

# 概念辞書と結び付いた物語生成システムのための映像表現機構

小野 淳平<sup>†</sup>  
Junpei Ono

上田将礼<sup>‡</sup>  
Masayoshi Ueda

小方 孝<sup>‡</sup>  
Takashi Ogata

<sup>†</sup>岩手県立大学大学院ソフトウェア情報学研究科

<sup>‡</sup>岩手県立大学ソフトウェア情報学部

## 1. はじめに

筆者らは現在、これまでに行って来た物語生成システムに関連する諸研究を一旦集大成することを目的として、「統合物語生成システム(第一版)」の開発を進めている(小方・金井 2010; Akimoto & Ogata 2012). 統合物語生成システムは、出来事の生起時間順の並びを含む木構造を生成する物語内容機構、出来事の提示の仕方を変化させて物語内容の実際の語りの構造を構成する物語言説機構、映像・文字・音楽の媒体により表層的な表現を生成する物語表現機構の三つのフェーズを持つ。それに加えて、名詞概念・動詞概念及び修飾概念(形容詞概念及び形容動詞概念)をそれぞれ階層的に格納した概念辞書、断片的な物語コンテンツを格納した物語知識ベースを持つ。物語内容機構と物語言説機構は、この概念辞書や物語知識ベースを利用する物語技法によって、それぞれの構造を形成する。物語表現機構としては、上述のように、通常の物語生成システムで用いられている自然言語だけではなく、筆者らのシステムでは、物語内容や物語言説との構造的対応性に基づき構築された音楽自動生成機構(Akimoto, Endo & Ogata 2012)や、映像表現機構も用いる。

物語生成と関連した映像表現の研究としては、例えば Piacenza, et al.(2011)が Interactive Storytelling のための映像表現を報告している。このシステムでは、シェイクスピアの「ヴェニスの商人」を題材として、キャラクターの視点を通じた物語生成とその3D映像による表現が行われる。この種の物語世界を限定してアニメーション映画をモデルに精密で柔軟な映像表現を行う研究はこれまでも多く存在した(Bates 1992; 土佐・中津 1997).

これに対して統合物語生成システムの第一版では、上記システムのように物語の世界を限定することなく、映像表現においても、概念辞書に格納された数多くの概念要素に対応した映像素材を用意し、原始的で簡易なものであれどんな物語にも対応する映像表現が可能な機構の開発を目指している。すなわち、汎用的な物語生成能力に対応した映像表現機構を目指す。具体的には、統合物語生成システムで共通に利用する上述の概念辞書のうち、名詞概念辞書についてはその概念要素を静止画像と結び付け、一方動詞概念辞書の概念要素は動画像(アニメーション)と結び付ける。以下、本稿では、概念辞書について説明した後それらの機構を提案する。さらに、統合物語生成システムの中でこの機構を完成する前にそれを応用したシステムの中に試験的に組み込んだので、それについても述べる。

## 2. 概念辞書の概要

概念辞書(Oishi, et al. 2012)は、統合物語生成システムが物語の基本要素である事象もしくは事象概念(後述)を生成する際、その背景となる知識である動詞概念・名詞概念及び修飾概念の情報を提供する。ここでは名詞概念辞書と動詞概念辞書のみを例に説明する。

図1の右に示すように、名詞概念辞書の全体は is-a 関係による階層構造として構成される。終端概念は具体的な名詞概念であり、中間概念は複数の名詞概念を束ねる分類項目か、複数の分類項目をさらに束ねる上位の分類項目に相当する。現在、5808個の中間概念と115765個の終端概念を含む。一方動詞概念辞書は、図1の左に示すように、動詞概念を意味的に分類した階層を持ち、その最下層は個々の動詞概念に相当する。一つの動詞概念の記述は、自然文表現のための「文型パターン」、その深層格を定義する「格フレーム」、それぞれの格が取り得る名詞概念の範囲を制約しない規定する「制約条件」を持つ。制約条件には名詞概念辞

書中の中間概念が記述される。この辞書では、概念の名称は表層的な言語表現に準拠しているため、表層的な言語表現が同じであっても意味が異なる複数の動詞概念が存在する。その場合は、「食べる2」のように意味の異なる概念毎に番号を付与し、それぞれに異なる文型パターン・格フレーム・制約条件が記述する。図7の左下に示す「食べる2」の記述は、文型パターン「N1がN2を食べる」を持ち、N1は主格で中間概念「人」より下位の名詞概念を、N2は対象格で中間概念「食物」より下位の名詞概念を取り得る、ということの意味している。現在、12176個の動詞概念がすべてこのような形式で登録されている。

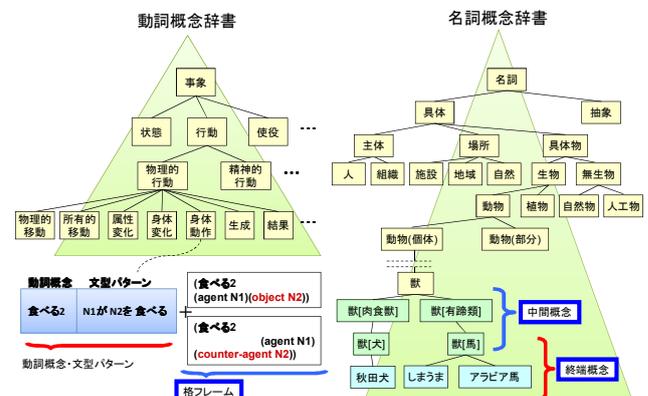


図1 動詞概念及び名詞概念の概念辞書の構成

以上のような動詞概念辞書に含まれる動詞概念にアニメーションを、名詞概念辞書に含まれる名詞概念に画像素材を結び付けることで、概念辞書をベースとする簡易な映像化機構を実現する。

## 3. 映像表現機構の構成と過程

### 3.1. 映像表現機構の概要

映像表現機構の役割は、前述の物語内容機構ないし物語言説機構で生成された物語の概念構造を映像表現に変換することである。物語の概念構造とは、図2に示すように、複数の事象概念どうしの関係からなる木構造により表現された物語の意味構造である。事象概念はその中で個々の出来事を表す要素であり、動詞概念ひとつとその格の値としてのひとつ以上の上の名詞概念インスタンスにより表現される。図3は事象概念の記述形式である。

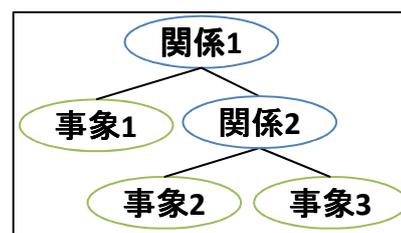


図2 物語の概念構造の構成

(event [動詞概念])	;事象
(type [種類])	;actionかhappening
(ID [事象ID])	;事象のID番号
(time [時間])	;事象の発生時間
(agent [インスタンスID])	;事象の行為主体
(counter-agent [インスタンスID])	;行為の客体
(object [インスタンスID])	;行為の対象物
(location [インスタンスID])	;事象の発生場所
(instrument [インスタンスID])	;行為に用いられる道具
(from [インスタンスID])	;事象の開始地点
(to [インスタンスID])	;事象の終了地点

図 3 事象概念の記述形式

図 4 に映像表現機構の概要を示す。大きく分けて、二種の処理機構（コード変換機構と出力ユーザインタフェース）と概念辞書を含む三種の知識ベースから構成される。映像化過程は二つに分かれ、一つ目の処理では物語の概念構造が映像コードに変換され、二つ目の処理ではこの映像コードの読み込みと実際の画面への出力が行われる。映像コードへの変換は Common Lisp で開発したコード変換機構が担い、実際の画像処理は HSP で開発した出力ユーザインタフェースが行う。この過程で概念辞書以外に利用する知識ベースとして、「画像素材知識ベース」と「アニメーション知識ベース」を用意した。前者は、人物・物・場所を表現する画像素材そのものの集合と、特定の画像素材と名詞概念辞書における特定の名詞概念とを関連付ける情報を格納する。一方アニメーション知識ベースは、動詞概念辞書における動詞概念の簡易なアニメーション表現の定義を格納する。

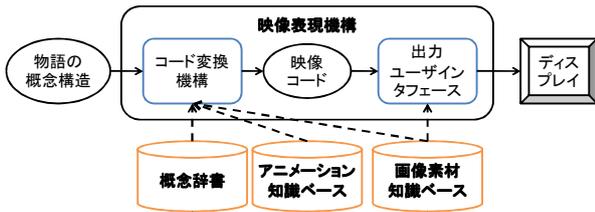


図 4 映像表現機構の構成

### 3.2. 物語の概念構造の変換過程

図 5 の例を用いて、物語の概念構造から映像コードへの変換処理並びに画像素材による映像化の方法を説明する。なおこの概念構造を文生成機構で簡易な単文（現在形）に変換すると次のようになる。「坊主が遠野南部神社から焼走りへ移動する。龍が焼走りで暴れる。坊主が焼走りで傷を負う。坊主が焼走りです勝利する。龍が焼走りから飛び去る。」

(\$継起	
(event 移動する1 (type action) (ID 1) (time (time1 time2)) (agent age%坊主#1) (location loc%遠野南部神社#1) (from loc%遠野南部神社#1) (to loc%焼走り#1))	
(\$継起	
(event 暴れる1 (type action) (ID 2) (time (time2 time3)) (agent age%龍#1) (location loc%焼走り#1))	
(event 負う1 (type action) (ID 3) (time (time3 time4)) (agent age%坊主#1) (location loc%焼走り#1) (object obj%傷#1))	
(event 勝利する1 (type action) (ID 4) (time (time5 time6)) (agent age%坊主#1) (location loc%焼走り#1))	
(event 飛び去る1 (type action) (ID 5) (time (time7 time8)) (agent age%龍#1) (location loc%焼走り#1) (from loc%焼走り#1))	

図 5 物語の概念構造の例

コード変換機構は、物語の概念構造内の事象概念を順番通りに映像コードに変換する。映像コードとは、画面に対する画像処理命令を記述したファイルであり、図 6 のように、一行につき一つの命令が記述される。例えば、「画像の水平方向への移動」を行う「mov」命令や、「背景の変更」を行う「ch\_BG」命令等がある。変換過程では、図 7 の変換ルールに基づき、事象概念単位で

映像コードへの変換を行う。

fnc,\$継起	利しました。、やきばしり
txt,坊主が遠野南部神社よ	でぼうずがりゅうにしよう
り焼走りに移動しました。	りました。
ぼうずがとおのなんぶじ	pos,30,200,350
んじゃよりやきばしりにい	txt,龍が焼走りから飛び去り
どうしました。	ました,りゅうがやきばし
ev,移動する,//////////	りからとびさりました。
pos,30,200,500	ev,飛び去る,//////////
mov,30,h,300,25,1	pos,32,200,500
go_action,//////////	mov,32,v,-50,-25,1
ch_BG,焼走り	mov,32,h,50,25,-1
... <中略>	... <後略>
txt,焼走りで坊主が龍に勝	

図 6 映像コードの一例

概念構造の『関係』は、構造を示す「fnc」に変換 事象概念は、agent格の値を人物・物の配置を表す「pos」 に変換 事象の文表現を示す「txt」を記述 動詞概念をアニメーションに変換
常に適用されるルール
一つ前の事象とlocation格の値が一致しない場合、背景変更 を表す「ch_BG」を記述 counter-agent格、object格、instrument格に値がある場 合、それぞれを人物・物の配置を表す「pos」に変換
適用条件が存在するルール

図 7 変換ルール

出力ユーザインタフェースは、映像コードを読み込み、一行ずつ命令の実行を行う。具体的には、名詞概念と対応する背景画像の上と同じく人物画像や物画像を重ねることで、全体の画面構成を作り出し、次に動詞概念に基づき人物や物のアニメーションをその画面上で動かす。図 8 は連続する映像表現の例で、図の左上の数字（実際には表示されない）の順に画面の映像が推移している。

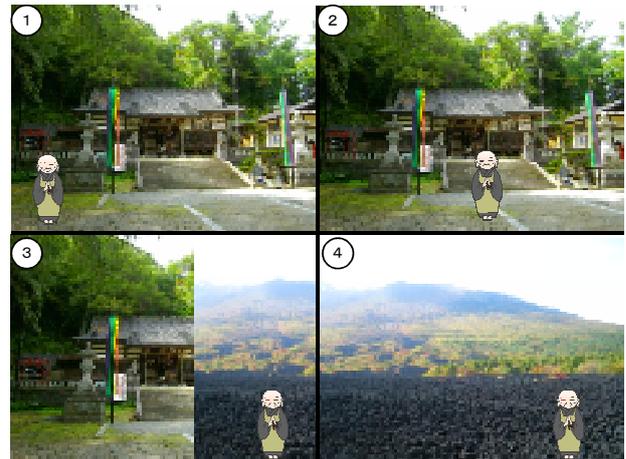


図 8 「坊主が遠野南部神社から焼走りへ移動する。」

## 4. 概念辞書と映像素材の対応付け

### 4.1. 名詞概念と画像素材

現状では、人・物・場所を表現する画像素材が、名詞概念辞書における「具体」範疇以下の 947 個の中間概念に対して用意されている。原則として、一つの中間概念に対して一つの画像素材を用意している。画像素材には、イラスト調のものと写実的なもの含まれ、前者は人や物の表現に使用し、後者は場所の表現に使

用することとしている。中間概念「場所」以下に含まれる名詞概念 130 個には写実的な画像を割り当て、それ以外にはイラスト調の画像を割り当てている。

## 4.2. 動詞概念とアニメーション

個々の動詞概念に対して、一つ以上の簡易な動きによってアニメーションを定義する。具体的な動きは、「画像の移動」、「画像の回転」、「画像の拡大・縮小」の三種類の画像処理タイプについて、表 1 のように五つのパラメータ（方向・変化量・変化の速度・加速量・画面に入出入りするかどうか）を用意し、それらの組み合わせから作成する。表 1 の“—”の部分は値が存在しない事を意味している。

表 1 アニメーションのための動きのパラメータ

	方向	画面の出入り	変化量	変化の速度	加速量
移動	上, 下, 左, 右	画面内, フレームイン, フレームアウト	小, 中, 大	遅, 並, 速	等速, 加速, 減速
回転	前, 後, 左, 右	—			
サイズ変更	拡大, 縮小	—			

特定の動詞概念とアニメーションの対応付けは、アニメーション知識ベース内で、「(移動する 1 (1 移動 A 中 右 並 等速 画面内))」のように記述している。これは、動詞概念「移動する 1」は、「主格が右水平方向に、並程度の速度で、加減速せずに、画面から見切れずに、中程度移動する」ことを意味している。これが映像コードに変換されると、「mov,1,h,300,25,0」という記述になる。これは「60fps で 1 フレーム毎に x 座標に対して +25 ピクセルずつ、300 ピクセル分まで人物 ID1 番の画像を動かす」ことを意味する。さらに、「移動する 1」は人物の現在位置を示す値を変化させる動詞概念であるため、上記の「mov」命令とは別に、「ch\_BG, 焼走り」という命令も映像コードに記述される。これは、「背景を焼走りの画像に変化させる」ことを意味しており、上述の「mov」命令と組み合わせることで、図 8 のような人物と背景が共に変化するアニメーションを含む映像表現が実現される。

上記の方法により、第一段階の作業として、動詞概念辞書の中概念「物理的行動」以下の 6264 個の動詞概念を対象にアニメーション定義を試み、そのうち 3078 個はアニメーションとの対応付けが可能であった。その結果について次に述べ考察する。

## 4.3. アニメーション作成から見た動詞概念の特徴の分類

動詞概念における「物理的行動」の範疇における、1. 身体全体、2. 物全体、3. 身体全体と物全体、これら三つのパターンの空間的な移動を伴う動詞概念、並びに、4. それぞれのパターンが反復する動詞概念、5. それぞれのパターンの組み合わせによる動詞概念の 5 種類からなる 3078 個の動詞概念に対しては、以上の方法によってアニメーションを定義した。しかし残りの物理的行動 3186 個に関しては以上の方法のままではアニメーション定義は不可能であった。その理由を考える中で、物理的行動の動詞概念の特徴を以下の三種類に分類した——(1) 詳細な動作への分割を必要とする抽象的な動詞概念、(2) 特定の物や場所に依存する意味を含む動詞概念、(3) 頭部や四肢等身体部位の動きを必要とする動詞概念。

### (1) 詳細な動作への分割を必要とする動詞概念

例えば「戦う」や「喧嘩する」のような動詞概念が物理的行動の範疇に含まれているが、それらの概念自体が直接動作を指定する訳ではない。それらは下位レベルの具体的な動作の集合に分割・詳細化されることによって初めて実現される抽象的な動作概念であり、またその抽象化の仕方にも多様性がある。

小方(2008)は、動詞的概念を「行為」と「動作」に分ける考えを議論している。動作が直接具体的に実現され得る動詞的概念であるのに対して、行為は動作の集合への詳細化によって実現され

る概念である。小方(2008)は図 9 の例を挙げている。「誘拐する」という行為としての動詞的概念は、「現れる」等の下位の行為に分割され、最終的に個々の最下位行為は「歩く」等の動作としての動詞的概念として実現される。実際は、両者の分割は総体的なものであり、何をもって動作とするかは、それを利用する目的によって可変的である。例えば、「歩く」を動作とするのではなく、それよりも詳細化された動きを動作とすれば、「歩く」の方はより抽象的な行為に相当する概念ということになる。このような考え方からすれば、本研究における現状でのアニメーションの定義はそのレベル設定が恣意的であり、例えば「現れる」のような動詞的概念も直接アニメーション化出来る動作として扱っている。今後より組織的に作業を進めるためには、ここでの動作に相当する基本的なアニメーションの動きを定義し、その組み合わせによって行為を実現するようにすることが必要である。真部・小方(2007)はもともと筆者らの物語生成システムとの共同研究において、複数の関節の動きを基本動作とするアニメーション定義の方法を提案しているが、当面の試作として動作を伴うすべての動詞概念のアニメーションを定義することを目指す本研究では、そのレベルを少し上げてその第一版の完成を目指している。

他方、動詞概念の具体的な実現方法の多様性の問題であるが、これはアニメーション表現の問題と言うより、物語生成におけるスクリプト的な事象展開と関連する問題である。行為と動作の分割との関わりでは、行為の詳細化は物語内容機構における物語の概念構造生成によって行われ、本アニメーション表現機構ではそれを具体的なイメージを伴う動作集合として実現する処理が行われる。今後、このような役割分担を明確にして行く。

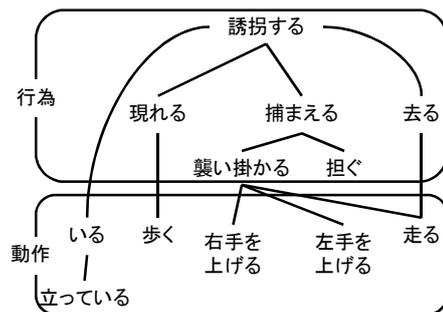


図 9 行為と動作の階層性の例(小方 2008)

### (2) 特定の物や場所に依存する意味を含む動詞概念

「乗車する」「航海する」のように動詞概念自体に特定の物や場所を指す意味が含まれているものである。既に、「乗る」のような動詞概念についてはアニメーション定義を行っているが、「乗車する」はそれに対して object 格が「乗り物」に限定されていると考えることが出来る。従って、この種の動詞概念については、「乗る」のようにより広い範囲の制約条件を持った動詞概念の特殊形として処理することが出来る。

### (3) 頭部や四肢等身体部位の動きを必要とする動詞概念

「蹴る」や「採取する」のように、頭部や四肢等体幹以外の部分の動きを含む動詞概念や、空間的な移動を伴わない動詞概念が存在する。現状の機構で実現している身体動作は身体全体の移動と頭部の動きであり、上記のような動詞概念にアニメーション定義の範囲を拡大するためには、可能な身体部位の動きを拡張しなければならない。これは上記(1)における、動詞概念の基本単位としての動作の集合をどのように作成するかという問題と関連する。すなわち、まず現状のような手順でボトムアップに作業を進めることを通じて、アニメーション定義として実現が必要な最小限の身体動作が明らかになる。それによって、現在アニメーション定義可能なものに分類されている動詞概念の定義の改訂(詳細化)を含めて、作業をより組織的なものにして行くことが可能であろう。

#### 4.4. 課題

名詞概念辞書に存在する中間概念及び終端概念の総数 120590 個のうち残りの 119643 個に対して画像素材を割り当てる予定である。そのために Web からの検索を予定している。動詞概念へのアニメーション定義の対象については、4.3 節の(1)で述べたように、動詞的概念を行為と動作に分けて処理する仕組みを整備することで、動作だけに絞りその数を大幅に減らせる可能性がある。

個々のアニメーションの質に関しては、定義したアニメーションを文表現と同時に確認して最低限認識可能かどうかを最初の段階の評価とする予定である。

現在は事象の連続性に関する考慮はされていないので、不自然な場合が生じる。事象概念毎にキャラクターの座標をリセットし、またキャラクターの画面上での座標位置や移動方向を定数で指定しているためである。連続する事象かどうかを判断し、その場合連続する事象間でのキャラクターの位置関係を考慮し、背景に制約的な情報を持たせる等のことで、座標や移動方向を適宜変更する処理を付加する予定である。

#### 5. 応用システム KOSERUBE での利用

「いわての民話 KOSERUBE」(今渕 他, 2012)と呼んでいるのは、岩手県に因んだ登場人物や場所や物品が登場する民話風物語の概念構造・文表現・音楽を自動生成し、さらに以上で述べた映像表現機構を利用して図 10 に示すような画面を表現するシステムである。KOSERUBE は、統合物語生成システムの名詞概念辞書に、「カッパ」「南部鉄器」「遠野南部神社」等岩手独自の要素として人物 34 種類、物 21 種類、背景 22 種類を追加しており、映像表現機構ではそれらの名詞概念に対して一つずつ画像素材を割り当てている。また KOSERUBE の主要な物語生成技法は、プロップの提唱した 31 個の「機能」に関する文学理論 (Propp1969) に基づいて筆者らが作成したストーリーグラマー (Imabuchi & Ogata 2012) に基づいており、アニメーション化される動詞概念はそこで使用されている 147 種類に絞られている。その多くは物理的行動であり上述の方法でアニメーション化されているが、その中の 41 個は「要求する 1」や「決意する 2」のような、動詞概念辞書における中間概念「精神的行動」の範疇に含まれるものであり、仮のアニメーション定義を割り当てている。精神的行動のアニメーション定義は本稿の内容を超える議論が必要であり、今後の課題とする。



図 10 上演中の物語 (左: 物語の最初, 右: 最後)

なお、KOSERUBE で生成された音楽と映像は、物語の概念構造をもとに同期する。音楽表現機構は、物語の概念構造をマクロなレベルで三つもしくは四つに分割し、分割した各部分毎に異なる手法で音楽を自動生成する。映像表現機構は、物語の概念構造を映像コードへ変換する際物語概念構造に関する情報もコードとして記述する。図 10 のように実際に物語を上演する段階で、出力ユーザインタフェースは、映像コードに記述された「fnc」命令により、現在上演中の内容が物語の概念構造のどの位置に当たるかをチェックしている。音楽生成の分割単位に当たる位置まで物語の上演が進んだ時その物語の概念構造に対応する音楽を再生することで、出力ユーザインタフェースは映像と音楽を同期させる。

#### 6. おわりに

統合物語生成システムの映像表現機構として、原始的で粗野ではあっても、その名詞概念体系と動詞概念体系における概念要素

と対応付けられた画像素材とアニメーション定義によって、物語世界を限定しない物語生成に対応した汎用的な映像表現を可能とする機構の開発を目指している。現状で、名詞概念辞書中の名詞概念 947 個に対応する人物・物・場所を表現する画像素材を収集し、動詞概念辞書に対しては「物理的行動」に含まれる動詞概念 3078 個のアニメーション定義を作成した。また、上記の数を絞った映像表現機構を統合物語生成システムの応用システム KOSERUBE の中に導入した。今後特に簡易アニメーション定義の作業を組織化して行くための方略として、動詞的概念を行為と動作に分け、さらに動作を基本的動作と複合的動作に分ける考えも示した。それを整理すれば、行為の構造は物語の概念構造生成の部分に委ね、アニメーション定義は動作の動詞的概念だけに絞ることが出来るようになる。また、画像素材の収集と格納は Web の自動検索で行う等、様々な工夫をして行く予定である。

#### 参考文献

- Akimoto, T., Endo, J. and Ogata, T. (2012). "The Expansion of Paths in the Mutual Transformation Mechanism of Music and Narrative." *Proc. of the 11th IEEE International Conference on Cognitive Informatics & Cognitive Computing*, pp.230-239.
- Akimoto, T. and Ogata, T. (2012). "Macro Structure and Basic Methods in the Integrated Narrative Generation System by Introducing Narratological Knowledge." *Proc. of the 11th IEEE International Conference on Cognitive Informatics & Cognitive Computing*, pp.253-262.
- Bates, J. (1992). *The Nature of Characters in Interactive Worlds and The Oz Project*, Technical Report CMU-CS-92-200, School of Computer Science, Carnegie Mellon University, 1992.
- Imabuchi, S. and Ogata, T. (2012). "A Story Generation System based on Propp Theory: As a Mechanism in an Integrated Narrative Generation System." *Lecture Notes in Artificial Intelligence 7614 (Isahara H, Kanzaki K. Eds.), Springer-Verlag*, pp.312-321.
- 今渕祥平, 小野淳平, 遠藤順, 栗澤康成, 道又龍介, 鎌田まみ, 秋元泰介, 小方孝 (2012). 『いわての民話 KOSERUBE』—プロップによるストーリー生成システムをベースに文・音楽・視覚表現の生成を統合したシステム—, 日本認知科学会第 29 回大会発表論文集, pp.733-742.
- 真部雄介, 小方孝 (2007). 物語生成における映像からの動作概念記述のボトムアップアプローチ, 2007 年度人工知能学会全国大会 (第 21 回) 論文集, 1F1-7.
- Oishi, K., Kurisawa, Y., Kamada, M., Fukuda, I., Akimoto, T. and Ogata, T. (2012). "Building Conceptual Dictionary for Providing Common Knowledge in the Integrated Narrative Generation System." *Proc. of 34th Annual Conference of the Cognitive Science Society*, pp.2126-2131.
- 小方孝 (2008). 物語生成システムにおける映像構成へ向けて, 金井明人, 丹波美之 編. 映像編集の理論と実践, pp.165-235. 法政大学出版局.
- 小方孝, 金井明人 (2010). 物語論の情報学序説—物語生成の思想と技術を巡って—. 学文社.
- 小野淳平, 秋元泰介, 小方孝 (2012). 視覚表現を物語生成とつなげる方法の検討—ふたつの応用システムを素材として—, 日本認知科学会第 29 回大会発表論文集, pp.353-362.
- Piacenza, A., Guerrini, F., Adams, N., Leonardi, R., Porteous, J., Teutenberg, J. and Cavazza, M. (2011). "Generating Story Variants with Constrained Video Recombination." *Proc. of the 19th ACM International Conference on Multimedia*, pp.223-232.
- Propp, V. (Пр о п п, В. Я.) (1969). *Мор ф о л о г и я с к а з к и*, Изд.2е. Москва: Наука. (北岡誠司, 福田美智代 訳 (1987). 昔話の形態学. 水声社.)
- 土佐尚子, 中津良平 (1997). 感情に反応する自律型バーチャルアクターと仮想世界の生成, 日本バーチャルリアリティ学会論文集, 2(1), pp.11-18.