

## DMP : シナリオベースの個人用 CG 映画制作システム

宮崎 誠也 青木 輝勝 安田 浩

東京大学大学院工学系研究科

本稿では、誰もが容易に動画コンテンツを作成、発信を行えるようにするための技術として、DMP ( Digital Movie Producer ) を提案する。DMP の目標は、ユーザによる入力はシナリオ文章のみと簡便化し、コンピュータが人に代わって映画を自動的に制作することである。ここで映像中必要な CG 背景や登場人物、音声などの素材コンテンツはインターネット上から検索、取得することを前提とする。また、映画コンテンツ制作上有効なルールを持つ知識ベースを事前に準備し、カメラアングルやライティング等の設定を、シナリオを基に自動的に決定し、取得したコンテンツを組み合わせることで動画を作成する方式をとる。

本稿ではこの DMP システムの全体像とシナリオの XML スキーマを規定し、CG 動画の自動生成に必要な機能を考察する。

## DMP – an innovative personal Digital Movie Producer system

Seiya Miyazaki, Terumasa Aoki, Hiroshi Yasuda

School of Engineering, the University of Tokyo

In this paper, we propose an innovative personal CG movie production system called DMP (Digital Movie Producer) system, which generates a movie automatically from XML based scenarios and CG content from the Internet. In order to achieve this purpose, this system has some knowledge bases, with appropriate rules for producing movies. This system also has the sub function to search CG objects (studios, human characters, sounds, etc) from the Internet and Content Databases.

### 1. はじめに

FTTH(Fiber To The Home)に代表される超高速ネットワークがすみずみまで行き渡った IT 社会が実現する時代には、「マルチメディアコンテンツ流通」が主要なアプリケーションになると言われている。この時、マルチメディアコンテンツの受信のみならず、ユーザの誰もが容易にコンテンツを作成、発信できる環境を整えることがもっとも重要な要件になると考えられる。

このような背景のもと、本研究では CG を主体にした映画を誰でも容易に制作するための技術として DMP ( Digital Movie Producer ) を研究開発する。DMP の目標は、ユーザによる入力はシナリオ文章のみと簡便化し、コンピュータが人に代わって映画を自動的に制作

することである。ここで映像中必要な CG 背景や登場人物、音声などの素材コンテンツはインターネット上から検索、取得することを前提とする。DMP は映画コンテンツ制作上有効なルールを持つ知識ベースをあらかじめ有し、カメラアングルやライティング等の設定を、シナリオを基に自動的に決定することにより、取得したコンテンツを組み合わせることで動画を作成する。また意図しない動画や不自然な動きがある場合は、手動で修正できるインターフェースを備える。従来の CG 制作ソフトウェアでは、キャラクターの動きやカメラワーク等をひとつひとつ手作業で設定しなくてはならないが、このシステムでは、制作者はストーリー制作のみに注力して動画を作成することができるのが特徴である。DMP は個人制作の短時間映画の他に、日記や旅行記等

のパーソナル面での動画コンテンツ制作や、企業における動画プレゼンテーション等でも活用されることを想定している。

本稿では第 2 章で CG 動画制作ツールに関する研究を紹介し、第 3 章で DMP の全体構成の提案と研究課題を示す。続く第 4 章ではシナリオや動画スクリプトを XML で記述した XDMPML の提案、第 5 章では素材となるコンテンツの検索、取得に関する考察、第 6 章で動画の自動生成に関する考察と、カメラワークやライティングの自動化について述べる。

## 2. CG 動画制作ツールに関する従来研究

本章ではまず番組記述言語として TVML (TV Making Language) を紹介する。TVML[1] は TV 番組作成ツールとして、NHK 技術研究所で開発されたシナリオ記述言語である。あらかじめツール側で用意している CG キャラクターやセットを用い、番組内容を抽象度の高い表現(「タイトル#1」「～をズームイン」「ジャンプ」等)で記述することができる。また、TVML プレイヤはスクリプトを処理し、ニュース、イベントガイド、対談番組を中心としたテレビ番組をリアルタイムに出力することができる。しかしこの TVML のような番組記述言語により映画を制作する場合、登場人物、背景オブジェクトの挙動をすべて設定しなければならず、実際の制作は非常に煩雑で高度な技術が必要である。

次に動画の自動生成の研究例を紹介する。

矢部らの研究[2]では、ネットニュースの議論から脚本を作成し、セリフを CG キャラクターに割り当て、議論をテレビ番組化するツールを提案している。しかし、このツールは対象をネットニュースの議論に限定しており、また演出方法もパターン化しており、汎用用途には使用できない。

一方、服部、灘本、田中らは、Web ページの TV 番組化の研究[3]を行っている。これは Web 文書を解析し、TV 動画形式に自動変換することで、より内容を分かりやすくするという試みである。しかし、HTML で記述されている Web 文書を番組化する上で、文書中の平叙文を番組用対話文に変換することの問題や、満足した演出を自動で行えないなどの問題点があり、完全な自動化は現状では難しいと結論付けられている。またこれをふまえ、Web 文書からの TV 番組化を半自動化するオーサリング機構の提案や、Web 作成時に番組作成に用いる演出情報を記述可能な Scripting-XML(S-XML)[4]の提案により、TV 番組化を実現させている。S-XML では、S-XML コードから TVML コードへの変換を行う際に、XSLT スタイルシートのテンプレートを用い

て、その演出スタイルが指定できる。また「明るい雰囲気」など雰囲気を指定するタグや起承転結を表すタグがあり、すでに実装を行っている。

道家ら[5]は、TVML を用いて、ユーザがニュース番組に必要な内容と演出を設定するだけで、さまざまなニュース番組を自動生成する手法を考案した。

また橋本ら[6]は、番組インデックスを持つ映像から生成されるダイジェスト映像シーン集合を、TV 番組風に提示するシステムを提案している。このシステムには仮想キャラクター、映像、スーパーキャプションなどの異なる素材の同期を制御しながらダイジェストシーン間の接続表現を生成する機能がある。ただし実装では、対象は野球番組に限られており、番組全般を対象としているわけではない。

植田ら[7]の研究ではスポーツ中継の TV 映像から、パターン化されたシナリオテンプレートを使用してストーリー性を重視したダイジェストの自動生成を提案している。

上述のように、分野を限定した映像の自動生成や、オーサリングツールを併用する半自動生成は可能になってきている。しかし、映画のような自由度の高い動画を作成することを前提とすると、CG の動作や番組の表現力は依然としてユーザの満足いくものではないのが現状である。この問題を解決するには、汎用的かつ効率的な知識ルールと、表現力の高いコンテンツ記述能力が重要である。

## 3. DMP (Digital Movie Producer)概要

DMP の処理を図 1 に示す。まずユーザはテキストベースのシナリオを入力する。ここでシナリオとは、放送番組用のシナリオなどと同形式であり、タイトル、場所、人物の説明、セリフ、動作、カメラワークや音楽等その他特筆すべき事柄が形式に従って記述されているものとする。

続いて、補助システムによりシナリオの構造情報を用いてシナリオは XML の形式(XDMPML)に変換される。次にシステムはシナリオ情報を解析し、必要な背景素材やオブジェクトの素材、音楽、音声などをインターネット上、もしくはコンテンツ DB から取得する。ここで候補は一つに絞れないことが多いので、ユーザが候補を見て選択する。どのコンテンツを用いるかが決まれば、これをコンテンツ情報として XDMPML に追加する。

必要な素材を取得した後に、どのように自動的に動画を生成するかが大きな問題になる。シーンに適した効果的なショットを決定するためには、背景や登場人物の位置、構図、カメラワーク、ライティング、アクション、カ

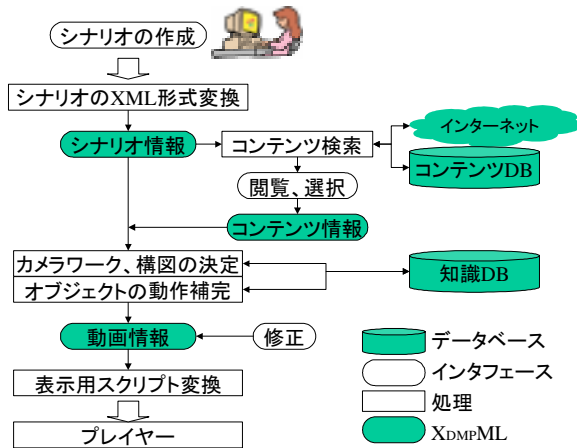


図1 システムの処理

ット割り、時間配分、分かりやすい表現構成など、映画作成にかかわる専門知識が必要である。さらにシナリオには省略された内容（演者の動きや表情など）や、抽象的な表現（「～のところに走りよる」など）があるため、知識ベースを参照しながら自動的に補完したり、制作者に問い合わせたりする必要がある。

また自動生成されたCG映画が必ずしも自分のイメージ通りとは限らないため、編集ツールが必要になるが、本研究では、シナリオ修正とマウスによる映像修正の2つのツールを用意する。

これにより動画生成に必要な完全な情報がXDMPMLに出力される。その後は、XDMPMLインタプリタにより動画出力するか、他の表示用スクリプト（TVML等）へのコンバーターを通して動画を再生する。

このように、DMP自体は素材となるコンテンツを作成するツールではなく、他のソフトウェアにより作成された素材コンテンツをシナリオと知識ベースに従って組み合わせることで動画を自動生成するシステムである。

#### 4. XDMPML

この章では、XDMPMLの言語仕様について簡単な解説を行う。XDMPMLドキュメント全体を表す構造木は図2のように表すことができる。すべての要素の先祖要素になるのがタグ属性<XDMPML>である。タグ要素<XDMPML>はその属性としてidを持つ。idは要素を識別するために用いられる識別値であり、ほとんどの要素に記載することができる。また、子要素は、テキスト文で記述されたシナリオに構造情報をタグにより付与したシナリオ情報、コンテンツの場所や初期化に用いる情報を記述したコンテンツ情報、動画を生成する上で必要な情報すべてを記述した動画情報の3つのパートに大きく分けられ、

それぞれ<Scenario><Content><Movie>のタグで表される。以下では、各要素に関して順次説明する。

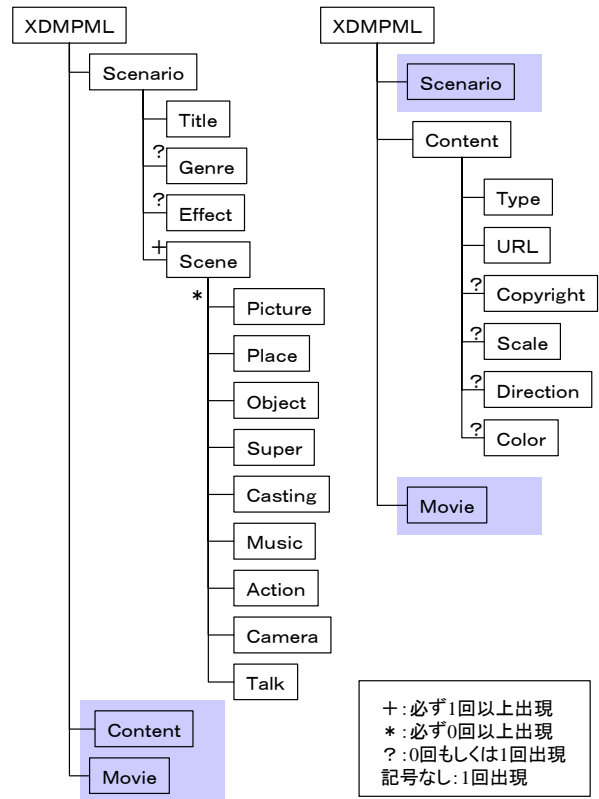


図2 XDMPML

#### 4.1. Scenario タグ要素

<Scenario>...</Scenario>で囲まれたタグ要素には、シナリオに書かれた抽象的な表現も含むテキストを、意味内容にしたがって構造化させた情報を記述することができる。

タグ要素<Title>にはシナリオのタイトル文が記述される。

またタグ要素<Genre>は番組のジャンル（地理紹介、ショートコント等）を記載できる。これは動画を生成するときの、カメラワークなどに影響する。例えば、地理紹介であれば背景を中心にしたカメラワークになる。

タグ要素<Effect>は番組の演出情報（ホラー、コミカル等）が入る。これは動画を生成するときの、ライティングや人物の動作などに大きく影響する。

<Scene>...</Scene>タグで囲まれたタグ要素には、シーン毎の情報が記述される。このタグは<Scenario>タグ内にシーンの数だけ記載される。

タグ要素<Picture>は、画面分割など画面構成に関する情報や、スローモーションなど動画の特殊な指定に用いる。

タグ要素<Place>は、シーンの場所、背景情報が記述される。例えば“セントラルパーク”“オフィス”などである。動画を生成するときは、このキーワードを用いてインターネットもしくはコンテンツDBから素材コンテンツを取得する。

またタグ要素<Object>は、シーンに出てくる小道具を指定する。

タグ要素<Super>は、映像に重ねる文字や画像（スーパーインポーズ）を指定する。

タグ要素<Casting>は、登場人物情報を記述し、タグ要素<Music>は、音楽や効果音を指定するのに用いる。

タグ要素<Action>は、登場人物の動きを記述できる。ここでは“画面にでてきて”などの抽象的な動作を記述してもかまわない。

タグ要素<Camera>は、カメラワークを記述する。例えば“～をズームイン”などである。

タグ要素<Talk>は、シナリオ中大部分を占める会話内容を記述するタグである。“早口で”“大声で”などの話し方の記述も可能である。

#### 4.2. Content タグ要素

<Content>...</Content>で囲まれたタグ要素には、先述のタグ要素<Scenario>で指定された素材となるコンテンツに関する詳細情報を記述する。

タグ要素<Type>は、素材となるコンテンツが背景、人物、小道具、音楽音声、図や絵のどれに属するか、また2Dコンテンツか3Dコンテンツか、またそのファイルフォーマット等を記述できる。タグ要素<URL>は、素材となるコンテンツのURLを記述する。

タグ要素<Copyright>は、素材コンテンツの著作権情報を記述する。色合いの変更など、部分的な改変を許可するかなどの情報も含まれる。また著作権情報を出力された映像に記述する時に用いられる。

タグ要素<Scale>は、コンテンツの大きさの初期値を変更するときに用いることができる。

タグ要素<Direction>は、コンテンツの向きを初期値を変更するときに用いることができる。例えば人物コンテンツであれば、初期値は上がY軸+方向でかつ正面がZ軸+方向になるように統一する。

タグ要素<Color>は、コンテンツの色の解像度や色合いを他のコンテンツと統一するときに用いることができる。

#### 4.3. Movie タグ要素

<Movie>...</Movie>で囲まれたタグ要素には、先述のタグ要素<Scenario>で指定されたシナリオ内容を、映像として再生する上で過不足ない具体的な情報を記述するのに用いられる。例えばシナリオ内には省略された動きやカメラワークなどすべての情報が記述される。この<Movie>タグ内容を参照して、XDMPMLインタプリタは動画を出力するか、他の表示用スクリプト（TVML等）へのコンバーターにより動画を再生する。

XDMPMLは、現在の仕様ではTVMLと相互変換可能なXML形式を採用している。例えば、TVMLでの以下のスクリプト表現（Guideという人物がx座標0.7のポジションに歩きながら、Acamカメラを見る）

```
character: walk( name=Guide, x=0.7, wait= no )
character: look( name=Guide, what=Acam)
```

は、XML形式で表現すると、

```
<Character name="Guide">
  <Par>
    <Walk x="0.7"/>
    <Look what="Acam"/>
  </Par>
</Character>
```

と相互に変換可能なXML形式でも表現できる。ここで<Par>...</Par>タグで囲まれた部分は同時に実行することを示し、一方<Seq>...</Seq>タグで囲まれた部分は逐次的に実行することを示している。

上述の仕様を用いて、例えば<Scenario>タグ内で記述された抽象的なシナリオの内容

```
<Action name="Guide">画面にでる</Action>
```

は、<Movie>タグ内では「画面右側の外に初期配置し、画面内に歩き手前に振り向く」という動作を意味する次の例のように詳細なパラメータを含む記述であらわされる。

```
<Character name="Guide">
  <Seq>
    <Position x="-0.3"/>
    <Walk x="0.4"/>
    <Turn d="20.0" speed="-3.0"/>
  </Seq>
</Character>
```

## 5. 素材コンテンツの検索と取得方法

素材コンテンツの検索キーとして、<Scenario>タグの情報を利用する。例えば背景コンテンツは<Place>タグの内容（「天橋立」など）を、人物コンテンツは<Casting>タグの内容（「バスガイド」など）を、小道具コンテンツは<Object>タグの内容を、音声音楽タグは<Music>タグの内容をそれぞれ参照する。このキーを利用してローカルDB、特定のコンテンツDBもしくはWeb上で公開されているコンテンツファイルの検索を行い、複数の候補を画面上に表示し、ユーザが選択取得できるようにする。

また素材コンテンツにはあらかじめメタデータが付与されており、名称、ファイル形式、コンテンツのタイプ、サイズや向き、アニメーションに必要な情報、構造情報が記述されているものとする。これらの情報は、動画を自動生成するとき用いられる。

## 6. 動画の自動生成

この章では、シナリオをもとに素材コンテンツを組み合わせ、適切な動画を自動生成するのに必要な機能を考察する。

### 6.1. 自動生成に必要な要素

素材となるコンテンツを取得した後に、シナリオのシーン情報を基に、まず構図を決定する必要がある。構図は、スタジオ内の人物や小道具（オブジェクト）の配置に影響する。例えば、人物の会話が中心の場面では、複数の人物を適切な距離をおいてスタジオ内に配置する。ここで「椅子に座る」等の指定がある場合、スタジオの椅子の場所がメタデータに記述されている場合はそれを利用し、記述がない場合は、椅子オブジェクトをスタジオの適切な位置に配置し、CG人物を座らせる。

構図を決定し、人物やオブジェクトの初期座標が決定されると、次は動作を生成する。ここで、シナリオ中には常識的な動きは省くことが多いので、知識ベースを参照して動作を補完する必要がある。例えば<Action>タグに「店の中に入る」という記述がある場合、人物はまず店の入り口前まで歩き、ドアノブをつかんでドアを開け、中に入るという一連の動作を自動的に生成する必要がある。またよりリアルな演出を行うには、一般的な常識を準備し、利用する必要がある。例えば、状況が雨の日の場合には、外では周りの歩行者は傘をさして歩くだろうし、夜という状況なら、街灯がついているなどの補完が必要となる。

構図と動作を決定したのち、カメラワークとライティングの詳細なパラメータを自動生成する必要がある。これらは次節で解説する。他にもタイミングや音楽、音声、スーパーなど、動画に必要な諸要素を<Scenario>タグの指定通りに設定をし、ポストプロセッシングを経て動画を生成する。この一連の処理の流れを図3に示す。

一方、シナリオにはあいまいさを多く含んでいるため、自動生成した構図や動きが意図したものとは異なる場合が多くある。この場合は、ユーザにシステムが問い合わせを行い、インタラクティブに修正できる機能が必要である。

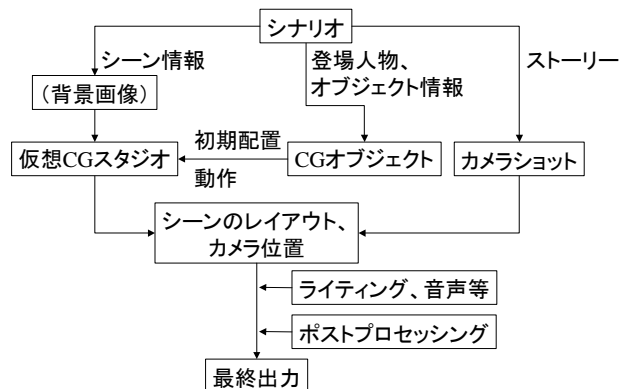


図3 フローチャート

### 6.2. 映画制作に関するノウハウ

カメラワークやライティングについては従来からある程度「定跡」となっている技法やノウハウが確立しており、例えばカメラショットには一つずつ名前がついていて、それぞれに特有の状況説明的な効果や心理的效果がある。

一例を挙げると、胸から頭までを入れる人物撮影時の標準サイズのバストショット、地形や空、天気、時刻などを示す効果のある、一帯の風景を映し出すロングショットなどがある。またカット割りは映像にメリハリを出す効果があるが、一方で多用すると視聴者を疲れさせてしまう。また画面中にまっすぐな線がエッジと並行もしくは垂直にないと画面に不必要な緊張感や不調和が生じるなどの、構図上のノウハウも多い。

これらの設定や演出は多くの知識と経験が必要とするため、素人には非常に難しい分野である。もちろんどの手法が最適なのかというルールはなく、実際映画などでは作成者の個性や感性、センスに従う部分が多い。しかしそれでも、視聴者が自然に感じる、素直なメッセージ伝達にはノウハウや基本的事項などがあるため、守るべき最低限のルールや演出方法を知識ベースに登録することで、常識

的なカメラワークやライティングを生成することができる。

### 6.3. カメラワークの自動化

シナリオに記述されているシーンの変化情報や人物の動き情報などを基に、カメラワークをシステムが自動生成するために、次の処理を行う。

まずシナリオからスタジオ内のオブジェクトの初期配置と最適なカメラショットの時間配分を決定する。ここで最適なカメラショットの決定方法は、カメラショットにおけるさまざまなノウハウやルールを準備し、それぞれに評価値を与え、最大の評価値をとるショットを選択することにする。このルールモデルは、視聴者が見たい部分や雰囲気、ショットの重要度などを踏まえ、視聴者を不必要に飽きさせたり疲れさせたりしない心理的な効果を考慮する必要がある[8][9]。例えば「シーンの変化を X<sub>DMPML</sub> の場所、背景タグが新たに指定される時として検出し、新しいシーンに変わったときは、まずワイドショットで周囲の環境を映す。」等のルールである。

ショットの種類を決定した後に、オブジェクトの大きさと動き、色合い等を考慮して、カメラの座標、傾き、ズームの度合い、レンズの種類、モーションパスなど詳細なカメラパラメータを決定する。例えば、話者のバストショットをするシーンであれば、話者 CG オブジェクトの位置と向き、大きさを取得し、他のオブジェクトの隠れないように胸から頭までが画面におさまるようカメラを移動させる。

### 6.4. ライティングの自動化

CGにおけるライティングには、

1. どのタイプの光源を
2. どのような属性で
3. どの場所に配置するか

を決定する必要がある。ライティングはシーンの雰囲気を決めるプロセスとして重要なステップである。光源の種類には、最もシンプルな点光源、注意を自然に惹きつけるスポットライト、太陽光などの無限遠光源、窓から屋内に光が差し込んでくる様子をシミュレートするのに良く使われる面光源、蛍光灯をシミュレートするのに使われる線光源、そしてシーン全体の明るさを決定する環境光がある。光源の対象との位置関係としては両サイド 45 度の斜光がもっとも基本的（標準光）であり、ドラマティックな演出の時のみ、下からの順光や後方上からの側面光を用いる。最終的に、演出方法やステージの形状、大きさ、登場人物の場所、話題の対象、全体の明るさ

などをパラメータとして、ライティングを決定する。

## 7. まとめと今後の課題

本稿では、個人用 CG 映画制作を非常に容易にするシステム - DMP について提案した。

DMP では、通常の CG 制作ソフトウェアと異なり、素材となるコンテンツ作成を対象とするものではなく、シナリオ文章を DMP が持つ知識ベースにより解釈し、カメラワークやライティングなど動画生成に必要な演出を自動化し、素材コンテンツを組み合わせで動画を生成する。

今後は試作を通し、問題点や要望点を検証し、より詳細な設計を行っていく予定である。最後に本研究を進めるにあたり、有益なご助言を頂いた放送作家の松岡剛様に感謝する。

### 参考文献

- [1] 林正樹ほか, “テレビ番組記述言語 TVML の言語仕様と CG 記述方法”, 第 3 回知能情報メディアシンポジウム, pp.75-80, 1997 年
- [2] 矢部純, 高橋伸, 柴山悦哉, “ニュースレッドからの番組自動生成”, IPSJ 第 85 回 HI 研究会
- [3] 服部多栄子, 灘本明代, 田中克己ほか, “番組メタファーによる Web ページの利用者適応型呈示方式”, 情報処理学会研究報告, Vol.99, No.61 99-DBS-119-69, 1999 年 7 月
- [4] 灘本明代, 服部多栄子, 田中克己ほか, “Web 情報の番組化のためのオーサリング機構”, 情報処理学会研究報告, Vol.2000, No.10 00-DBS-120-14, pp.99-106, 2000 年
- [5] 道家守, 林正樹, 牧野英二, “TVML を用いた番組情報からのニュース番組自動生成”, 映像情報メディア学会誌, No.7, pp.1097-1103, 2000 年
- [6] 橋本隆子ほか, “ダイジェスト映像シーンとマークアップ言語に基づく TV 番組生成システム”, 情報処理学会研究報告, Vol.42, No.SIG01 2000 年 1 月
- [7] 植田和憲, 鎌原淳三, 下條信司, 宮原秀夫, “シナリオテンプレートによるストーリー性を重視したダイジェスト生成機構”, 情報処理学会研究報告, Vol.99, No.61 99-DBS-119-70, 1999 年 7 月
- [8] コンプリート 3DCG, Isaac Victor Kerlow, MdN Corporation
- [9] Computer Graphics, James D. Foley ほか, オーム社