

## シナリオ情報によるオーケストラ演奏のカメラワーク生成手法

井上 亮文<sup>†1,†2</sup> 平石 絢子<sup>†3</sup> 柴 貞行<sup>†4</sup>  
 市村 哲<sup>†2</sup> 重野 寛<sup>†5</sup>  
 岡田 謙一<sup>†5</sup> 松下 温<sup>†2</sup>

シナリオの存在するシーンを一般のユーザが撮影する場合、被写体の数が多かったり、撮影環境がそのつど異なったりするため、効果的なカメラワークを事前に計画することは困難である。そこで本論文ではシナリオ情報に基づいてカメラワークを自動的に計画する手法を提案する。撮影対象としてオーケストラ演奏を想定し、シナリオである楽譜から被写体の候補を決定する方法と、各カメラのショットを決定するための優先度の計算方法を定義している。実際にショットを接続して1本の映像を編集する実験の結果、提案手法はシナリオや優先度を考慮しないショットで編集したものよりも変化に富んだ映像を制作できることが分かった。また、カメラの配置に適応したショットを提示できることが分かった。

## Scenario Based Camerawork Planning for Orchestral Performances

AKIFUMI INOUE,<sup>†1,†2</sup> JUNKO HIRAISHI,<sup>†3</sup> SADAYUKI SHIBA,<sup>†4</sup>  
 SATOSHI ICHIMURA,<sup>†2</sup> HIROSHI SHIGENO,<sup>†5</sup> KEN-ICHI OKADA<sup>†5</sup>  
 and YUTAKA MATSUSHITA<sup>†2</sup>

Planning effective cameraworks for scenario-based scenes is difficult for average users, because there are a lot of subject to shoot and the deployments of cameras are different every time. In this paper, an automatic camerawork planning method for scenario-based scenes is proposed. The orchestral performance is assumed as an application. Then two features — (1) how to determine candidates for shooting subjects from the score and (2) the priority calculation to determine the shot of each camera — are defined. From the experimental results, the prototype system could provide necessary shots to produce interesting video picture. Approximately equivalent result was obtained even the camera deployment was different.

## 1. はじめに

高速インターネット環境とデジタルビデオカメラの普及により、自分で撮影した映像をネットワーク上で公開する個人レベルの映像配信が可能になった。これにともない、映像コンテンツ制作を支援する自動撮影技術への注目が高まっている。撮影対象は大きく分けて講義<sup>1)~3)</sup> やプレゼンテーション<sup>4)</sup>、スポーツ<sup>5)</sup> の

ようにその場の状況に応じて流れが決まるイベント型と、舞台<sup>6)</sup> や料理番組<sup>7)</sup> のようにあらかじめ進行が決まっているシナリオ型<sup>8)</sup> が存在するが、本研究では後者のシナリオ型に注目する。

シナリオ型における映像制作のプロセスは、いつ、どのカメラで、何を撮るかを決める計画フェーズ、その計画に従ってカメラを制御し撮影を行う実行フェーズ、撮影された映像を編集する編集フェーズの3つに分類することができる。実行フェーズではシナリオと実際のシーンの時間的・空間的ズレの修正<sup>9)</sup>、編集フェーズでは特定場面の検索機能<sup>10)</sup> などが必要となるが、良い映像を作るには、編集フェーズで必要とされる映像素材を確実に提供できるカメラワークを計画することが重要になる。

しかし、映像制作は一般的に難しく、シーンに多くの被写体が存在する場合、どのカメラで何を撮るのかを1つ1つ決めていくことは、映像知識のないユーザ

†1 慶應義塾大学大学院理工学研究科  
 Graduate School of Science and Technology, Keio University

†2 東京工科大学  
 Tokyo University of Technology

†3 株式会社エヌ・ティ・ティドコモ  
 NTT DoCoMo, Inc.

†4 株式会社 NTT データ  
 NTT Data Corporation

†5 慶應義塾大学理工学部  
 Faculty of Science and Technology, Keio University

にとって大きな負担である．また，テレビ番組が何十台ものカメラを用いて撮影し，その中から最高の表情や動きを選び抜いているのに対して<sup>11)</sup>，個人レベルの撮影ではカメラの台数が限られ，設置位置もそのつど異なることが多い．その結果，撮影計画がうまく練れていないと，編集フェーズで必要となるショットが撮影されていないといった問題が発生する．

そこで本研究では，シナリオからカメラワークを自動的に計画することを目的とする．撮影対象として，アマチュアレベルのオーケストラ演奏を想定する．オーケストラは，プロの交響楽団以外にも大学のサークルや地域の同好会など数多くのアマチュア団体が存在する．また，演奏は楽譜に基づいてあらかじめ進行が決まっており，シナリオ型のシーンとして適していると考えた．提案手法では，事前に楽譜から作成したシナリオから，多くのユーザの注目が集まる被写体を抽出する方法と，その中から優先度の高い被写体をカメラに割り当てる方法を実現している．これにより編集時に発生する様々なショット要求を満足し，撮影環境に応じたカメラワークの自動生成を目指す．

提案手法は，いい換えれば，シナリオで何を表現したいかというシナリオライタの知識，シーンの被写体をどのように撮影すればよいかというカメラマンの知識，どのような映像を作りたいかというビデオ編集者の知識を合わせて，演出が施された新しい映像作品を作り出すものである．

以下，2章では本論文が撮影対象とするオーケストラについて，3章では提案するカメラワーク生成方式について，4章ではプロトタイプを用いた実験方法について，5章では実験の結果および考察を述べ，6章を本論文のまとめとする．

## 2. オーケストラ

本章では，撮影対象であるオーケストラの特徴，撮影の現状と要求，映像表現について述べる．

### 2.1 特徴

オーケストラとは，様々な種類の楽器を指揮者のもとの大合奏すること，またはその団体を指す．オーケストラの代表的な楽器編成と略称を図1に，舞台配置を図2に示す．各楽器は木管楽器，金管楽器，打楽器，高弦楽器，低弦楽器という5つのグループから成り立ち，これらは演奏形態の違いから管打楽器，弦楽器にまとめられ，指揮者を含めて全体となる<sup>12),13)</sup>．各

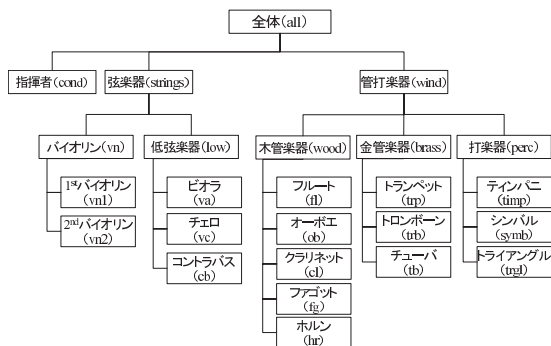


図1 オーケストラの編成例

Fig. 1 An example of orchestra's organization.

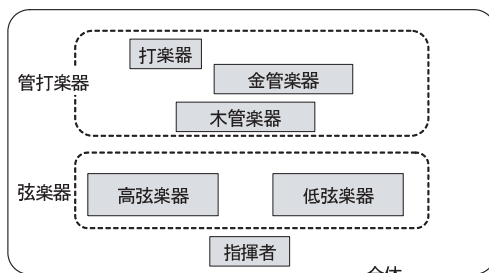


図2 オーケストラの配置例

Fig. 2 An example of orchestra's arrangement.

楽器はグループごとにまとめられた位置に配置される．

また，オーケストラには各楽器がどのような演奏をするかが音符で表現された楽譜（以下スコア）が存在する．これは「どこで，誰が，何を」するというイベントが時間軸に沿って書かれたシナリオとしての性質を持っている．

### 2.2 現状と要求

サークルや同好会などによるオーケストラ演奏の映像は，演奏者の記念や団体の紹介用として撮影されている．しかし，会場によってはカメラの設置位置が制限されており，つねに同じカメラ配置で撮影できるとは限らない．ある配置では正面から撮影できる演奏者は，別の配置では真横からのショットになってしまうこともある．これらを考慮して毎回カメラワークを計画するのは，映像編集に関する専門知識の少ないものにとっては大きな負担である．

また，現在は各シーンに適した被写体を1つに限定し，テレビ番組のような「お勤めの1本」にするカメラワークを計画している．つまり，全員が共通の1本の映像を視聴するよう編集されている．しかし，オーケストラには舞台上に複数の演奏者が存在し，同時に演奏をすることが多い．各演奏者からは，このようなお勤めの1本をベースにしながらも，自分の担当楽器

管楽器と打楽器は1つの旋律を1人で演奏するのに対し，弦楽器は同じ旋律を複数人で演奏する特徴がある．

の登場シーンを増やすよう編集したい、という要求がある。このような様々な編集要求にこたえるためには、各被写体に常時1台のカメラを割り当てておく必要があるが、個人レベルの撮影で用意できるカメラの台数は被写体である演奏者の数よりも少ない。

以上より本研究では、オーケストラ演奏の撮影において、映像知識の少ないユーザに対して、

- (1) 映像技法に基づいてお勤めのショットを確保する、
- (2) 編集フェーズで要求されるその他のショットを予測し、残されたカメラで撮影しておく、
- (3) 撮影環境の変動に対応する、

の3点を考慮して、シナリオからカメラワークを自動的に計画することを目標とする。なお、システムを完全に自動化するには、スコアからシナリオを自動的に生成する必要もあるが、これに関しては本論文では扱わない。

### 2.3 映像分析

オーケストラ撮影に適した映像技法を得るために、プロのカメラマンが撮影・編集したサンプル映像3本を分析した。いずれの映像もマイスタージンガー前奏曲の演奏シーン約10分(223小節)を撮影したものである。

映像分析の一般的な方法として、(1)ショットの内容に応じた分類、(2)ショット切替えのタイミング、(3)どの種類のショットからどの種類のショットへ移るかという1階の遷移、の3点がある<sup>14)</sup>。本研究ではこれに加え、(4)スコアとショットの関連性、についても分析を行った。

その結果、ショットの内容は図3に示す4種類に分類できた。「指揮者ショット」は演奏を指揮する指揮者を単独で映したものである。「メロディー楽器単独ショット」はフレーズの中心となるメロディーを演奏する楽器1種類を映したものである。「組合せショット」は音を出している楽器(演奏楽器)を複数まとめて映したものである。「全体ショット」はオーケストラ全体を映したものである。

次に、ショットの切替えは、メロディーのまとまりであるフレーズ単位で行っていることが分かった。図4にフレーズの例を示す。小節はスコアの定期的な時間の区切りであり、これを基準にするとメロディーの途中でショットが切り替わってしまうことがある。これに対しフレーズは、文章における節ないし文に相当する音楽的に意味を持ったまとまりであり、人間の心理に安定感を与える<sup>15)</sup>。その境界で映像を切替えることで無理のない映像にしていると考えられる。

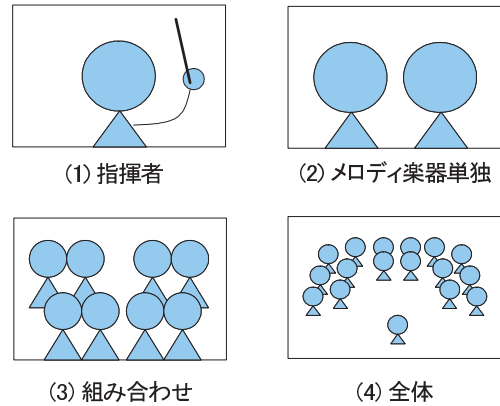


図3 オーケストラ映像におけるショット分類

Fig. 3 Shot categories in the orchestra video images.

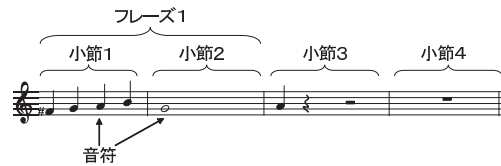


図4 フレーズの一例

Fig. 4 A phrase example.

ショットの1階の遷移では、切替えの前後でカメラのズーム値(ショットサイズ)に変化をつける傾向があった。同じ被写体のショットが続いたり、被写体を変えたとしても構図が似通ったりしていると退屈な印象を与えてしまう。アップショットとワイドショットを交互にしたり、視点を変えたりすることで映像に変化を与えていると考えられる。

最後に、スコアとショットは、演奏楽器の数と関連があることが分かった。ごく一部の楽器だけが演奏している「ソロ」状態(全楽器数の2割未満)では、ほとんどがメロディー楽器ショットであり、稀に指揮者のショットが挿入されていた。いくつかの楽器が演奏している状態では、メロディー楽器ショット、組合せショットが多く存在した。ほとんどの楽器が演奏している「大合奏」状態(全楽器数の6割以上)では、メロディー楽器ショット、組合せショットに加えて、指揮者または全体ショットが多く存在した。

### 3. 提案手法

本章では、2.3節の分析の結果得られた知見を利用したカメラワークの計画手法について述べる。提案手法の流れを以下に示す(図5)。

- (1) 最初にシナリオを一括して読み込む。
- (2) 読み込んだシナリオを解析する。
- (3) 注目が集まりそうな演奏楽器(被写体候補)を

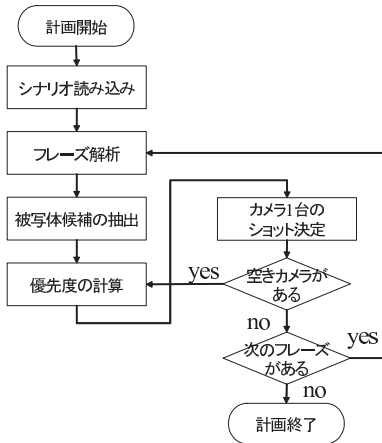


図 5 カメラワークの計画手法  
Fig. 5 Camerawork planning method.

推測する。

- (4) (3) の被写体候補の優先度を計算する。
- (5) 最も優先度の高い候補を、最も良く撮影可能なカメラに割り当てる。
- (6) 何も撮影していないカメラ（空きカメラ）がなくなるまで、(4) の優先度付けステップ以降を繰り返す。
- (7) 最後のフレーズまで、(2) からの処理を繰り返す。

ここで、以降で用いる用語について説明する。オーケストラにおける演奏者の数は、用意されるカメラの台数よりも多い。多くの楽器が同時に演奏した場合は、すべての演奏パートにカメラを割り当てることができない。そこで、シナリオで何らかの役割（演奏）を果たし、カメラが割り当てられる“可能性のあるもの”を「被写体候補」と呼ぶ。一方、カメラワークを計画した結果、実際にカメラが割り当てられたものを「被写体」と呼ぶこととする。

このように、提案手法はどのカメラがどの演奏楽器のショットを撮影するかを、先頭フレーズから1つずつ決定していく。なお、スコアからシナリオへの変換に関しては事前に手作業で作成しておくものとする。以降、計画の開始から終了までの各ステップの詳細を述べる。

### 3.1 シナリオの読み込み

スコアからシーンの演奏状況を把握するため、本研究ではXMLで次の2点を記述したものをシナリオとして扱うこととする。このシナリオをカメラワークの計画時に一括して読み込む。

#### 3.1.1 楽器の編成および位置

オーケストラでは、楽器の編成およびその配置は

```
<!ELEMENT stage_map (stage, orchestra)>
<!ELEMENT stage (name, w, h, d, sw, sh, sd)>
.....
<!ELEMENT fl (level, position)>
<!ELEMENT ob (level, position)>
<!ELEMENT cl (level, position)>
<!ELEMENT fg (level, position)>
.....
```

図 6 シナリオの DTD (舞台情報)  
Fig. 6 The DTD of the scenario (stage).

```
<!ELEMENT music (info, phrase*)>
<!ELEMENT info (title*, composer*, year*)>
<!ELEMENT title (#PCDATA)>
.....
<!ELEMENT phrase (no, start, end,
    main_part*, sub_part*)>
<!ELEMENT no (#PCDATA)>
<!ELEMENT start (#PCDATA)>
<!ELEMENT end (#PCDATA)>
<!ELEMENT main_part (#PCDATA)>
<!ELEMENT sub_part (#PCDATA)>
```

図 7 シナリオの DTD (フレーズ情報)  
Fig. 7 The DTD of the scenario (phrase).

楽曲によって決まる。そこで舞台の大きさ、楽器の編成および配置を記述しておく。これを記述した DTD (Document Type Definition) の一部を図 6 に示す。舞台の大きさなど (stage), 舞台上に存在する楽器名の略称 (fl, ob など), および三次元座標 (position) が記述されている。

#### 3.1.2 フレーズと各楽器の役割

カメラワークはフレーズ単位で計画されるため、シナリオにはフレーズに関する情報を記述しておく。DTDの一部を図 7 に示す。フレーズ番号 (no), 開始時刻 (start), 終了時刻 (end), メロディーを演奏している楽器のリスト (main\_part), 伴奏を演奏している楽器のリスト (sub\_part) が記述されている。

### 3.2 フレーズの解析と被写体候補の抽出

読み込んだシナリオを解析し、ユーザの注目が集まりやすい被写体候補を抽出する。この決定には、2.3 節の分析結果に基づき演奏楽器の数を利用する。フレーズ  $i (i = 1, 2, \dots, n)$  における演奏率  $E(i)$  は、シナリオから抽出可能な舞台上の全楽器数  $A$  と演奏楽器数  $I(i)$  から次のように求まる。

$$E(i) = \frac{I(i)}{A} \tag{1}$$

次に、フレーズ  $i$  における被写体候補  $H_i$  が  $a$  と  $b$  である場合を  $H_i = \{a, b\}$  と表すものとし、この  $H_i$

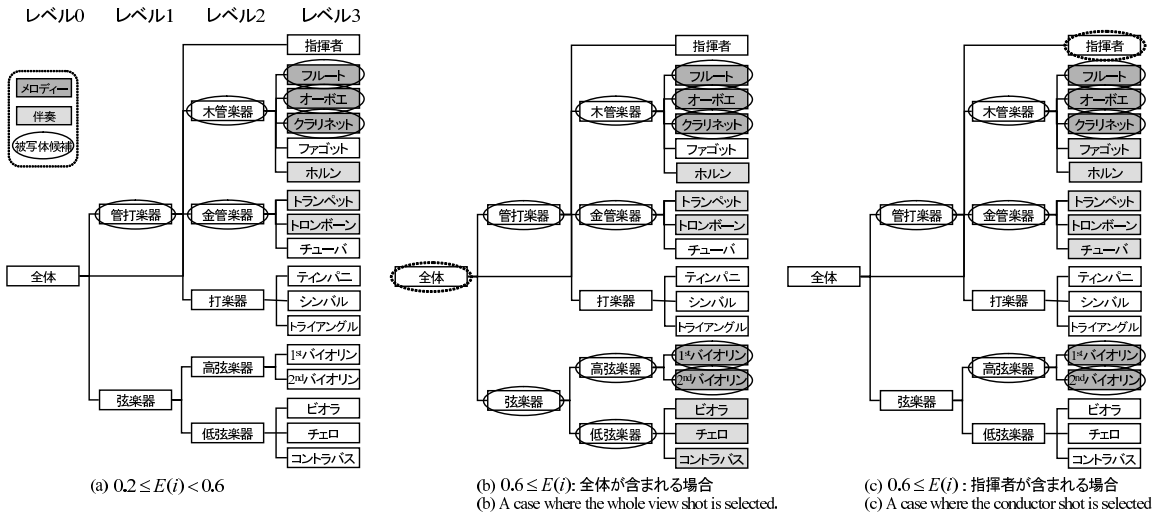


図 8 被写体候補の決定

と演奏率  $E(i)$  の関係を次のように定める .

$$H_i = \begin{cases} \{M_i, C\} & (E(i) < 0.2) \\ \{M_i, G_i\} & (0.2 \leq E(i) < 0.6) \\ \{M_i, G_i, C\} & (0.6 \leq E(i)) \\ \{M_i, G_i, W\} & (0.6 \leq E(i)) \end{cases} \quad (2)$$

ここで  $M_i$  はフレーズ  $i$  におけるメロディー楽器,  $G_i$  は演奏楽器の組合せ,  $C$  は指揮者,  $W$  はオーケストラ全体の各ショットを表す. 演奏率の境界値である 0.2, 0.6 という値は 2.3 節のビデオ分析から決定した .

この式 (2) より,  $E(i)$  が 2 割未満と小さい「ソロ」の場合は, メロディー楽器  $M_i$  と指揮者  $C$  を被写体候補とする .

$E(i)$  が 2 割以上 6 割未満の「いくつかの楽器が演奏している」場合は, メロディー楽器と演奏楽器の組合せを被写体候補とする . この組合せを作るために, 図 8 に示す楽器の階層関係を反映した 4 階層のツリー構造を定義する . そして, 同一階層のノードの過半数が演奏楽器である場合, その親ノードを「組合せ」として被写体候補に加えていく .

例として図 8(a) ではフルート・オーボエ・クラリネットがメロディーを, ホルン・トランペット・トロンボーンが伴奏を担当している (演奏率 0.38) . 式 (2) より, まずメロディー楽器であるフルート, オーボエ, クラリネットのショット 3 つが被写体候補になる . このとき, 木管楽器ノードに属する子ノードの過半数がアクティブであるため, 木管楽器全体を映すショットを「組合せ」として被写体候補にする . 金管楽器ノード

に属する子ノードの過半数もアクティブであるため, 金管楽器全体を映すショットが同様に被写体候補になる . さらに, 管打楽器の子ノードのうち 2 つがアクティブなので, 管打楽器全体を映すショットも加えて計 6 つが被写体候補となる .

$E(i)$  が 6 割以上の「大合奏」の場合は, 組合せの作り方によって指揮者ショットが全体ショットを被写体候補とする .

例として図 8(b) はフルート・オーボエ・クラリネット・1st バイオリン・2nd バイオリンがメロディーを, ホルン・トランペット・トロンボーン・ピオラ・チェロ・コントラバスが伴奏を担当している (演奏率 0.69) . まずメロディー楽器の 5 つが被写体候補になる . さらに子ノードの状態から, 木管楽器・金管楽器・高弦楽器・低弦楽器・管打楽器・弦楽器の 6 つが被写体候補となる . ここで管打楽器と弦楽器がアクティブであるため, その親ノードである全体を含めた計 12 個が被写体候補となる .

一方図 8(c) (演奏率 0.63) では, まずメロディー楽器の 5 つ, 組合せとして 4 つが被写体候補になる . しかし全体の子ノードのうちアクティブなのは管打楽器だけであり, 過半数に満たない . この場合は指揮者を被写体候補に加えることとし, 計 10 個が被写体候補となる .

このように, カメラを割り当てる可能性がある楽器の決定はフレーズ  $i$  の演奏率のみが影響する . フレーズ  $i$  以前, 以降の事象は影響しない .

### 3.3 優先度の計算

式 (2) で決定した被写体候補の数がカメラの台数よ

り多い場合、どの候補をカメラに割り当てるのか取捨選択する必要がある。本研究では、各被写体候補に対し優先度を計算し、この値に応じてこの判断を行う。現在のフレーズ  $i$  における被写体候補  $x$  の優先度  $P(i, x)$  の計算方法を式 (3) に示す。

$$P(i, x) = -\alpha F_p(i, x) - \beta F_f(i, x) - \gamma C(i, x) + \delta D_p(i, x) + \epsilon D_c(i, x) \quad (3)$$

ここで、 $\alpha, \beta, \gamma, \delta, \epsilon$  は楽曲や楽器の構成によって値が異なる重みを表す正の定数である。以降は  $F_p, F_f, C, D_p, D_c$  の詳細について述べる。

### 3.3.1 出現頻度 $F_p, F_f$

候補  $x$  がフレーズ  $i (i = 1, 2, \dots, n)$  までにカメラが割り当てられた頻度  $F_p(i, x)$  を計算し、その大きさに応じて優先度を  $\alpha F_p(i, x)$  マイナスする。同様に、シナリオを先読みすることでフレーズ  $i+1$  以降に被写体候補になる頻度  $F_f(i, x)$  を計算し、優先度を  $\beta F_f$  マイナスする。 $F_f(i, x)$  および  $F_p(i, x)$  はそれぞれ次の式 (4), (5) で表すことができる。

$$F_p(i, x) = \sum_{k=1}^{i-1} s_{k,x} \quad (4)$$

$$F_f(i, x) = \sum_{k=i+1}^n c_{k,x} \quad (5)$$

ただし

$$s_{k,x} = \begin{cases} 1 & (x \text{ が被写体のとき}) \\ 0 & (\text{それ以外}) \end{cases} \quad (6)$$

$$c_{k,x} = \begin{cases} 1 & (x \text{ が被写体候補のとき}) \\ 0 & (\text{それ以外}) \end{cases} \quad (7)$$

であり、 $F_p(0, x), F_f(n, x)$  は 0 とする。

$F_p(i, x)$  が大きい候補は、過去に何度もカメラが割り当てられ、多くのショットが撮影されている。フレーズ  $i$  でこのような候補にカメラを割り当てると、同じようなショットばかりになってしまう。また、 $F_f(i, x)$  が大きい候補ほど、フレーズ  $i$  以降の未来で被写体候補になる回数が多く、結果としてカメラが割り当てられる可能性が高い。2.2 節で述べたように、提案手法はなるべく多くの編集要求を満足することを目的とするため、このような候補の優先度を下げることで登場機会の少ない候補を撮影するようにする。

$F_p(i, x)$  は被写体に関連する値なので、直前フレーズまでのカメラワークが決定した後でなければ計算

できない。よってフレーズごとに逐次求める必要がある。これに対して  $F_f(i, x)$  は被写体候補に関連する値なので、カメラワークが決定している必要はない。ステップ (1) でシナリオを一括して読み込んだ時点で、式 (2) を用いてすべて計算することができる。

### 3.3.2 前後フレーズにおける類似度 $C$

候補  $x$  がフレーズ  $i-1$ 、つまり直前フレーズにおいてカメラに割り当てられたショットと類似性がある場合は優先度を下げる。この類似度の判定には図 8 の階層構造を利用する。

例として、フレーズ  $i-1$  において 3 台のカメラがフルート、金管楽器、ティンパニのショットを撮影していたとする。一方フレーズ  $i$  ではトランペット、指揮者、1st バイオリン、フルートの 4 つが被写体候補にあがっているとする。この場合、フレーズ  $i$  のトランペットは直前の金管楽器と図 8 において親子関係にある。2.1 節に述べたように、オーケストラでは各楽器がグループごとにまとめて配置されているため、金管楽器の中にはトランペットが映り、ショットに類似性があるといえる。このような場合、類似度  $C(i, x) = 1.0$  とする。一方、フルートは直前のフレーズでも撮影されており、親子関係にある場合よりもさらに類似性が高い。この場合は類似度  $C(i, x) = 2.0$  とする。

同様に、シナリオを先読みし、直後のフレーズ  $i+1$  における類似性も判定する。フレーズ  $i+1$  では、まだどの候補をどのカメラに割り当てるか決定していないため、自分と同じ、もしくは親子関係にある候補があれば、上記と同様の類似度に従って優先度をマイナスする。

映像文法上、同じ被写体のショットが続くと退屈でつまらないものになるといわれる。この類似度に応じた優先度  $\gamma C(i, x)$  をマイナスすることで、なるべく異なる種類の楽器が撮影されるようにする。

### 3.3.3 ショットサイズの差 $D_p, D_c$

2.3 節の分析より、フレーズの前後ではショットサイズを変化させる傾向があることが分かっている。そこでショットサイズの差に基づく優先度を計算する。ここでショットサイズの差は、図 8 におけるレベル値の差で表現する。たとえば「全体 (レベル 0)」と「金管楽器 (レベル 2)」の差は 2 である。

まず、直前フレーズとの差を考える (図 9)。フレーズ  $i$  における被写体候補  $x$  と、フレーズ  $i-1$  におけるすべてのショットとのサイズ差  $D_p(i, x)$  は次の式で表される。

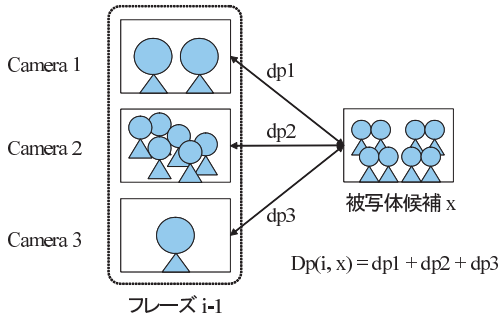


図 9 フレーズ間のショットサイズの差

Fig. 9 Shot size distance among phrases.

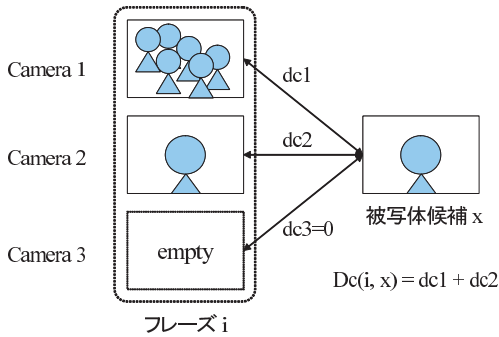


図 10 カメラ間のショットサイズの差

Fig. 10 Shot size distance among cameras.

$$D_p(i, x) = \sum_{j=1}^{N_c} |L(x) - L(S(i-1, j))| \quad (8)$$

ここで  $N_c$  はカメラの台数,  $S(i, j)$  はフレーズ  $i$  におけるカメラ  $j$  のショット,  $L(S)$  はショット  $S$  のレベル値を表す. この値に応じた優先度  $\delta D_p$  をプラスすることで, 直前フレーズになかったショットサイズの候補, つまり映像に変化を与える演出効果の高い候補の優先度を上げることができる.

次に, 同一フレーズとの差を考える (図 10). 被写体候補  $x$  と, フレーズ  $i$  においてすでに決定したショットとのサイズ差  $D_c(i, x)$  は次の式で表される.

$$D_c(i, x) = \sum_{j=1}^{N_c} |L(x) - L(S(i, j))| \quad (9)$$

ただし  $S(i, j)$  が empty (カメラ  $j$  のショットが決定していない) の場合,  $L(x) - L(S(i, j)) = 0$  として計算する. この値に応じた優先度  $\epsilon D_c$  をプラスすることで, 同一フレーズ内で様々なサイズのショットを提供する. 同じサイズのショットは, 被写体が異なっても構図が似てしまうためである.

### 3.4 位置関係を考慮したショット決定

フレーズ  $i$  において, すべての被写体候補の優先

```
<!ELEMENT camera_map (camera*)>
<!ELEMENT camera (name, position, pan, tilt, vangle)>
<!ELEMENT name (#PCDATA)>
<!ELEMENT position (area, x, y, z)>
<!ELEMENT area (#PCDATA)>
<!ELEMENT x (#PCDATA)>
<!ELEMENT y (#PCDATA)>
<!ELEMENT z (#PCDATA)>
<!ELEMENT pan (#PCDATA)>
<!ATTLIST pan unit (degree | radian) "degree">
<!ELEMENT tilt (#PCDATA)>
<!ATTLIST tilt unit (degree | radian) "degree">
<!ELEMENT vangle (#PCDATA)>
<!ATTLIST vangle unit (degree | radian) "degree">
```

図 11 カメラマップの DTD

Fig. 11 The DTD of the camera map.

度計算が終わった後, その優先度に従い被写体候補をソートする. そして最も優先度の高い被写体候補を, 最も良く撮影可能なカメラに割り当てる.

この「映り具合」の判断には, 撮影環境におけるカメラの設置状況を記録したカメラマップを利用する. この DTD を図 11 に示す. カメラの三次元座標 (position) と, シナリオから抽出可能な演奏者の位置と向きを利用し, カメラから見た被写体の向きが正面 > 斜め > 横の順に「より良く撮影可能」と定義する. 被写体とカメラの位置関係を考慮することで, 撮影環境の違いを考慮したカメラワークを計画することができる. また, 背後からのショットなど映像表現上不適切なものを排除できる.

カメラと被写体のペアが 1 つ決定するごとに, ステップ (4) の優先度付け以降を繰り返し, すべてのカメラに被写体を割り当てるまで処理を繰り返す.

## 4. 実験

### 4.1 プロトタイプシステム

3 章の提案手法に基づいて, シナリオからカメラワークの計画を自動的に行うプロトタイプシステムを実装した. システム全景を図 12 に示す.

システムはまず, XML 形式で記述したシナリオとカメラマップを読み込む. 次にカメラワーク計画部において, 3 章の手法に従って, どのカメラでどの被写体を撮影するかを計画していく. 計画されたカメラワークは, TVML 変換部において TVML スクリプトに変換される. カメラワーク実行部では, 計画されたカメラワークの映像を楽曲の MIDI 再生と同期させながら, TVML による仮想空間の映像として確認することができる.

現段階では計画したカメラワークの映像を確認する

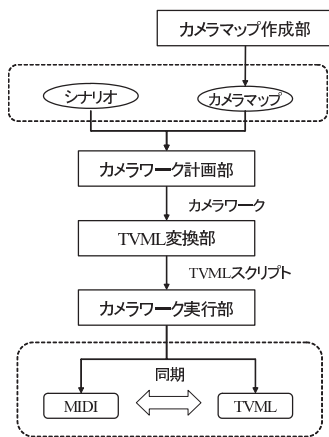
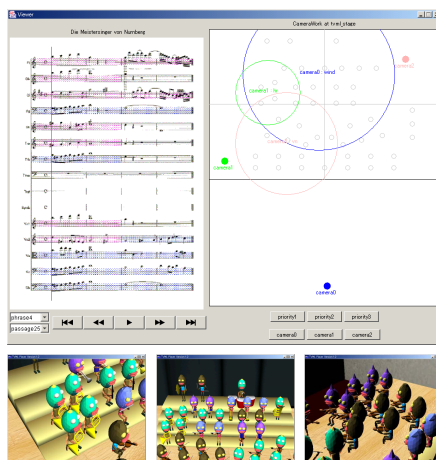


図 12 システム全景  
Fig. 12 System overview.



カメラ0:ホルン カメラ1:管楽器 カメラ2:高弦楽器

図 13 実装画面とショット例

Fig. 13 The camerawork executor and examples of planned shots.

ため、TVML1.2<sup>16)</sup> をシミュレータとして用いている。実際のカメラを用いて撮影すると、カメラの精度による実行時のズレの影響が大きくなり、本研究が目指す映像評価に支障が生じる可能性があるためである。ただし、実世界での利用を考慮し、カメラを任意の位置に移動させるといった仮想空間特有の機能はいっさい利用していない。

次に、システムの各機能の詳細について示す。シナリオは 3.1.1 項および 3.1.2 項に示した DTD に従い、事前に手作業で作成しておいた。カメラマップは、3.4 節に示した DTD に従って自動的に生成する GUI (カメラマップ作成部) を用意した。

カメラワーク実行部と、TVML によるショットの例を図 13 に示す。画面左側にはスコアが表示され、現在の演奏地点と各楽器の役割 (メロディー・伴奏) が表示されている。画面右側では、各カメラの配置と撮影範囲を確認することができる。

カメラマップ作成部、カメラワーク計画部、TVML 変換部は Java で作成した。カメラワーク実行部では、MIDI ファイルの再生と同期させるために、TVML プレイヤの外部制御 API を Visual C++ から利用した。

## 4.2 方法

### 4.2.1 被験者の選択傾向

マイスタージンガー前奏曲の中の連続した 10 シーン (約 2 分 30 秒) を取り出し、A, B の 2 種類の配置でカメラワークを計画した (図 14)。配置 A は 1 台のカメラを客席に、2 台を舞台上に設置し、配置 B では 3 台のカメラを客席に設置したものである。実験に用いた式 (3) のパラメータは、 $\alpha = \beta = 0.1, \gamma = 0.5, \delta = \epsilon = 0.25$  とした。これらは予備的な実験に基づいて決定した。

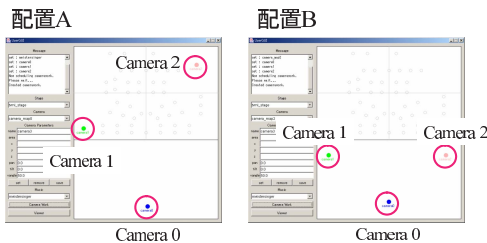


図 14 実験に用いたカメラ配置

Fig. 14 The camera arrangements for the experiment.

次に、A, B 両配置に共通し、どの演奏者も無理なく撮影可能な客席正面のカメラ 0 から、単独・組合せを含めた全 22 種類の被写体を撮影した。被験者 ( $N = 12$ 。全員音楽に関する知識を有する) には、各シーンに該当するスコアの一部を提示し、状況に合致していると思うもの 3 つを 22 種類の中から選択してもらった。この際、全体の構成などは考えずに判断するよう指示し、スコアも順序を入れ替えて段階的に提示した。このようにシナリオを意識せずに選ばれた被写体を調べ、提案手法で計画した被写体とどの程度一致するかを求めた。

### 4.2.2 映像編集

被験者に各シーンごとに 3 つのショットを提示し、その中から 1 つを選択して、1 本の映像作品を作ってもらった実験を行った。その際、被験者にはスコア全体を提示し、楽曲全体の構成や前後関係を考慮するよう指示した。まず、各配置で提案手法で計画したショット 3 つで実験を行った。さらに、4.2.1 項の実験において、各フレーズで被験者の支持を集めた被写体上位 3 つをランダムにカメラに割り当てたもの (以降、人



表 1 プロトタイプで計画した 3 ショットと人気ベースの上位 3 ショットのカメラへの割当て  
Table 1 Three shots by the prototype and three shots by poll.

scene	プロトタイプ A			プロトタイプ B			人気ベース AB		
	camera 0	camera 1	camera 2	camera 0	camera 1	camera 2	camera 0	camera 1	camera 2
1	all	cl	vn	all	wood	vc	all	wood	wind
2	low	timp	strings	timp	vn	strings	wood	strings	timp
3	wind	trp	wood	fg	wind	low	vn	brass	low
4	fl	cl	cond	fl	fl	cond	cond	cl	fl
5	va	ob	cond	cond	va	ob	low	ob	cond
6	wood	wood	cond	wood	wood	cond	fl	cond	wood
7	strings	hr	va	va	strings	vn	strings	fg	hr
8	vn	cond	vn	vn	vn	cond	vn1	cond	vn2
9	brass	symb	wind	symb	brass	wind	wind	wood	brass
10	all	fl	low	all	hr	low	vn	all	low

略称は図 1 を参照．N=12

気ベース)で同様の実験を行った．この 2 つを比較し，出来上がった映像にどのような特徴や違いがあるかを調べた．

## 5. 結果および考察

### 5.1 被験者の選択傾向との比較

4.2.1 項の結果として，プロトタイプで計画したショットと，人気ベースの上位 3 つの被写体を表 1 に示す．プロトタイプの被写体と，人気ベースの被写体が完全に一致したのは，配置 A におけるシーン 4 のみであった．それ以外のシーンでも，両配置とも，人気ベースの被写体と一致するショットが 1 つまたは 2 つ確保されているのが分かる．

プロトタイプで計画した被写体が，被験者のニーズとどれだけ一致しているかを見るには表 1 だけでは不十分である．そこで新たな評価尺度として一致率  $C_i$  を定義する．先の 4.2.1 項の実験で，シーン  $i$  で被験者が選択した 3 つの被写体には各 1 点を与えることとし，被験者数を  $N$ ，プロトタイプが計画した被写体すべての得点の合計を  $N_p$  とすると，一致率  $C_i$  は次の式で表される．

$$C_i = \frac{N_p}{3N} \times 100(\%) \quad (10)$$

各カメラ配置における一致率を表 2 に示す．この結果，シーン 3 が際立って低く，シーン 4, 5 が高くなった．

シーン 3 は，式 (2) における「演奏率 6 割以上」の状態であった．舞台上では様々な楽器が演奏しているため，被験者の注目する被写体が分散したためと考えられる．シーン内に複数の被写体が存在する場合，様々な編集要求があること，演奏率が高いほど被写体候補

表 2 各カメラ配置における一致率

Table 2 Shot corresponding ratio for each camera arrangement.

シーン	配置 A (%)	配置 B (%)
1	47	44
2	53	50
3	28	33
4	78	78
5	69	69
6	50	50
7	67	50
8	50	56
9	64	64
10	53	50

を増やしていくという提案手法の前提が確認できたと見える．

一方で，同様に演奏率が高かったシーン 1, 10 では約 50%で，シーン 3 と比較すると高い．この 2 つのシーンでは，提示したスコアから演奏の最初と最後であると被験者が判断したためか，全体を映したショットに投票が集中した．プロトタイプで計画した被写体の 1 つも全体であったため，一致率はそれほど下がらなかった．

シーン 4, 5 は，演奏率が 2 割未満の状態であった．舞台上で演奏する楽器もほとんどなく，注目が集まる被写体が限定されたためだといえる．シーン 8 も演奏率が 2 割未満であったが，シーン 4 より低くなっている．これはプロトタイプで計画した被写体が 2 種類だけだった (表 1 参照) が影響している．同様のことはシーン 6 にもいえる．1 つの被写体を重複して選択することを認めていないため，この 2 シーン (6, 8) は一致率を正しく評価できず，除外して考えるべきである．

シーン 7 は，配置 A と B とで一致率に差があった．シーン 7 では木管楽器と弦楽器が演奏をしている．そのため，被験者の多くがホルンなど木管楽器系の被写

同一の被写体を重複して選択することは認めていないため，1 つの被写体の得点の最大値は  $N$  となる．

体を選択した。配置 A では弦楽器とともにホルンが計画されていたが、配置 B ではホルンの代わりにピオラが計画されたため値が低くなった。両配置ともシーン 6 までにホルンのショットがなく、直前・直後のシーンにおける 3 ショットが共通であること、同じスコアを利用しているため未来の優先度も同じであることを考えると、被写体とカメラの位置関係による可能性がある。

これ以外では両配置で一致率に大きな差はなく、各シーンで最も多く投票された被写体（1 位タイを含む）は、すべてプロトタイプで計画されたものと一致していた。また、一致率を評価できないシーン 6, 8 を除いた平均は配置 A で 57%、配置 B で 55%であった。このことから、提案手法は撮影環境の違いに対応しつつ、そのシーンで最もニーズのある被写体のほかに、何人かの被験者が選択する被写体を確保するようにカメラワークを計画しているといえる。

## 5.2 映像編集方法の分析

4.2.2 項の結果を以下に示す。人気ベースで選んだショットは、前後関係や全体の構成を意識して選ばれたわけではないため、映像編集を行う際になんらかの違いが出ると考えられる。そこで、以下の 4 点について分析した。

### 5.2.1 ショットサイズ

2.3 節の分析で述べたように、プロが撮影したオーケストラの映像は、ショットの 1 階遷移の際にサイズに変化をつける傾向がある。そこで、出来上がった映像で、図 8 における同一レベル値のショットを接続した回数を測定した。その結果、配置 A でプロトタイプが 2.3 回、人気ベースが 3.4 回、配置 B でプロトタイプが 2.4 回、人気ベースが 3.2 回となった。

提案手法では、シナリオからカメラ間距離・フレーズ間距離を計算し、ショットのサイズによる優先度を付加している。プロトタイプでは、連続したシーン 4 と 5 では演奏楽器の割合が 2 割未満のケースであったため、両配置ともに計画したショットはどれも同じ階層レベルのものであった（メロディー楽器および指揮者）。よって、実質 1.0 回はシステムによって強制されている。

これに対し人気ベースでは、A, B 両配置ともに、そのような強制箇所は存在しなかったにもかかわらず、同一レベルでの切替え回数が大きくなってしまった。このことから、ショットサイズによる優先度付けが映像表現の向上に一定の効果があるといえる。

### 5.2.2 カメラ切替え

ショットを接続する際、サイズだけでなく、カメラ

の位置を変えれば映像に変化をつけることができる。そこで、カメラの切替え回数を測定した。結果、配置 A でプロトタイプが平均 8.0 回、人気ベースが 6.8 回。配置 B でプロトタイプが 5.6 回、人気ベースが 5.6 回となった。

3.4 節にあるように、提案手法では被写体とカメラを結びつける際に位置関係を考慮する。配置 A では、プロトタイプが人気ベースを上回り、一定の効果が見られた。一方配置 B では大きな差は見られなかった。

配置 A では、ホールの様々な位置にカメラが配置されており、各カメラからの映り具合が大きく異なる。カメラ 1 と 2 は、カメラ 0 に比べて正面方向から撮影可能な被写体は少ないが、カメラ 1 は低弦楽器全般、カメラ 2 はバイオリンと指揮者を撮影するのに適している。これらの被写体を撮影していた場合に、一時的にカメラ 1, 2 へと切り替える被験者が多く見られた（シーン 5, 8）。このような特殊な配置では、位置による優先度が有効に機能し、撮影環境の変動に対応して適切な被写体を撮影できていることが分かる。

配置 B では、すべてのカメラが客席側にある。カメラ 1, 2 ともに、正面・斜め前方から撮影できる被写体は配置 A に比べて多く、カメラ自体を切り替えなくてもある程度の構図で撮影できる。配置 A に比べて切替え回数自体が少ないことからそれがうかがえる。このため、位置の優先度があまり大きく働かなかったものと考えられる。将来的には、現在の正面 > 斜め > 横という単純なものではなく、たとえば楽器の見え具合など位置より複雑な優先度を適用することで、B のような安定した配置でも差が出るように対応する必要があるだろう。

### 5.2.3 全体の構成

実験で用いた 10 シーンでは、打楽器の演奏機会が非常に少ない。トロンボーン、コントラバスは伴奏が多く、図 8 における 1 レベル上の管楽器・低弦楽器ショットで撮影されることが多い。これに対しティンパニはシーン 2、シンバルはシーン 9 でしか演奏をせず、打楽器ショットとしても撮影されない。

プロトタイプでは、A, B 両配置ともこの 2 つのショットを確実に撮影できていた。これに対し、人気ベースの手法でも、ティンパニは上位 3 つに入ったため撮影することができたものの、シンバルの順位は 4 番目であったために撮影されなかった。このことから、提案手法による過去 ( $F_p$ ) および未来 ( $F_f$ ) に対する優先度付けが有効に機能していることが分かる。

### 5.2.4 ショットの優先度

最後に、各シーンで選択されたショットの優先度の

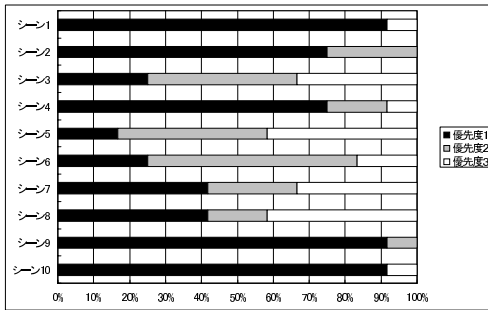


図 15 優先度の内訳 (配置 A)

Fig. 15 The priority distribution at arrangement A.

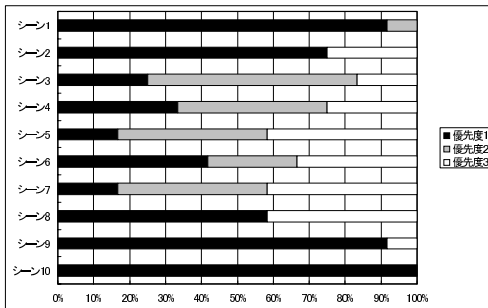


図 16 優先度の内訳 (配置 B)

Fig. 16 The priority distribution at arrangement B.

内訳を調べた。結果を図 15 および図 16 に示す。配置 A, B とともに、優先度の高いショットが積極的に選択されているシーンと、分散したシーンとに分かれた。

選択が集中したシーン 1 とシーン 10 は、優先度 1 のショットは「全体」であった。これらシーンは演奏の開始と終了という特殊なシーンであったため、選択が偏ったものと考えられる。また、シーン 2 とシーン 9 は 5.2.3 項で述べたティンパニとシンバルの演奏シーンであったため、これを選択した被験者が多かった。

それ以外の演奏の中盤で選択が分散した理由としては、被験者の経験や嗜好によるところが大きい。たとえば両配置のシーン 5 において、弦楽器の経験がある被験者はピアノを、木管楽器の経験がある被験者はオーボエを選択していた。2.2 節で述べたように、プロトタイプがお勧めの映像をベースにしなが、各自の嗜好にあわせて適宜差し替えるような形で利用されていることが分かる。将来的には、システムを利用するユーザのプロフィールを事前に収集しておき、これを優先度付けに利用するとよいと思われる。

## 6. おわりに

本論文では、オーケストラ演奏を撮影対象とし、シ

ナリオ情報に基づいたカメラワークの計画手法を提案した。提案手法では、シナリオからシーンの状況を把握して多くのユーザの注目を集める被写体の抽出方法と、優先度の高い被写体候補をカメラに割り当てていく方法を実現した。評価実験では、提案手法によって計画されたショットは、シナリオを意識せずに支持されたショットよりも変化に富む映像を制作できることを示した。また、一般的な配置とは異なり、舞台上にカメラを設置した場合でも、それに適したショットを提示できることが分かった。

提案手法はオーケストラを対象にしたが、各要素技術は様々なシナリオ型シーンに適用できると考えている。たとえばシナリオの進行役である司会がいて、様々な肩書きを持つ人たちが集まっている場面である結婚披露宴へ適用すれば、ショットサイズによる優先度付けで主役の新郎新婦を様々なサイズで撮影し、後の記念ビデオ編集作業で効果的なショットをたくさん確保しておくことができる。階層モデルによる被写体候補の作り方は、父、家族、親族一同、新郎側出席者など、様々な被写体のショットを撮影するのに役立つと考えられる。

今後の課題としては、今回は手動で行ったスコアからシナリオを作成する部分の自動化や、実機のカメラを用いた撮影実験があげられる。また、今回は比較的短いシナリオを想定していた。長いシナリオを用いた場合での優先度計算方法やパラメータの変更なども検討していく必要がある。

謝辞 本研究の一部は、21 世紀 COE プログラム研究拠点形成費補助金のもとに行われた。ここに記して謝意を表す。

## 参考文献

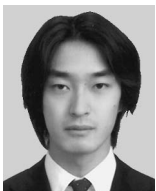
- 1) 大西正輝, 泉 正夫, 福永邦雄: 情報発生量の分布に基づく遠隔講義撮影の自動化, 電子情報通信学会論文誌, Vol.J82-D-II, No.2, pp.1590-1597 (1999).
- 2) 大西正輝, 村上昌史, 福永邦雄: 状況理解と映像評価に基づく講義の知的自動撮影, 電子情報通信学会論文誌, Vol.J85-D-II, No.4, pp.594-603 (2002).
- 3) 先山卓朗, 大野直樹, 棕木雅之, 池田克夫: 遠隔講義における講義状況に応じた送信映像選択, 電子情報通信学会論文誌, Vol.J84-D-II, No.2, pp.248-257 (2001).
- 4) Ozeki, M., Nakamura, Y. and Ohta, Y.: Camerawork For Intelligent Video Production — Capturing Desktop Manipulations, *International Conference on Multimedia and Expo*

(*ICME2001*), pp.41-44 (2001).

- 5) 井口泰典, 土居元紀, 眞鍋佳嗣, 千原國宏: スポーツ映像放送のための実時間映像解析によるマルチカメラの自動制御と自動スイッチング, 映像情報メディア学会誌, Vol.56, No.2, pp.271-279 (2002).
- 6) 灰塚凡樹, 井上誠喜: カメラワーク生成に関する一考察, 電子情報通信学会技術研究報告, PRMU96-202, pp.69-74 (1997).
- 7) Pinhanez, C. and Bobick, A.: Approximate World Models: Incorporating Qualitative and Linguistic Information into Vision Systems, *Proc. AAAI'96*, pp.1116-1123 (1996).
- 8) 住吉英樹, 有安香子, 望月祐一, 佐野雅則, 井上誠喜: 階層化した情報管理構造を用いた番組制作, 映像情報メディア学会誌, Vol.55, No.3, pp.397-404 (2001).
- 9) 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, CVIM-121-10, pp.73-80 (2000).
- 10) 秦 淑彦, 廣瀬竜男, 中西吉洋, 田中克己: カメラメタファによる多視点映像の検索, 情報処理学会論文誌: データベース, Vol.42, No.SIG 4 (TOD 9), pp.14-26 (2001).
- 11) 津田貴生, 加藤大一郎, 畠山祐里一, 井上誠喜: 番組制作とロボットカメラ, *NHK 技研 R & D*, 64, pp.4-9 (2000).
- 12) 満富俊郎: オーケストラとは何か, 新潮社 (1992).
- 13) 鈴木織衛: オーケストラを読む本, ヤマハミュージックメディア (2000).
- 14) 井上智雄, 岡田謙一, 松下 温: テレビ番組のカメラワークの知識に基づいた TV 会議システム, 情報処理学会論文誌, Vol.37, No.11, pp.2095-2104 (1996).
- 15) 岩宮眞一郎: 音楽と映像のマルチモーダル・コミュニケーション, 九州大学出版会 (2000).
- 16) 林 正樹: テキスト台本からの自動番組制作—TVML の提案, 情報処理学会研究報告, DBS-120-13, pp.91-98 (2000).

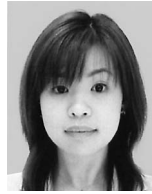
(平成 16 年 5 月 7 日受付)

(平成 16 年 11 月 1 日採録)



井上 亮文 (学生会員)

1999 年慶應義塾大学理工学部計測工学科卒業. 2001 年同大学院前期博士課程修了. 2004 年東京工科大学コンピュータサイエンス学部助手.



平石 絢子

2001 年慶應義塾大学理工学部情報工学科卒業. 2003 年同大学院前期博士課程修了. 現在株式会社 NTT ドコモに勤務.



柴 貞行

2002 年慶應義塾大学理工学部情報工学科卒業. 2004 年同大学院前期博士課程修了. 現在株式会社 NTT データに勤務.



市村 哲 (正会員)

1966 年生. 1989 年慶應義塾大学理工学部計測工学科卒業. 1994 年同大学院理工学研究科博士後期課程修了. 博士 (工学). 同年富士ゼロックス (株) 入社. 1997 年 ~ 1999 年富士ゼロックスパロアルト研究所駐在. 2002 年 4 月より東京工科大学助教授. グループウェア, ネットワークサービス, 生体情報活用等の研究に従事. 『IT TEXT 基礎 Web 技術』, 『IT TEXT 応用 Web 技術』 (オーム社). DICOMO 2003 優秀論文賞受賞. ACM, 電子情報通信学会各会員.



重野 寛 (正会員)

1990 年慶應義塾大学理工学部計測工学科卒業. 1997 年同大学院理工学研究科博士課程修了. 1998 年同大学理工学部情報工学科助手 (有期). 現在, 同大学理工学部情報工学科助教授. 工学博士. 計算機ネットワーク・プロトコル, モバイル・コンピューティング, マルチメディア・アプリケーション等の研究に従事. 著書 『~ ネットワーク・ユーザのための ~ 無線 LAN 技術講座』 (ソフト・リサーチ・センター), 『コンピュータネットワーク』 (オーム社) 等. 電子情報通信学会, IEEE, ACM 各会員.



岡田 謙一(フェロー)

慶應義塾大学理工学部情報工学科教授，工学博士。専門は，CSCW，グループウェア，ヒューマン・コンピュータ・インタラクション『ヒューマンコンピュータインタラクション』

(オーム社)、『コラボレーションとコミュニケーション』(共立出版)をはじめ著書多数。情報処理学会誌編集主査，論文誌編集主査，GW研究会主査等を歴任。現在，情報処理学会 GN 研究会運営委員，BCC 研究グループ幹事，日本 VR 学会 CS 研究会副委員長。情報処理学会論文賞(1996年，2001年)，情報処理学会40周年記念論文賞，日本 VR 学会サイバースペース研究賞，IEEE SAINT'04 最優秀論文賞を受賞。IEEE，ACM，電子情報通信学会，人工知能学会各会員。



松下 温(フェロー)

1963年慶應義塾大学工学部電気工学科卒業。1968年イリノイ大学大学院コンピュータサイエンス専攻修了。工学博士。1989年より2002年3月まで慶應義塾大学理工学部教授，2002年4月より東京工科大学教授および慶應義塾大学理工学部客員教授，2003年4月より東京工科大学コンピュータサイエンス学部学部長および慶應義塾大学理工学部客員教授。マルチメディア通信，コンピュータネットワーク，グループウェア等の研究に従事。情報処理学会理事，同学会副会長，マルチメディア通信と分散処理研究会委員長，グループウェア研究会委員長，電子情報通信学会情報ネットワーク研究会委員長，MIS研究会委員長，パーチャルリアリティ学会サイバースペースと仮想都市研究会委員長，情報処理学会 ITS 研究会委員長等を歴任。郵政省，通産省，建設省，農水省，都市基盤整備公団，行政情報システム研究所等の委員長，座長，委員等を多数歴任。特に国土交通省，住宅情報化標準策定委員会委員長，経済産業省総合エネルギー調査会電子計算機と磁気ディスク委員会委員長を務める。現在，経済産業省総合エネルギー調査会ルータ装置基準委員会委員長，最高裁判所専門委員。『やさしい LAN の知識』(オーム社)，『201x 年の世界』(共立出版)等著書多数。1993年情報処理学会ベストオーサ賞，1995年および2000年情報処理学会論文賞，2000年10月20日情報処理学会40周年記念90年代学会誌論文賞，2000年10月2日電子情報通信学会フェロー，2000年10月 VR 学会サイバースペース研究賞，2001年5月情報処理学会功績賞，電子情報通信学会，人工知能学会，ファジイ学会，IEEE，ACM 各会員。