

手指動作入力を併用した音声駆動型身体引き込みキャラクターの映像コンテンツ制作への応用

大崎 浩司^{†1} 渡辺 富夫^{†2,†3} 山本 倫也^{†2}

InterPuppet は、発話音声からコミュニケーションの引き込み動作を自動生成する音声駆動型身体引き込みキャラクター InterActor に、手指動作による片手使い人形のような操作入力を併用したインタラクションキャラクターである。本論文では、InterPuppet を映像コンテンツ制作に応用し、InterPuppet によるコンテンツ制作の評価システムを構築している。さらに、システムを 5 日間使用したコンテンツ制作作業の評価と、制作されたコンテンツの視聴評価を行うことで、制作者と視聴者双方の視点から提案手法の有効性を示している。

Application of Speech-driven Embodied Entrainment Character by Speech and Hand Motion Input to Video Contents Production

KOUZI OSAKI,^{†1} TOMIO WATANABE^{†2,†3}
and MICHIIYA YAMAMOTO^{†2}

InterActor is a speech-input-driven CG-embodied interaction character that can generate communicative movements and actions for entrained interaction. InterPuppet, on the other hand, is an embodied interaction character that is driven by both speech input, similar to the InterActor, and hand motion input, like a puppet. In this study, we apply InterPuppet to video contents production, and we construct a system to evaluate the contents production by InterPuppet. The self-evaluation of long-term video contents production for 5 days and the sensory evaluation of the contents demonstrate the effectiveness of the developed system.

1. はじめに

今日 CG キャラクターを使用した映像コンテンツがますます増えており、子供から大人まで人気を集めている。俳優やぬいぐるみのキャラクターなどと同じように親しみの持てる存在としての地位を確立しつつある。今後は、テレビの多チャンネル化¹⁾ やストリーミング放送、消費者発信型メディアによる個人での動画配信の増加²⁾ にともない CG キャラクターを活用する機会がさらに増えることが予想される。このような CG キャラクターによるアニメーション制作支援を目的としてキャラクターの動作生成や CG キャラクターを使用した映像コンテンツに関する研究³⁾⁻⁶⁾ も多く行われており、道家らは、番組の制作や公開、視聴を行うテレビ環境を作成している⁷⁾。また、CG キャラクターの動作はリアルタイムに合成することも可能であり、映像コンテンツにおける CG キャラクターの役割はさらに大きくなると考えられる。

一方、人は対面で情報を伝えるとき、単に言葉によるバーバル情報だけではなく、音声に対するうなずきや身振り・手振りなどのノンバーバル情報が相互に同調して、対話者が引き込みあうことで円滑なコミュニケーションを行っている。著者らはこの身体的リズムの引き込みに着目し、発話音声のみを入力としてキャラクターの自然なコミュニケーション動作を自動生成する音声駆動型身体引き込みキャラクター InterActor を開発し、有効性を示している^{8),9)}。また、InterActor に手指動作による操作入力を併用した身体引き込みキャラクター InterPuppet を開発し、音声からの自然なコミュニケーション動作と手指からの意識的な動作を併用することの有効性をコミュニケーション支援において示している^{10),11)}。

映像コンテンツにおいても CG キャラクターは音声をともなって動作を行うことが多いことからキャラクターのかかわり動作は非常に重要と考えられる。InterPuppet のような引き込み動作と意識的な動作を組み合わせる仕組みを導入することで、生命感のある豊かなコミュニケーション動作により映像コンテンツ制作を支援し、ニュースや生中継など幅広い分野への応用が期待できる。

本研究では、InterPuppet のコンテンツ制作への応用を目指して、InterPuppet によるコンテンツ制作の評価システムを構築している。次に、長期間にわたりコンテンツ制作作業を

†1 岡山県立大学大学院情報系工学研究科

Graduate School of Systems Engineering, Okayama Prefectural University

†2 岡山県立大学情報工学部

Faculty of Computer Science and Systems Engineering, Okayama Prefectural University

†3 科学技術振興機構戦略的創造研究推進事業

CREST of Japan Science and Technology Agency

行うことで制作者の視点からコンテンツ制作における InterPuppet の有効性を評価している。さらに、制作した映像コンテンツを視聴・評価することで、視聴者の視点からも提案手法の有効性を示している。

2. InterPuppet

操作できる人形をメタファとするキャラクタ操作に関する研究^{12)–15)}に加え、物理メディアとしての人形あるいはぬいぐるみを介したインタラクションに関する研究も多く行われている^{16)–19)}。これに対して、著者らは、キャラクタを介したコミュニケーションを支援するために、InterActor に手指動作入力を併用した、手指動作併用音声駆動型身体引き込みキャラクタ InterPuppet を開発している¹⁰⁾。

InterPuppet の動作部位を図 1 に示す。動作部位はうなずき、瞬き、口の開閉に加え、頸部、腰部、肩関節、肘関節、手首関節部であり、豊かなコミュニケーション動作が可能である。音声に基づくコミュニケーション動作生成モデルは、音声の ON-OFF の時系列からうなずき動作、瞬き、身体動作のタイミングを推定する移動平均モデル (MA モデル: 30 Hz でパラメータ数 60) を使用した⁹⁾。

人と人が対面でコミュニケーションを行うときは、このような引き込み動作に加えて、会話にあわせて肯定や否定、リズム同調をあえて崩すなどの意思や感情を身体動作で表すといったように、意識的な身体動作を行っている^{20),21)}。InterPuppet では、この意識的な身体動作を手指から入力し、図 1 の動作部位を操作することで、引き込みや一体感を確保しつつ使用者の意識的な動作を加えることを可能とし、より多様なコミュニケーション表現をリアルタイムにキャラクタに行わせることができる。



図 1 InterPuppet
Fig. 1 InterPuppet.

手指動作の入力デバイスとして、データグローブと携帯型キーボードを使用したコミュニケーション支援システムを開発し、遠隔コミュニケーション実験を行った結果「楽しさ」や「キャラクタになった感覚」といった観点で高い評価を得ており、コミュニケーション支援において有効であることが示されている¹⁰⁾。

3. コンテンツ制作評価システム

3.1 コンテンツ制作の概要

InterPuppet を応用したコンテンツ制作の概要を図 2 に示す。先行研究^{10),11)} で開発した InterPuppet によるコミュニケーション支援システムは、対話者が互いに自分の代役となる CG キャラクタを介してコミュニケーションを行うものであり、対話相手が CG キャラクタの後ろに必ずいることを前提としている。これに対し映像コンテンツ制作では、キャラクタが対話者の代役となるのではなく、キャラクタそのものに役割を持たせ、制作者がそれを演じる。さらに制作者が演じたキャラクタ動作を不特定多数の視聴者が見ることになり、その視聴者を引き込む魅力的なコンテンツの制作を行う必要がある。提案システムでは、コンテンツに登場するキャラクタを InterActor として動作させることにより、身体的リズムの共有による引き込みや、キャラクタの生命感、自然な動きが確保される。特に情報を伝えたり、説明を行ったりするような映像コンテンツの制作では、キャラクタが話し手として視聴者に語りかける動作を行い、視聴者を引き込むことができる。さらに、手指からの動作入力によりキャラクタに片手使い人形のように動作を入力することで、映像コンテンツに必要な意識的な身体動作を発話内容に合わせて行わせることができる。つまり、キャラクタの台詞として入力される制作者の声から InterActor として話し手の動作を生成しつつ、必要に応じて物を指し示すなどの映像コンテンツとして必要な動作をリアルタイムで

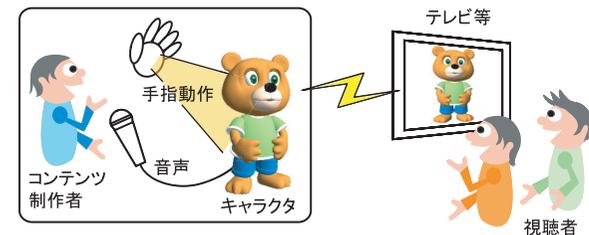


図 2 コンテンツ制作の概要
Fig. 2 Outline of the contents production.

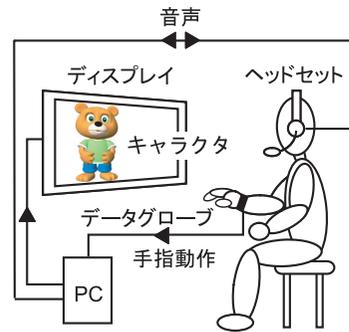


図 3 システム構成図
Fig. 3 System configuration.

与え、番組の司会進行や、画面に表示される情報の説明、他の登場人物とのインタラクシオンを行うことができる。操作入力のみでキャラクターの動作を作る手法に比べ、身体的リズムの引き込みに基づいた豊かなコミュニケーション動作を容易に生成し、映像コンテンツ制作を支援することができる。

3.2 システム開発

InterPuppet によるコンテンツ制作の評価システムを構築した。システム構成図を図 3 に示す。システムは片手使い人形のようにキャラクターを動作させるデバイスとして、データグローブ (Immersion, CyberGlove) を用いており、PC、映像を提示するディスプレイ、音声入出力用のヘッドセットにより構成した。PC は、音声入出力、シリアルポート (データグローブ接続用)、DirectX9.0 対応ビデオカードを備えたものである。本システムでは WindowsXP を搭載した AT 互換機を使用した。作成した映像コンテンツは、PC のハードディスクに記録され、自由に再生することができる。

映像コンテンツの一例として草花を紹介する情報番組を想定した。画面構成を図 4 に示す。画面には使用者が操作する CG キャラクターと背景画像として静止画と動画を配置した。動画は画面中央に、CG キャラクターは右側に配置した。CG キャラクターは、愛・地球博プロトタイプロボット展 (2005 年 6 月 9 日 ~ 19 日) で男女ともに人気であった熊型ロボット「トミー」と同一デザインの CG を使用し、立ち位置はどのような動作を行っても画面からはみ出ないように調整した。動画の位置や大きさは図のように必要に応じて変えることができ、拡大縮小はテレビ番組の構成を模してスムーズに拡大縮小のエフェクトを加えること



図 4 画面構成
Fig. 4 Screen composition.

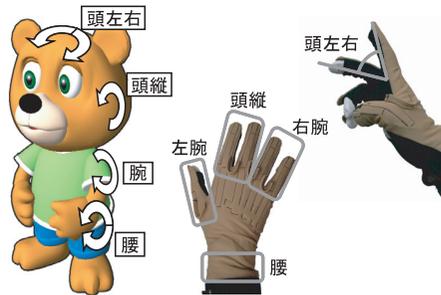


図 5 データグローブでの操作

Fig. 5 Data glove operation to InterPuppet.

が可能である．さらに，動画にあわせて BGM の再生を行うことができる．それぞれのオブジェクトは Microsoft, DirectX9.0c により描画した．背景や動画，画面効果を変えることで，さまざまな映像コンテンツの制作を行うことができる．

3.3 キャラクタの動作生成

音声からの動作生成では，台詞音声から InterActor としての動作を生成している．著者らはこれまでに，音声からの話し手動作により映像コンテンツを制作する InterCaster を開発しており，研究室ベンチャーのインタロボット株式会社*1より，テレビ番組などへ応用されている．この InterCaster において，話し手動作に聞き手動作を付加し，自分の発話音声に対してもうなずきなどの聞き手動作を行うことで，コミュニケーション動作の効果が高まることが実証されていることから，本システムでは，入力された音声から話し手としての動作を生成すると同時に，聞き手動作も生成し合成している．話し手，聞き手の身体動作のタイミングの推定は先行研究¹⁰⁾と同じものを使用した．音声データは 16 bit 22,050 Hz でサンプリングした．

手指動作からの動作生成では，データグローブから取得した手指動作をキャラクタの動作に線形的に対応付けることで操作を行う．データグローブでは手袋部に取り付けられた計 18 個のセンサにより，手指の動きを手首や指の曲げ角度情報として分解能 0.5 度で取得できる．手指部の動作とキャラクタの動作との対応は，図 5 のような対応にした．この対応は片手使い人形の操作²²⁾を参考にして，人形を操るのに似た感覚で操作できるようにした．

表 1 手指動作とキャラクタの動作の対応

Table 1 Correspondence of hand motion and character motion.

キャラクタ稼動部	手指動作との関係
頭縦	人差し指と中指の第2第3関節の曲げ角の合計を0.2倍した角度で頭が縦に回転
頭左右	中指の第2第3関節の曲げ角の合計から人差し指の第2第3関節の曲げ角の合計を引いた値を0.5倍した角度で頭が鉛直方向に回転
右腕	小指と薬指の第2第3関節の曲げ角の合計を0.3倍した角度で肩が回転，肘の関節が肩の0.7倍の角度で回転
左腕	親指の第1第2関節の曲げ角の合計を0.5倍した角度で肩が回転，肘の関節が肩の0.7倍の角度で回転
腰前後	手首の曲げ角の0.3倍の角度で回転

手指動作からキャラクタ稼動部の動作への変換は初期値を表 1 のように設定し，必要に応じてキャリブレーションを行い値を調節できるようにした．InterActor と操作入力と同時に同じ部分进行操作しようとしたときは InterActor を優先するように設定した．なお，手指部の動作は 20 Hz で各関節の曲げ角度を取得した．

4. コンテンツ制作実験

4.1 実験概要

コンテンツ制作者側からの評価実験として，テレビの生放送を想定した番組制作を行った (図 6)．先行研究では日常のコミュニケーションでの有効性を評価するために自由対話を行ったが，本実験ではコンテンツを制作するという観点から，あらかじめ決められた内容に従って作業を行うこととした．また，被験者にコンテンツ制作作業を意識させるために，仕事としてのコンテンツ作りを想定し，映像コンテンツを収録する緊張感の中でのコンテンツ作りを行わせた．実験システムの配置を図 7 に示す．制作現場としての緊張感を与えるためにダミーの放送機材 (SONY, FXE-120) を配置し，制作者役としての被験者 1 人に対して制作スタッフ役の実験者 2 人で実験を行った．実験者は被験者への指示や DV カメラの映像のチェックのほかにダミー機器の操作などを行った．さらに，制作者から不特定多数の視聴者への情報の伝達を想定し，制作した作品は第三者に公開するものとして，各被験者に伝えることで，作品としての完成度を要求した．草花を紹介する 1 分間の番組の制作にあらかじめ原稿と動画を用意し，動画にあわせて原稿を読み上げながらキャラクタ

*1 インタロボット株式会社 : <http://www.i-robot.co.jp/>



図 6 実験風景
Fig. 6 Experimental scenery.

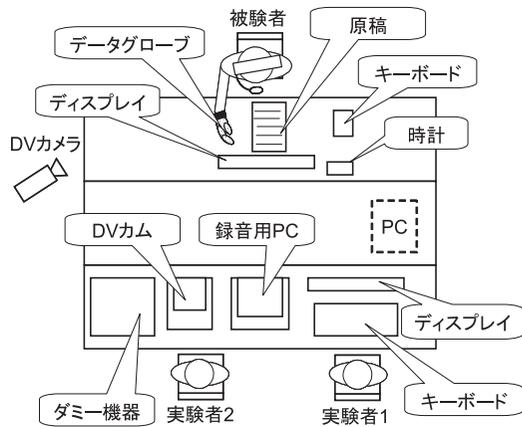


図 7 機材の配置
Fig. 7 Arrangement of machine parts.

の動作生成を行わせた。定期的なコンテンツ制作作業を行った場合の影響を調べるために、番組は全 15 回で構成され、1 日あたり 3 回の収録を 16 日以内に 5 日行った。草花の動画と原稿は番組ごとに違うものを用意し、原稿は平均 261.8 モーラで、すべて 1 分程度で読めるように調整した。また、BGM と動画の拡大縮小のタイミングは 15 回とも同じになるように設定した。

実験ではキャラクターの動作生成に、次に示す A~C の 3 つのモードを用意し、それぞれのモードで番組を作成することで、InterPuppet の評価を行った。

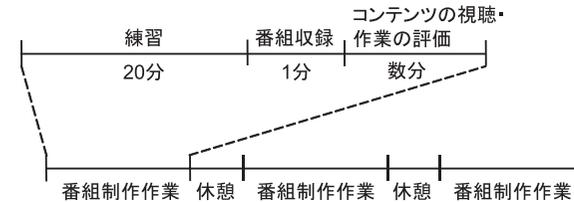


図 8 1 日のタイムテーブル
Fig. 8 Time table.

- A 手指動作のみによる操作 (口の開閉と瞬きは B と同様)
- B InterActor
- C InterPuppet (A + B)

被験者には、各モードの違いとキャラクターの操作法を、実際にシステムを使用させながら説明し、自由に操作できていることを確認している。また、必要に応じてデータグローブのキャリブレーションを行った。番組の制作は、図 8 の手順で行い、1 日につき 3 つのモードをランダムな順で使用した。練習では原稿を読む練習やキャラクターの操作を自由に行わせ、動画の再生は被験者がキーボードから自由に操作できるようにした。収録ではテレビの生放送を想定したことから、やり直しを許可しないものとし、実験者が合図を出して、時間どおりに動画の再生を行った。各モードでの収録が終わるごとに制作したコンテンツを視聴させ、作業の評価を行った。作業の評価では、作業としてキャラクターの動作生成を行ったことからメンタルワークロードを評価するための NASA-TLX²³⁾ とキャラクター動作生成における身体的インタラクション支援の観点から「楽しさ」、「話しやすさ」、「キャラクターになった感覚」、「キャラクターを操作している感覚」の 4 項目についての 7 段階評価を行った。NASA-TLX では、知的・知覚的負荷 (mental demand)、身体的負荷 (physical demand)、タイムプレッシャ (temporal demand)、作業成績 (own performance)、努力 (effort)、フラストレーション (frustration level) の 6 項目について一対比較による重み付けと 0~100 の評定値の決定を被験者に行わせ、加重平均することでメンタルワークロードの総合得点である WWL 得点 (Mean Weight Workload Score) を求めた。また、モード A と C の収録中にグローブの操作によるキャラクターの操作量を記録した。番組制作作業の間には任意で数分間の休憩をとった。

さらに、5 日の実験終了後にすべての作業を通しての全体評価としてアンケートを行った。アンケートの項目は、A、B、C の各モードで生成したキャラクターの動作についてコンテン

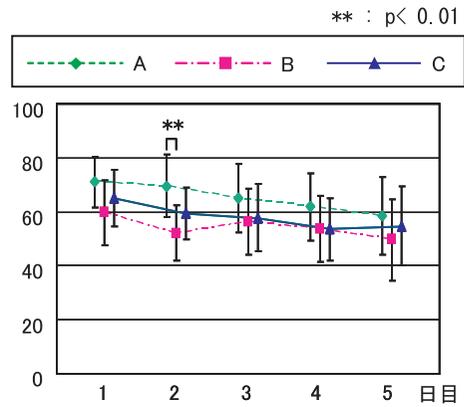


図9 5日間の評価 (NASA-TLX)
Fig. 9 Result of NASA-TLX in 5 days.

ツ中の CG キャラクタの役割を演じて動作を作成したことから「生命感のある動き」、「番組の内容を伝える動き」、「豊かな動き」、決められた原稿を読む作業を行ったことから、原稿の読みについて「うまく読む」、第三者に見せるコンテンツの制作を行ったことから、作品の出来について「満足できる」、「人に見せたい」を7段階評価させた。そのとき同時に「練習すればもっとうまくなるか」という質問を行った。

被験者は18~22歳の男女学生10人であった。

4.2 実験結果

NASA-TLXのWWL得点の平均と標準偏差、分散分析によるモード間の統計的有意差を図9に示す。各モードとも日ごとに負荷が下がっているが、すべて手指で操作するAのモードの負荷が他のモードより高い。

7段階評価の結果を図10に示す。図では、日ごとの被験者間の平均と標準偏差、Friedmanの分散分析法²⁴⁾によるモード間の統計的有意差を表している。「楽しさ」ではInterPuppetと手指動作のみによる操作が高い評価を得た。「話しやすさ」ではInterActorが5日間通して高く評価され、InterPuppetも手指のみの操作に比べ高い評価であった。「キャラクタになった感覚」と「キャラクタを操作している感覚」では手指での操作が最も高くなったが、「キャラクタになった感覚」についてはInterPuppetの評価も次第に高くなった。

モードAとCにおける手指動作からのキャラクタ操作量を図11に示す。図では、番組収録1分間のキャラクタ操作に使用した指および手首関節動作量の累計の平均と標準偏差、

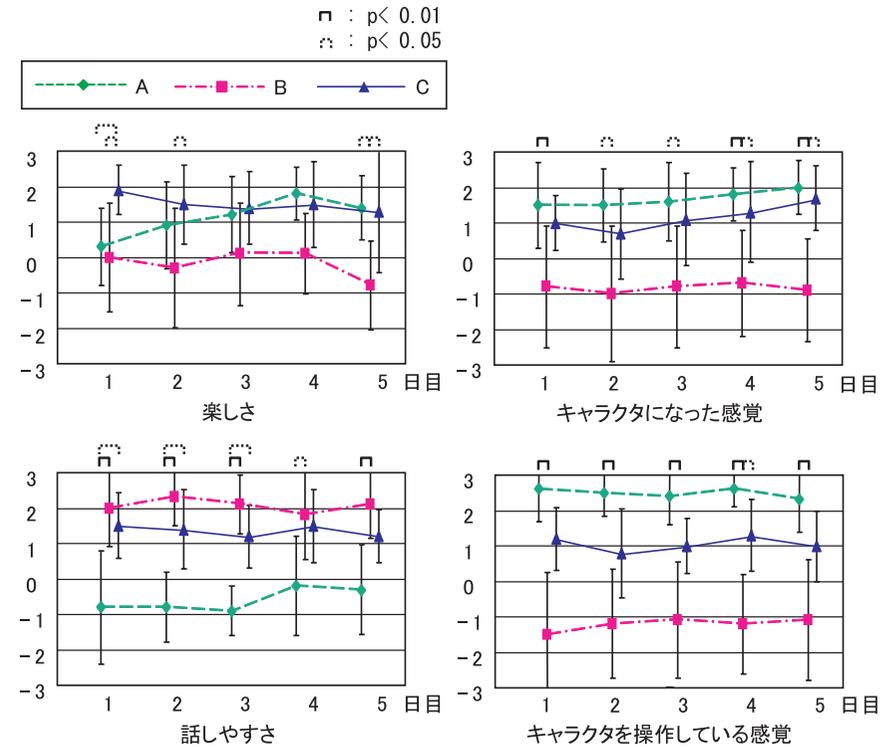


図10 5日間の評価 (7段階評価)
Fig. 10 Sensory evaluation in 5 days.

t検定によるモード間の有意差を表している。3日目以降に有意な差が見られる。

また一例として、モードA~Cの頭部の縦方向動作と横方向動作を図12に示す。図はある被験者が5日目に制作したコンテンツの開始後20秒間のキャラクタの頭部動作である。縦方向動作ではうなずき方向、横方向動作ではキャラクタが右を見る方向の回転を正としている。図中グレーのハイライト部分はInterActorのうなずき動作が行われている部分である。InterActorの動作がないモードAではうなずき動作の回数が少なくなっている。また、手指からの操作を行うモードAとCではそれぞれ開始後8秒付近と12秒付近で大きく右方向を向いている部分が見られる。この部分ではキャラクタが動画部分を見る動作を行っている。モードCにおける11.5秒と、12.0秒時点のスクリーンショットを図13に示す。

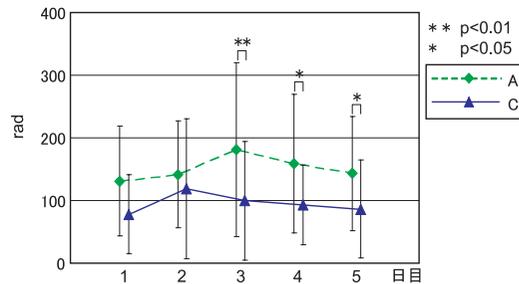


図 11 キャラクタ操作量

Fig. 11 Amount of character operation.

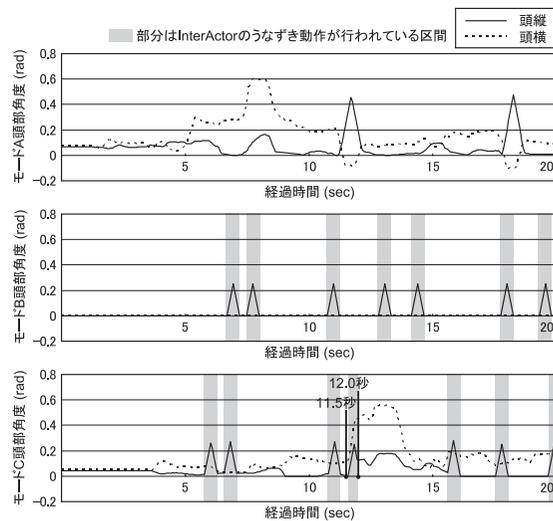


図 12 頭部の動作例

Fig. 12 Example of head movements.

さらに、5日間のモード A~C におけるキャラクタ動作量を図 14 に示す。モード A では図 11 の手指動作により動作したキャラクタ関節の動作量、モード B では A と同じ関節における InterActor によって生成された動作量、モード C では同関節における手指動作と InterActor から生成された動作量を平均と標準偏差で表している。分散分析によるモード間の有意差を同図に示す。キャラクタの動作量はモード A よりモード C の方が多くなって

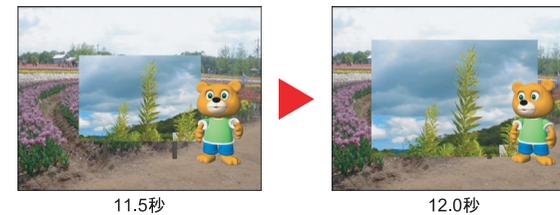


図 13 頭部を横に振る様子

Fig. 13 Screenshot of head movement (Left-Right).

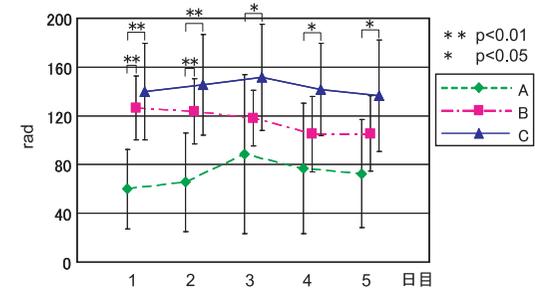


図 14 キャラクタ動作量

Fig. 14 Amount of character movement.

いる。

実験終了後に行った全体評価についてのアンケートの結果を図 15 に示す。図では全体の平均値と標準偏差、Friedman の分散分析法によるモード間の統計的有意差を表している。「生命感のある動き」、「豊かな動き」、「満足できる」、「人に見せたい」の項目では InterPuppet、InterActor、手指動作のみによる操作の順に高く評価され、特に生命感のある動きと豊かな動きで InterPuppet がほかの 2 つのモードに比べ高い評価を得ている。「うまく読む」の項目では InterActor が最も高く評価された。また、「番組の内容を伝える動き」では InterPuppet が InterActor よりも高く評価された。

「練習すればもっとうまくなるか」という質問に対する回答を図 16 に示す。手指動作のみによる操作では 10 人全員が、InterPuppet では 8 人がうまくなると回答した。一方、InterActor のモードでは 10 人中 2 人がうまくなると回答し、早期の習熟がうかがえた。

4.3 考 察

実験期間中の習熟の結果、被験者の評価に変化が現れた。NASA-TLX の結果からはコン

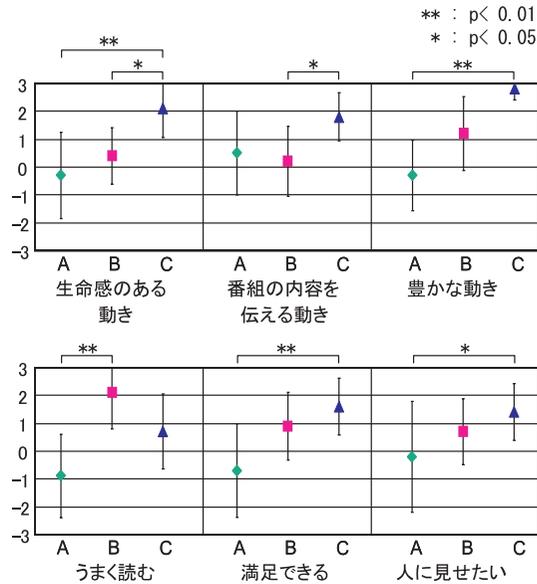


図 15 全体評価
Fig. 15 Self evaluation of contents.

テンツ制作作業の負荷が減少していることが分かる。「キャラクタになった感覚」では手指動作のみによる操作と InterPuppet の評価が次第に高くなっており、手指からのキャラクタ操作の習熟がうかがえる。また、「楽しさ」では初めは InterPuppet の評価が手指動作のみによる操作より高いが 5 日目には手指動作のみによる操作の評価も同等に高くなっている。ただし、キャラクタ操作量において 3 日目以降に有意な差が見られることなどから、手指動作のみによる操作と InterPuppet では使用法に違いが生じて、どちらの評価も高くなったと考えられる。

「話しやすさ」では InterActor の評価が 5 日間を通して高いままで、従来からの InterActor は誰もが使えるコミュニケーション支援システムとして有効といえる。「キャラクタを操作している感覚」では逆に InterActor の評価が低かった。InterPuppet に注目すると、この「話しやすさ」、「キャラクタを操作している感覚」で、他のモードの間の評価を得ている。さらに、NASA-TLX の結果や「キャラクタになった感覚」においても他の 2 つのモードの間の評価を得ており、これらの観点で他のモードの間の特性を持つバランスのとれたモードで

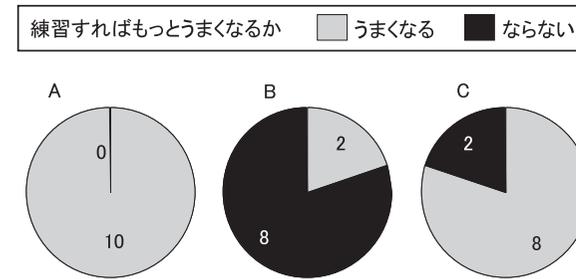


図 16 上達についての感想
Fig. 16 Impression of progress.

あるといえる。

操作量の変動を見ると、手指のみによる操作では 3 日目、InterPuppet では 2 日目にピークがあるが、キャラクタの動作量ではともに 3 日目がピークとなっている。この理由として、習熟により 3 日目以降に InterActor によって生成される動作と手指動作による動作をよりうまく使い分けることができた、あるいは InterPuppet の「キャラクタになった感覚」が 3 日目以降上昇していることから、キャラクタになりきって原稿を読むことで、読み方が変化し、InterActor の動作量も変化したことが考えられる。操作量、動作量ともに、ピークを越えた後も徐々に減少しており、習熟による収束傾向が見られる。しかし、InterPuppet の動作量には InterActor によるものが含まれるので InterActor を下回ることはなく、少ない操作量でコンテンツ制作が行えるといえる。

全体評価では、InterActor によって生成される身体動作に対して手指からの動作を重畳合成することから、InterPuppet が「豊かな動作」で高く評価された。音声からつねにコミュニケーション動作を生成しつつ、制作者の意識した動作を加えることができるために「生命感のある動き」でも高く評価されたと考えられる。「うまく読む」では原稿を読むだけでキャラクタ動作を生成できることから InterActor が高い評価を得たと考えられる。また、InterPuppet は 5 日間の「楽しさ」や「満足できる」といった項目で評価が高く、練習すればもっとうまくなると評価されていることから、コンテンツ制作者が長期にわたって使い続ける制作手法として有効であると考えられる。これらの結果は CG キャラクタのコミュニケーション動作生成法として好ましい結果である。さらに、「番組の内容を伝える動き」、「人に見せたい」といった項目が高く評価されたことから、提案手法はコンテンツ制作支援において有効といえる。

5. コンテンツ視聴評価実験

5.1 実験概要

4章で制作した映像コンテンツを視聴評価することで、視聴者側からの評価を行った。実験風景を図17に示す。比較には制作作業最終日の5日目に制作した映像コンテンツを使用した。同じ制作者による3モードの映像コンテンツから、2つを選んで「キャラクタの説明する動作が良い」方を選択する比較をすべての組合せで行い、一対比較した。視聴評価の被験者は男女学生各5人計10人であり、4章の被験者10人分のコンテンツを評価することで、各モードの組合せをそれぞれ100回比較した。

映像コンテンツの比較に用いたソフトウェアの画面構成を図18に示す。ソフトウェアは、映像表示部と操作部、決定ボタンを持ち、一対比較のためにそれぞれ左右に配置した。操作



図 17 実験風景
Fig. 17 Experimental scenery.

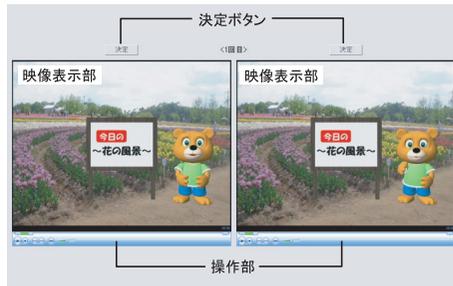


図 18 比較用インタフェース
Fig. 18 Interface for comparison.

部は、映像コンテンツの再生に広く使われている Microsoft, WindowsMediaPlayer と同様のインタフェースを有しており、再生、一時停止、停止、再生位置の変更、ボリュームの調整を自由に行うことができる。さらに、映像表示部をクリックすることでも再生と一時停止を交互に切り替えることができる。片方の映像コンテンツを再生中にもう一方の映像コンテンツを再生したときは、強制的に一時停止になるようにし、片方ずつしか再生できないようにした。左右に割り当てられた映像コンテンツをそれぞれ見てから、良いと思う方の決定ボタンを押すことで比較作業が行えるようにした。決定ボタンを押すと次の比較用の映像コンテンツが自動的に割り当てられる。ただし、左右の映像コンテンツをそれぞれ5秒以上見なければ決定ボタンは押せないようにした。使用したPCは、17インチワイド液晶とステレオスピーカを有するノートPC(東芝, Qosmio G30/697HS)であり、ソフトウェアは全画面表示にして使用した。

実験ではまず、被験者に比較用ソフトウェアと同じインタフェースを持つ練習用ソフトウェアを使用させ、ソフトウェアの使い方と比較方法に慣れさせた。このときの映像コンテンツは、実験者が別途作成したものを使用した。次に実際の比較作業を行った。比較順序と、左右の割当ては、被験者ごとにランダムとした。また、比較終了後にどのような基準で比較したかを自由記述で聞いた。

5.2 実験結果

一対比較の結果を表2に示す。表中の数字は各モードの勝数、つまり各行のモードを選択した回数を表している。この結果に対して被験者による評価を定量的に評価するために、Bradley-Terryモデルを想定した²⁴⁾。図19に求めた強さ π (平均 50) を示す。InterPuppetのモードCおよびInterActorのBが手指だけの操作のAに比べ高く評価されていることが分かる。さらに、人に見せるために制作した映像コンテンツという観点から、4章の全体評価で「人に見せたい」の項目について、全モードで0以上の評価をつけた制作者4人の作品に対する評価のみを抜き出したものを表3に示す。図20は同様に、Bradley-Terryモデルを想定し、強さ π (平均 50) を求めたものである。InterPuppetがInterActorよりも

表 2 一対比較の結果

Table 2 Result of paired comparison.

	A	B	C	total
A		28	31	59
B	72		47	119
C	69	53		122

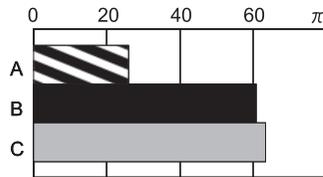


図 19 各モードの強さ π
Fig. 19 Comparison of π .

表 3 人に見せたい作者のコンテンツのみに基づく一対比較の結果

Table 3 Result of paired comparison based on the only contents the authors who want to show.

	A	B	C	total
A		15	10	25
B	25		18	43
C	30	22		52

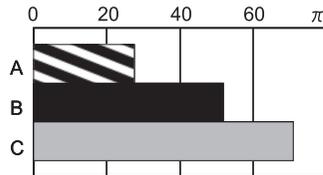


図 20 人に見せたい作者のコンテンツのみに基づく各モードの強さ π
Fig. 20 Comparison of π based on the only contents the authors who want to show.

高く評価されていることが分かる。

コンテンツの評価基準として視線と、うなずきやあいづちなどの頭部の縦振り動作に関することをあげる被験者が多く、視線は 10 人中 7 人、頭部の縦振り動作は 6 人があげた。4 章の実験で使用した CG キャラクタは目の向きが固定であり、視線は頭部の動作によるものである。

5.3 考 察

視聴者側の一対比較の結果からも InterPuppet の有効性が示された。すべての作者のコンテンツを評価した場合、InterActor の引き込み動作を行う B と C のモードが高く評価されたことから、映像コンテンツ中のキャラクタには音声からのコミュニケーション動作が必要といえる。InterActor のモードでも InterPuppet に近い評価を得ていることから、InterActor の身体的引き込みのメカニズムを使用すれば、誰もが手軽に CG キャラクタを

使った映像コンテンツを作成できると期待できる。人に見せたいと回答した制作者のコンテンツに限定した場合は、InterActor よりも InterPuppet を使用したコンテンツがさらに高く評価されており、これらのコンテンツを制作した制作者は、手指からの意識的な動作の提示と音声からのコミュニケーション動作の両方を活かして映像コンテンツを制作できていたと考えられる。また、InterActor により自動生成されるうなずきと、手指から入力される頭部動作による視線に注目する被験者がいたことから、映像コンテンツの制作において InterActor に手指からの意識的な動作を併用する InterPuppet のメカニズムが有効であると考えられる。

以上より、映像コンテンツの制作者だけでなく、視聴者側からも InterPuppet が高く評価されたことから、提案手法の有効性が示された。

本実験では記録した映像コンテンツを使用して視聴評価を行ったが、これらの映像コンテンツは 4 章において生放送を想定し、原稿を読みながら動作を作成したものである。よって、InterPuppet を使用すれば生中継などのリアルタイム性の強い映像の制作においても声優が CG キャラクタを使用して、評価の高い映像コンテンツを作成できると期待できる。

6. おわりに

手指動作入力併用音声駆動型身体的インタラクションキャラクタ InterPuppet を映像コンテンツ制作に応用し、制作者と視聴者の双方の視点から評価した。まず、InterPuppet の特徴を活かした映像コンテンツ制作手法の概要を示し、評価システムの構築を行った。次に、システムに手指動作入力のみでの操作、InterActor、InterPuppet の 3 つのキャラクタ動作生成モードを設け、5 日間コンテンツ制作を行うことで評価を行った。その結果、InterPuppet がコンテンツ制作作業において、「楽しさ」、「話しやすさ」、「キャラクタになった感覚」で高く評価された。また、作成したキャラクタの動作に対しても「生命感のある動き」、「番組の内容を伝える動き」、「豊かな動き」といった観点で高く評価された。さらに、制作したコンテンツを視聴・評価した結果、キャラクタの説明する動作が良いという観点で高く評価され、映像コンテンツ制作者、視聴者から提案手法の有効性が示された。

著者らはこれまでに、教育番組の映像に仮想観客として InterActor を複数体配置することで、学習効果を高める教育支援システムを開発し、その学習効果を明らかにしている²⁵⁾。コンテンツ制作支援においても、仮想観客としての InterActor を複数体配置することで、場を盛り上げ、魅力的なコンテンツを制作することができると考えられ、今後の課題である。

謝辞 本研究の一部は、科学技術振興機構 (JST) の戦略的創造研究推進事業 (CREST)

研究領域「デジタルメディア作品の制作を支援する基盤技術」における「人を引き込む身体性メディア場の生成・制御技術」プロジェクトの支援による。

参 考 文 献

- 1) 亀山 歩, 花村 剛: MPEG1/MPEG2/MPEG4 デジタル放送教科書(上), pp.1-24, IDG ジャパン (2003).
- 2) 総務省: 平成 19 年版情報通信白書, pp.217-222, ぎょうせい (2007).
- 3) 森脇 晃, 國分 剛, 尾下真樹: 自然言語からの仮想人間の複合動作アニメーションの生成, 情報処理学会研究報告グラフィックスと CAD, Vol.2006, No.91, pp.67-72 (2006).
- 4) 中野 敦, 塩入健太, 星野准一: 対話型キャラクタのための姿勢, しぐさ, ジェスチャの複合による心理状態表現, インタラクション 2006 論文集, pp.149-156 (2006).
- 5) 齋藤 豪, 井元崇之, 中嶋正之: モーショングラフ: 自律型キャラクタの動作生成手法, 芸術科学論文誌, Vol.1, No.1, pp.22-29 (2002).
- 6) Buttussi, F., Chittaro, L. and Nadalutti, D.: H-animator: A visual tool for modeling, reuse and sharing of X3D humanoid animations, *Proc. Web3D 2006*, pp.109-117 (2006).
- 7) 道家 守, 浜口斉周, 林 正樹: TV4U —番組の制作から発信, 視聴までを統合した, 新しいテレビ環境, 信学技報 ITS2004-72 IE-2004-206, pp.41-46, 電子情報通信学会 (2005).
- 8) Watanabe, T., Okubo, M., Nakashige, M. and Danbara, R.: InterActor: Speech-Driven Embodied Interactive Actor, *International Journal of Human-Computer Interaction*, Vol.17, No.1, pp.43-60 (2004).
- 9) 檀原龍正, 渡辺富夫, 大久保雅史: 音声駆動型身体引き込みキャラクタ InterActor が発話音声に与える影響, 日本機械学会論文集(C編), Vol.71, No.712, pp.152-159 (2005).
- 10) 大崎浩司, 渡辺富夫, 山本倫也: InterActor に手指動作入力を併用した身体的インタラクションシステムの開発, ヒューマンインタフェース学会論文誌, Vol.7, No.3, pp.399-408 (2005).
- 11) 大崎浩司, 渡辺富夫, 山本倫也: InterActor に手指動作入力を併用した身体的インタラクションシステム, インタラクション 2006 論文集, pp.165-166 (2006).
- 12) Mazalek, A. and Nitsche, M.: Tangible interfaces for real-time 3D virtual environments, *Proc. International Conference on Advances in Computer Entertainment Technology (ACE'07)*, pp.155-162 (2007).
- 13) Kim, S., Zhang, X. and Kim, Y.J.: Haptic Puppetry for Interactive Games, *Edu-tainment 2006*, pp.1292-1302 (2006).
- 14) 與田厚志, 天野直紀: マリオネットの操作性に基づいた入力デバイス, ヒューマンイ

ンタフェースシンポジウム 2004 論文集, pp.1171-1174 (2004).

- 15) Johnson, M., Wilson, A., Kline, C., Blumberg, B. and Bobick, A.: Sympathetic Interfaces: Using Plush Toys to Direct Synthetic Characters, *Proc. CHI'99*, pp.152-158 (1999).
- 16) 小泉直也, 清水紀芳, 杉本麻樹, 新居英明, 稲見昌彦: ハンドパペット型ロボティックユーザインタフェースの開発, 日本バーチャルリアリティ学会論文誌, Vol.11, No.2, pp.265-274 (2006).
- 17) Sato, K. and Lim, Y.: Physical Interaction and Multi-Aspect Representation for Information Intensive Environments, *Proc. 9th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN2000)*, pp.436-443 (2000).
- 18) 棚瀬篤史, 神原誠之, 竹村治雄, 横矢直和: 人形型インターフェイスを用いた共有仮想環境におけるコミュニケーションの実現, 日本バーチャルリアリティ学会第 5 回大会論文集, pp.331-332 (2000).
- 19) 米澤朋子, ブライアン・クラークソン, 間瀬健二: 文脈適応型音楽生成をともなうぬいぐるみインタラクション, 情報処理学会論文誌, Vol.43, No.8, pp.2810-2820 (2002).
- 20) 黒川隆夫: ノンバーバルインターフェイス, pp.41-68, オーム社 (1994).
- 21) 畠山 誠, 西田豊明: 同調動作に基づくロボットと人間のコミュニケーション, 第 17 回人工知能学会全国大会論文集, 1D1-05 (2003).
- 22) 川尻泰司: 人形劇で遊ぼう, pp.48-51, 玉川大学出版部 (1988).
- 23) 芳賀 繁: メンタルワークロードの理論と測定, pp.77-95, 日本出版サービス (2001).
- 24) 広津千尋: 実験データの解析—分散分析を超えて, pp.203-217, 共立出版 (1992).
- 25) 山本倫也, 渡辺富夫: 音声駆動型身体引き込みキャラクタを映像に重畳合成した教育支援システム, 情報処理学会論文誌, Vol.47, No.8, pp.2769-2778 (2006).

(平成 19 年 10 月 5 日受付)

(平成 20 年 5 月 8 日採録)



大崎 浩司

1981 年生まれ。2006 年岡山県立大学大学院博士前期課程修了。現在、岡山県立大学大学院博士後期課程在学中。主にヒューマンインタラクション・ヒューマンコミュニケーションの研究に従事。2004 年ヒューマンインタフェースシンポジウム優秀プレゼンテーション賞。ヒューマンインタフェース学会会員。



渡辺 富夫 (正会員)

1955 年生まれ。1983 年東京大学大学院工学系研究科博士課程修了 (工学博士)。同年山形大学工学部情報工学科助手, 1984 年同専任講師, 1989 年同助教授。1992 ~ 1993 年米国ブラウン大学客員研究員, 1993 年岡山県立大学情報工学部情報システム工学科教授, 現在に至る。身体的コミュニケーション, ヒューマンインタラクションの研究に従事。1998, 2003 年 IEEE ROMAN, the Best Paper Award, 2001, 2002, 2004, 2005 年ヒューマンインタフェース学会論文賞等受賞。日本機械学会フェロー, ヒューマンインタフェース学会, 計測自動制御学会, 日本バーチャルリアリティ学会, 日本赤ちゃん学会, 日本子ども学会, IEEE 各会員。



山本 倫也 (正会員)

1974 年生まれ。2002 年京都大学大学院エネルギー科学研究科博士後期課程修了, 博士 (エネルギー科学)。同年岡山県立大学情報工学部助手, 2007 年同助教, 現在に至る。身体的インタラクション, ネットワークバーチャルリアリティの研究に従事。2005 年ヒューマンインタフェース学会論文賞, 2005 年情報処理学会第 66 回全国大会大会奨励賞等受賞。ヒューマンインタフェース学会, 日本バーチャルリアリティ学会, 日本原子力学会各会員。