_s, j, c^{(j)}) \\ &\quad - \min_j \min_{c^{(j)}} \sum_{s \in A} W_A^s E_f(P_s, j, c^{(j)}) \\ &\quad - \min_j \min_{c^{(j)}} \sum_{s \in B} W_B^s E_f(P_s, j, c^{(j)}) \end{aligned} \quad (5)$$

となる。式 (5) は非負であるので、ノード統合はショット評価を常に悪化させる。

#### (ii) ショット間評価

ノード  $A$  からノード  $B$  への遷移の瞬間では、式 (2) において、 $j_t = j_B, j_{t-0} = j_A$  とするとショット間評価が得られる。これを  $E'_d(P, j_A, j_B)$  と表記する。ショット間評価では、統合後の評価式  $E'_d(P, j_M, j_M) = 0$  であることから、変化値の見積もり  $\Delta E_d$  は次式のように表せる。

$$\begin{aligned} \Delta E_d &= - \left( W_{AB} \frac{\sum_{s \in A} W_A^s E'_d(P_s, j_A, j_B)}{\sum_{s \in A} W_A^s} \right. \\ &\quad \left. + W_{BA} \frac{\sum_{s \in B} W_B^s E'_d(P_s, j_B, j_A)}{\sum_{s \in B} W_B^s} \right) \end{aligned} \quad (6)$$

なお、ノード  $A$  とノード  $B$  の統合に伴い、出力映像制御パラメータが変更されるため、ノード  $M$  への遷移時のカメラ切り替えに対応する  $E_d$  の増減と、ノード  $M$  からの遷移時のカ

メラ切り替えに対応する  $E_d$  の増減についても評価を行うべきであるが、これらはノード  $M$  に隣接するノードの統合によって影響を受けるため、局所的な計算だけで評価を行うことができない。そのためノードの統合段階ではショット間設定についてはこれ以上の評価を行わず、出力映像制御パラメータの決定段階で評価を行うことにする。

#### (iii) ショット構成

ショット構成については、ノード  $A$  とノード  $B$  の間での遷移がなくなるため、変化量  $\Delta E_s$  は下記のように示される。

$$\Delta E_s = - \left( W_{AB} + W_{BA} \right) \quad (7)$$

と表せる。

#### (iv) 総変化量

式 (5)(6)(7) から、総変化量の見積もり  $\Delta \tilde{E}$  は

$$\Delta \tilde{E} = w_f \Delta E_f + w_d \Delta E_d + w_s \Delta E_s \quad (8)$$

で与えられる。映像評価基準に基づくノードの統合処理は次のように行う。

- (1) 遷移の存在するノードの組について  $\Delta \tilde{E}$  を求める
- (2)  $\Delta \tilde{E}$  の最小値が非負の場合はノードの統合を終了
- (3)  $\Delta \tilde{E}$  の最小値を与えるノードの組について統合を行い、各ノードに関係する遷移もそれに伴い統合する
- (4) (1) へ戻る

### 3.3.2 固定ショット撮影実現のためのノード統合

映像評価基準に基づくノードの統合を行った後も、運動履歴グラフは遷移先選択カメラ条件、兄弟ノード選択カメラ条件を満たしているとは限らない。この状態を競合と呼び、本研究では競合を解決する手段としてさらなるノード統合を行う。

具体的には、以下の3種類の処理を実行する。

**競合ノード統合** 競合しているノードの組を統合する。統合前の出力カメラがそのまま統合後のノードの出力カメラとなる。  
**出力カメラ変更** 競合しているノードに割り当てられている出力カメラを変更する。

**隣接ノード吸収** 競合しているノードの遷移元ノード、もしくは遷移先ノード、もしくは同じ遷移元ノードをもつノードで、かつ別の出力カメラが割り当てられているノードに吸収させる。吸収するノードの出力カメラが吸収後のノードの出力カメラとなる。

競合解決処理は以下のように行う。

- (1) 競合しているノードの組数が0ならば終了
- (2) 1回の競合解決処理によって減少する競合数  $\Delta D$  が1以上である処理全てに対して、競合するノード1組あたりの映像評価値の変化値の見積もり  $\frac{\Delta \tilde{E}}{\Delta D}$  を求める
- (3)  $\frac{\Delta \tilde{E}}{\Delta D}$  が最小となる競合解決処理を実行する
- (4) (1) へ戻る

ここで  $\tilde{E}$  については、式 (8) と同様の定式化を行って得る。

### 3.3.3 出力映像制御パラメータの決定

これまでの段階で、各ノードに対して出力カメラが割り当てられており、遷移間カメラ選択条件、兄弟ノードカメラ選択条件を満たしたグラフが得られる。ここでは、そのグラフに従って、各ノードについて出力カメラ毎に出力映像制御パラメータを決定する。映像評価全てについて大局的に最適なカメラ制

御パラメータを求めることは困難であるので、各ノード毎に、当該ノードのショット評価関数と、そのノードからまたはそのノードへ遷移するノードとの間のショット間評価関数との和を最適にするようなカメラ制御パラメータを求め、出力映像制御パラメータとする。

### 3.3.4 非出力カメラの制御パラメータ決定

3.1節で述べたように、カメラ切り替え時に、新たに出力カメラになるカメラのカメラ制御パラメータを変化させないで済むように、非出力カメラのカメラ制御パラメータを決定する。得られているグラフは二つのカメラ選択条件を満たしているため、新たなショットが始まる時点で、非出力カメラのカメラ制御パラメータの変更値も同時に決定できる。つまり本手法では、1つのショットがごく短い時間で終了しない限り、必ず出力カメラになる可能性のある非出力カメラは固定値で映像切替を待つことができる。これによって、最終的な固定ショット撮影グラフを得る。

## 4. 実験と考察

京都大学学術情報メディアセンター南館 R202 教室を対象空間とし、シミュレーションと実験を行った。撮影対象は講義中の講師である。使用したカメラ数は  $N = 4$  であり、いずれのカメラも SONY 製 EVI-D30 である。講師の移動範囲を幅 14.0m、奥行き 4.0m としている。

運動履歴の獲得には、InterSense 製超音波式ポジジョンセンサー IS-600 MarkII を用い、ビーコンを講師の両肩に取り付けて  $P = (X, F, V)$  を得た。観測データの量子化については、 $X$  については 50cm 四方で床面を均等分割した。  $V$  については、 $0.0 \leq \arctan(V_Z, V_X) < 2\pi$  を均等に 10 分割し、 $F$  についても同様に  $0.0 \leq \arctan(F_Z, F_X) < 2\pi$  を均等に 10 分割した。時間については、 $\Delta t$  が 0.5 秒となるように処理を行った。

### 4.1 固定ショット撮影グラフの生成

講義 90 分の観測データを運動履歴として用いて、撮影対象である講師に対して固定ショット撮影グラフを生成した。

固定ショット撮影グラフのノードとアークを、図 2,3 に示す。あるノード  $Z$  内の数字は、左がそのノード  $Z$  に含まれる観測履歴データ  $s \in Z$  の数、右はその重みの総和  $\sum_{s \in Z} W_Z^s$  を表し、あるノードから別のノードへのアークの隣の数字はその重みを表す。図 2,3 は、撮影対象の講師が異なるほかは、いずれも、 $d = 1$ 、 $w_f = w_d = w_s = 1$  として映像評価基準を評価している。また、構図パラメータについては  $w_{iloc} = w_{isize} = 1$ 、 $w_{wspd} = w_{wdir} = 0$  として、対象の位置、大きさについて等しい重みとし、進行方向、向きについては考慮しない場合の結果である。

また、提案手法により約 6 分間の模擬講義の映像取得を行った。事前に与える運動履歴としては模擬講義 3 回分のデータを与えた。観測履歴データの長さは  $d = 0$  とし、映像評価基準の重みについては、 $w_f = 10$ 、 $w_d = w_s = 1$  とし、構図パラメータについては  $w_{iloc} = w_{isize} = w_{wspd} = 1$ 、 $w_{wdir} = 0$  として、対象の位置と大きさ、および顔の向きを考慮するよう設定した。実制御の実験では、観測データの獲得には超音波センサーを用いず、エビポーラ拘束を用いた画像からの講師の 3 次元位置推定を行っている。講師の向きは、速度方向から予測した。得ら

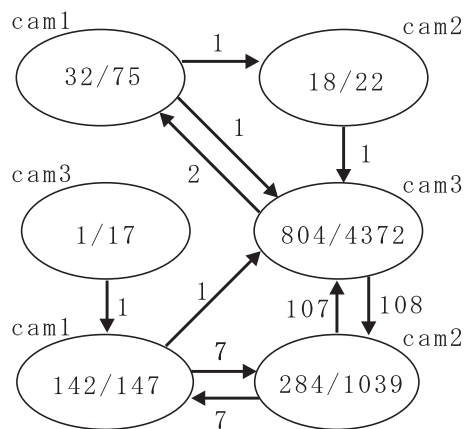


図 2 固定ショット撮影グラフの例 1

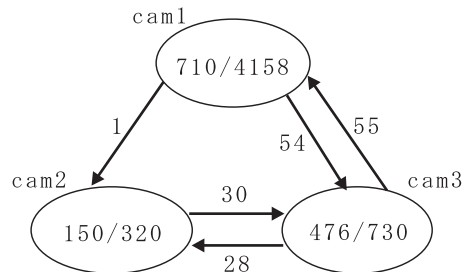


図 3 固定ショット撮影グラフの例 2

れた映像の系列を図 4 に示す。一枚の画像が一つの固定ショットを表し、白抜き文字がショットの継続時間を表す。時系列は、左から右の後、下の段となっている。



図 4 取得された映像

検証用に取得したポジジョンセンサーによる観測データを用いて、撮影に用いた制御ベクトルから計算した評価値と、シミュレーションを行って計算した評価値を表 1 に示す。

本実験では、運動履歴にないような観測データが撮影中に得られた場合、固定ショット切り替え式撮影システムと別に用意したマスターショットカメラに遷移し、固定ショット撮影グラフ内の遷移に戻れる状態になるまでマスターショットを使い続ける。表 1 のマスターショット率とは、撮影時間中のマスターショットの長さを百分率で示したものである。また、フレームアウト

実験方法	Master Shot[%]	Frame Out[%]	$\bar{E}$
シミュレーション	1.94	0.24	27.13
実制御	4.37	10.68	30.65

表 1 生成映像の評価

率は対象物体が映像中にいなかった時間の長さを示し、平均映像評価値  $\bar{E}$  は式 (4) を撮影時間で割った値である。シミュレーション結果と比較して、実制御の結果においてマスターショット率が増加し、映像評価値がよくない理由として画像処理による講師の三次元位置推定制度が悪いことが挙げられる。カメラ数を増加させ、対象の周囲に設置することができれば観測精度を向上させることは可能である。また、フレームアウト率が下がったのは、観測誤差の他、観測やカメラ切り替え制御の処理遅れが原因である。これについては、出力映像制御パラメータを決定する際に、観測履歴データに含まれる観測データのうち最新の観測データを基にするのではなく、予測したデータを基に決定することが改善策として考えられる。

最後に、模擬講義で得られた映像に対しアンケートによる評価を行った。比較対象として、1つのカメラにより追跡撮影を行った映像と、3つのカメラそれぞれに対して、撮影対象の位置に対して重なりがないように担当範囲を決め、担当範囲内では追跡撮影を行い、撮影対象が別のカメラの担当範囲内に移動した場合にはそのカメラに切り替えて再び追跡撮影を行った映像とを用意した。提案手法による映像と合わせて3種類の映像を9人の被験者に提示した。それぞれの映像は約6分である。1つのカメラによる追跡撮影を「追跡撮影」、3つのカメラによる切り替えを伴った追跡撮影を「切替付き追跡撮影」として、アンケートの質問項目と平均点を表2に示す。アンケートは5点を最大とする5点法と自由記述によって行った。なお、追跡撮影についてはカメラ切り替えが存在しないため、質問5の回答はない。

No.	質問項目	追跡 撮影	切替付き 追跡撮影	提案 手法
1	見やすい映像ですか？	2.33	2.67	3.44
2	映像に不快感を持ちますか？	3.89	3.44	2.33
3	長時間見られる映像ですか？	1.78	2.67	4.11
4	人物の姿はよく見えますか？	4.11	4.11	3.33
5	カメラ切り替えは自然ですか？	N/A	3.78	2.78
6	人物の撮影にこのシステムを使いたいですか？	2.56	2.78	3.33

表 2 アンケート結果

提案手法は、質問2,3に対する回答で高い評価を受けている。また、追跡撮影、切替追跡撮影に対しては、映像が細かく動いて見づらいとの意見が多く見られた。このことから、固定ショット切り替え式撮影に視聴者の不快感を抑える効果があることが分かる。特に質問3に対する回答結果に大きな差があることから、講義撮影など長時間にわたって見る映像に対して大きな効果があることが分かる。

一方、質問4では他手法よりも低い評価となっている。これに関連して自由記述では、あるショットにおいて対象が画像の端に留まる時間が長かったことや、フレームアウトが気になったという意見が見られた。これは、カメラ数を増やすことや、

ショット内設定の評価について指定値からの差が大きくなるほど評価の悪化率を大きくする指標を導入したりすることで改善されると考えられる。また、対象の向きに関して重視して欲しいとの自由意見があり、向きに対する重みが一般の視聴にあたっては重視されていることも分かった。質問5で評価が低い理由として、特に短時間だけマスターショットに切り替わる状況が問題であると、自由意見で挙げられていた。これについては、マスターショットとの切り替えアルゴリズムを改良することが今後の課題である。

全体として、質問1に対する回答から提案手法での映像の見やすさが高評価を受けており、本研究のアプローチの正しさが立証されている。また、質問7に対する回答から、提案手法が人物の撮影に有意義であることが示された。

## 5. おわりに

本研究では事前に獲得した対象の運動履歴から、映像評価基準を満たす固定ショット撮影グラフを生成し、そのグラフを用いて複数カメラの固定ショット切り替えによる撮影を行う手法を提案した。シミュレーションおよび実制御実験でその有効性を確認した。

本手法では、固定ショット撮影グラフの生成は、ショット・ショット間・ショット構成の3つの映像評価基準に基づく評価関数に従って行われるので、人手による状況の理解やカメラ制御パラメータの設定は不要である。今後の課題としては、予想外の観測データが到来した時のマスターショットへの切り替え方法の改善などが挙げられる。

## 文 献

- [1] 石塚健太郎, “被写体の位置・音声レベルに基づく講義の自動映像化手法”, 京大大学院情報学研究所修士論文, 2000
- [2] 大西 正輝, 村上 昌史, 福永 邦雄, “状況理解と映像評価に基づく講義の知的自動撮影”, 電子情報通信学会論文誌 (D-II), vol.J85-D-II, no.4, pp.594-603, 2002
- [3] 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 -”, 電子情報通信学会技術報告 PRMU 2000-104, pp.31-38, 2000
- [4] 先山卓朗, 大野直樹, 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”, 電子情報通信学会論文誌 (D-II), Vol.J.84-D-II, no.2, pp.248-257, 2001
- [5] 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”, 情報処理学会研究報告 CVIM121-10, pp.73-80, 2000
- [6] S.A.Hutchinson, G.D.Hager, and P.I.Corke. A tutorial on visual servo control. *IEEE Trans. Robot. Automat.*, 12(5):651-670, 1996.
- [7] 足立淳, 佐藤淳 未校正カメラによる未校正ロボットの視覚サーボ 電子情報通信学会論文誌 D-II, Vol.J83-D-II, No.4, pp.1110-1118, 2000.
- [8] 東 和秀, 亀田 能成, 美濃 導彦, “パン・チルトス ビード及びズーム制御による動物体の自動撮影手法”, 第6回知能情報メディアシンポジウム予稿論文集, pp.101-106, 2000.
- [9] 新 康孝, 亀田 能成, 美濃 導彦, “指定カメラワーク下でのカメラ動作変更の抑制”, 電子情報通信学会研報 MVE, Vol.102, No.445, MVE2002-79, pp.69-74, 2002.
- [10] ダニエル・アリソン著, 岩本憲児, 出口丈人訳, “映画の文法”, 紀伊國屋書店, 1980.