

# 自由視点映像における映画的カメラ配置知識の適用

A Method to Apply Camera Work Knowledge on 3D-videos

坂本 竜基\*<sup>1</sup>      里見 美香\*<sup>1\*2</sup>      北原 格\*<sup>1\*3</sup>      土川 仁\*<sup>1</sup>      小暮 潔\*<sup>1</sup>  
Ryuuki Sakamoto      Mika Satomi      Itaru Kitahara      Megumu Tsuchikawa      Kiyoshi Kogure

\*<sup>1</sup>ATR 知能ロボティクス研究所      \*<sup>2</sup>情報科学芸術大学院大学  
ATR Intelligent Robotics and Communication Laboratories      Institute of Advanced Media Arts and Sciences

\*<sup>3</sup>筑波大学大学院システム情報工学研究科  
Univ. of Tsukuba

This paper describes a method to select a series of camera shots that is proper in the sense that it satisfies the grammar called "grammar of the film language". This method is assumed to apply to 3D free view point videos and the output video aims to record big moments in our daily life as cinematographic videos by setting virtual cameras' positions according to the grammar of film language and expertise. In this paper, we also show the prototype system implemented the method. To place a camera for making suitable shots using the proposed method, the system must know positions of the target objects/actors. Our solution to this problem is the annotation to note the positions for notifying the system.

## 1. はじめに

個人が日常的に得ている経験や体験をカメラやセンサでキャプチャリングしてこれを外部記憶する研究がおこなわれている[角 03, Gemmell 04]. このような体験のログデータは、後に本人や他人が閲覧するのが主な利用法であると考えられるが、例えば、環境カメラで撮影した映像を後で閲覧しても変化のない映像が多いため閲覧者の注意が持続せず、本当に必要な情報を見逃す恐れがある。

環境カメラの映像に比べて TV や映画等のそれは、シーンによってカメラの配置を変えたり動きをつけたりすることによって頻繁に変化がつけられている。このようなカメラの配置や動き(以下、ショットと呼ぶ)は、その選択によって映像の受け手に与える印象や効果が異なり、これらは専門知識としてある程度の範囲で形式化されている[Arijon 91]. ここではこのような撮影の知識を映画の文法と呼ぶ。映画の文法は、ショットが与える印象などの他にも、ショットとショットの繋がりといったショット間の制約条件についても形式化されていて、例えば二名を撮影する際はカメラの切り替えはこの二名を通る直線、イマジナリラインを跨いではならないといった条件がある。

体験をキャプチャリングする既存の環境において映画の文法に従った効果的な映像を撮影をしようと思うと、あらかじめその状況がおこることを見越してそれに応じた位置にカメラを配置する必要があるため、事実上不可能である。しかし、近年複数のビデオカメラの映像を統合して3次元空間を構築する仮想化現実の研究[Kanade 97, 北原 02]がおこなわれてきており、実際にはカメラがない位置に仮想的なカメラを配置してそこから撮影したような映像を作成することが可能となってきた。しかし、自由視点映像を作成する技術の研究に比べて、仮想カメラの配置方法といった演出や編集に関する応用研究はあまりなされていないのが現状である。

本稿では、実世界での出来事を効果的に映像表現するための手段として、複数台の環境カメラから作成した自由視点映像に対して映画の文法や専門家の撮影手法を適用する手法を提案

し、これを実装したプロトタイプシステムについて説明する。専門的な撮影知識を適用した映像は、環境カメラや個人視点からの生の映像では困難な映像による恣意的な状況説明やストーリー表現が可能であるため、既存の映像に比べて閲覧者に興味をもたせ、スムーズな情報伝達が可能であると考えられる。このような専門知識のカメラ配置を自由視点映像に対しておこなうためには、ショットの定義や配置ルールの定式化はもとより、そのショットを適用する時間と対象を特定する必要がある。プロトタイプでは、撮影対象物の位置と領域を指示するためのアノテーションを入力可能なシステムを導入した。

## 2. 自由視点映像

本手法で適用する自由視点映像とは、キャリアレーション済みの複数枚のビデオ映像を計算機内部で統合することにより、実際にはカメラが設置されていない位置から撮影した映像を仮想的に作り出す技術である(図1)。自由視点映像の生成・提示を実現する手法としては、ボクセルと呼ばれる立方格子の集合体により撮影対象空間を3次元配列として表現したデータ構造に対し、Shape-From-Silhouette法[Laurentini 94]を適用して、撮影物体の形状を構築するものや、物体が存在する位置にビルボードと呼ばれる平面を立て、それに環境カメラによって撮影された画像をテクスチャとしてを貼り付ける手法等、多数の技術が考案されている。其々の手法には、撮影環境や仮想カメラの位置等によって再現精度の面で有利不利が存在するが、本研究はカメラ配置知識の実現を主眼とするため、自由視点映像の構築に関しては深く言及せず、後述のプロトタイプシステムでは、最も基本的な自由視点映像作成手法の一つであるShape-From-Silhouette法を用いて3次元物体モデルを構築する。

## 3. 関連研究

計算機の高速化によって3Dモデルのレンダリングが高速におこなえるようになるのに伴って、人工的に作成された3Dアニメーションに対して知的なカメラワークを適用する研究がおこなわれてきた。これらの研究は、台本にあたる情報を元に3Dアニメーションを自動的に動かしつつ、適切なカメラワー

連絡先: 坂本 竜基, 国際電気通信基礎技術研究所, 住所: 〒619-0288 京都府相楽郡精華町光台 2-2-2 知能ロボティクス研究所, 電話: 0774-95-2553, Fax: 0774-95-1408

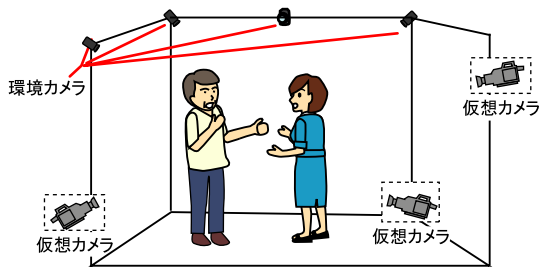


図 1: 自由視点映像の概念図

クをおこなうアプローチ [HAYASHI 99] と、あらかじめ決定されている 3D モデルの行動に対してカメラワークを適用するアプローチ [Bares 97, Halper 00] に分けられる。前者は、撮影対象である 3D モデルの動きそのものを制御可能なため、カメラと対象の相対的位置を都合よくとれるため、台本にあたる情報を作成しやすく、また精度の高いコントロールが可能である。一方、後者では台本の作成時に対象の位置取りや動きが未決定だとすると、対象の予定外の動きによって対象が画面上にうまく収まらないことや、遮蔽物などによって台本通りのカメラ配置が不可能な場合も考えられる。本稿がカメラ配置の対象としている実世界をキャプチャリングした自由視点映像は、対象が台本通りに行動しないため後者に近い。しかし、後者では対象物全体やパーツの位置や領域に関する情報及び、カメラが切り替わるべきトリガに関する情報が予め存在しているが、実世界を対象とするとこの情報の獲得も困難である。

一方、実世界の映像に対して映画の文法を適用する研究も始まっている。そのうち文献 [井上 04] では、テレビ会議システムにおいて映画の文法に則ったカメラ切り替えをおこなうシステムについて言及されている。この研究では、実カメラの切り替えをターゲットしているが、本研究では自由視点映像を対象としているため、実カメラでは困難なショットの適応が可能であり、より多彩な映像表現が可能である。

#### 4. 映画の文法に沿ったカメラ配置及び移動

本手法が扱う映画の文法は、ショットとショット同士の制約条件を基礎とする。ショットとは、映像全体からみれば断片的な映像であり、出力映像を完成させるにはショットを複数組み合わせる必要がある。このようなショットのシーケンスをシーンと呼ぶことにする。この際、このシーンを構成するショットは、それぞれのショット同士の制約条件を充足していなければならない。以下に、ショット情報と制約条件について説明する。

##### 4.1 ショット情報

ショット情報とは、撮影対象に対する適切なカメラパラメータを設定した初期状態と、初期状態からの移動やズームといった状態の変化という二つの情報から成るもので、カメラ制御に関する情報である。このうち前者を初期情報、後者を変化情報と呼ぶ。ショットによっては、ショットが終了するとカメラは初期位置とは別の場所に位置していることがあり、別のショットでは初期状態から全く変化しないこともある。

ショット情報における初期情報は、図 2 に示すような対象物体に対する角度とその物体が画像中に占める割合と位置が記述されている。新たにショット情報を追加する際に作成が簡便になるよう、双方の値とも数値での指定の他にエイリアスとなる予約語でも指定可能である。また、変化情報は初期情報からのカメラパラメータの変化が記述されているが、初期情報からの

時系列による相対的な変化量が記述される場合と、終了状態のカメラパラメータを設定して初期状態との変化途中を補間する場合がある。これらショット情報は、あらかじめ用意されて撮影知識ベースに保存される。

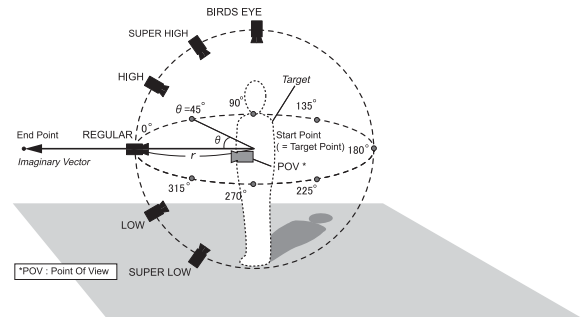


図 2: カメラ位置決定のための角度指定

##### 4.2 制約条件

制約条件とは、あるショットと次のショットとの間における禁忌のことであり、現在は、あるショットの終了時点の位置と次のショットの初期位置がイマジナリラインを跨ぐような組み合わせは受け手が混乱する恐れがあるため選択できないという制約が定められている。また、近いアングルのショットが連続しないような制約もかけられている。

##### 4.3 カメラ配置の適用

シーンは、ショット情報の集合を宣言することで作成する。シーンは、ショット情報のシーケンスとして保持され、各ショット情報は、ユーザが適合すると判断したショット情報があればその使用を宣言し、適合するショット情報がなければ、新しくショットを作成し、ショット情報として保存する。

このシーンを用いて実際に映像を出力する場合、内包するショット情報はそれ単体では機能せず対象物と対象時間を与える必要がある。また、有向なイマジナリラインであるイマジナリベクトルという仮想的なベクトル情報も与える必要がある。両者ともシーンを通じて共通することもあるし、個別に与える必要がある場合もある。対象物と対象時間は、後述するアノテーション付与インタフェースによって入力された注釈のインデックスを指定することによって与える。イマジナリベクトルは、ベクトルの始点と終点を同じくアノテーション付与インタフェースから入力することで与える。

#### 5. プロトタイプシステム

環境カメラから自由視点映像を作成し、本手法を用いて映像を出力するプロトタイプシステムを作成した。以下にシステム構成と出力例を示す。

##### 5.1 システム構成

システム構成は、撮影モジュール、自由視点映像作成モジュール、アノテーションモジュール、本稿で提案するカメラ配置をおこなうモジュールであるカメラワーク決定モジュールに大別される。図 3 にシステム構成を示す。

本システムにおいてショット情報は一つ一つ Java のクラスとして保存されており、新たなショット情報が必要になった場合は、既存クラスを継承したり、パラメータを変更したりすることも作成可能である。また、新たなクラスの中で複数の既存クラスのインスタンスを呼び出してカメラ位置を再計算する

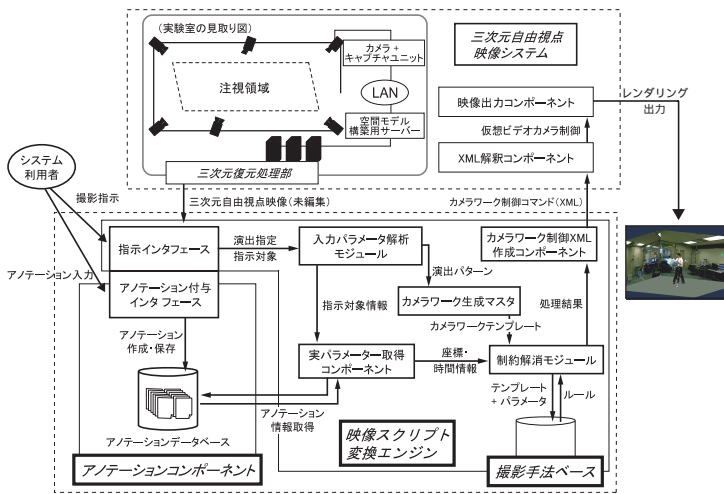


図 3: システム構成図

ことによって複数のクラスを擬似的に組み合わせたクラスも容易に作成可能である。

アノテーションは、ショット情報が実際にカメラを配置する際の対象を指示するために付与する。ショット情報が利用する対象とは、空間的对象と時間的对象に別れ、前者は対象の中心座標と領域という形で、後者はフレーム番号という形で、それぞれ名前であるインデックスが与えられ保存される。前者を空間的アノテーションと呼び、後者を時間的アノテーションと呼ぶ。

以上の情報をもったアノテーションを入力するインタフェースを作成した(図4)。空間的アノテーションは、マウスドラッグによって領域を入力し、時間的アノテーションは、タイムライン中の一点をクリックすることによって入力する。また、これらのアノテーションにインデックスと自由文記述で注釈をつけられる入力フォームを用意した。

すべてのアノテーションは人手で入力されるが、空間的对象はフレーム単位すべてに入力すること非常に手間がかかり、非現実的である。よって、異なるタイムラインに存在する同一インデックスの空間的对象に関しては線形補間することで入力の手間を省く機能をつけた。

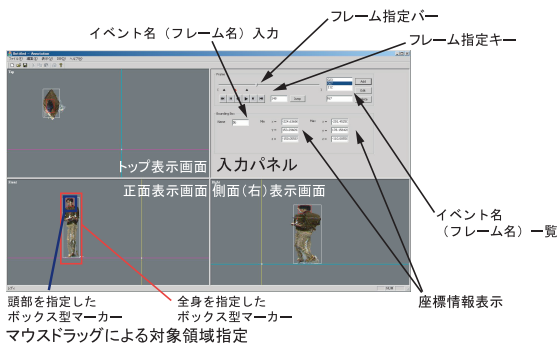


図 4: アノテーション付与インタフェース

本システムのプロトタイプを、約 27 平米の部屋に USB カメラとキャプチャユニット付き PC を接続したものを 6 セットから 8 セット配置した環境へと適用した。自由視点映像の生成処理は、毎秒 20 枚のフレームレートで撮影した映像を入力とし、約 2 倍の時間で実行されている。



図 5: 出力例

## 6. 出力例

プロトタイプシステムで作成された映像を図 5 に示す。撮影モジュールには、部屋の中を直線的に移動し、立ち止まって考え事をするという場面が撮影されており、自由視点映像作成モジュールがこの状況を自由視点映像化している。これに対して時間的アノテーションとして、開始時と対象の移動が終了する時点と終了時点にアノテーションをつけ、空間的アノテーションとして人物全身を内包するアノテーションと、顔を内包するアノテーションを付与した。

本映像では、2 つのショット情報を用いている。それぞれのショットの名前は「ズームアップドリー」「スピン・アラウンド」といい、ズームアップドリーは「ズームアップ」と「ドリー」という二つのショット情報を組み合わせたものである。ドリーとズームアップはシーンの切り替わり時や導入部など、人物が登場する場面によく利用され、スピン・アラウンドは、印象的なシーンで利用される撮影手法である。

図 5 の 1 では、後方からのカメラの直線移動であるドリーとカメラを対象物に向けてズームするズームアップが同時におこなわれている。このズームアップドリーは撮影対象が移動を停止するまで継続され、2 から 4 では、移動が終わった時点で撮影対象の周りを 1 回転するスピナラウンドに切り替わっている。

## 7. まとめ

本稿では、自由視点映像に対して効果的なカメラ配置をおこなう手法について提案した。また、自由視点映像の作成と、その映像に時間的、空間的アノテーションを付与するサブシステムを紹介し、撮影から映像出力までを実装したプロトタイプシステムについて述べた。

## 謝辞

本研究の機会を与えて頂いた ATR 知能ロボティクス研究所 萩田紀博所長に感謝する。また、研究に協力して頂いた北陸先端大の田中郁氏に感謝する。本研究は情報通信研究機構の委託研究により実施したものである。

## 参考文献

[Arijon 91] Arijon, D.: *Grammar of the Film Language*, Silman-James Press (1991)  
 [Bares 97] Bares, W. H. and Lester, J. C.: *Cinematographic User Models for Automated Realtime Cam-*

- era Control in Dynamic 3D Environments, pp. 215–230 (1997)
- [Gemmell 04] Gemmell, J., Williams, L., Wood, K., Lueder, R., and Bell, G.: Passive capture and ensuing issues for a personal lifetime store, in *CARPE'04: Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, pp. 48–55, New York, NY, USA (2004), ACM Press
- [Halper 00] Halper, N. and Oliver, P.: CamPlan: A Camera Planning Agent (2000)
- [HAYASHI 99] HAYASHI, M., UEDA, H., KURIHARA, T., and YASUMURA, M.: TVML (TV program making language) automatic TV program generation from text-based script (1999)
- [Kanade 97] Kanade, T., Rander, P., and Narayanan, P. J.: Virtualized Reality: Constructing Virtual Worlds from Real Scenes, *IEEE MultiMedia*, Vol. 4, No. 1, pp. 34–47 (1997)
- [Laurentini 94] Laurentini, A.: The Visual Hull Concept for Silhouette-Based Image Understanding, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 16, No. 2, pp. 150–162 (1994)
- [井上 04] 井上 亮文, 吉田 竜二, 平石 絢子, 重野 寛, 岡田 謙一, 松下 温: 映画の映像理論に基づく対面会議シーンの自動撮影手法, *情報処理学会論文誌*, Vol. 45, No. 1, pp. 212–221 (2004)
- [角 03] 角 康之, 伊藤 禎宣, 松口 哲也, Fels, S., 間瀬 健二: 協調的なインタラクションの記録と解釈, *情報処理学会論文誌*, Vol. 44, No. 11, pp. 2628–2637 (2003)
- [北原 02] 北原 格, 大田 友一: 大規模空間を対象とした自由視点映像生成のための3次元形状表現手法, *日本バーチャリアリティ学会論文誌*, Vol. 7, No. 2, pp. 177–184 (2002)