

# 非接触計測による自己キャラクタ対面合成型実映像対話システム E-VChat の評価

石井 裕\*<sup>1</sup> 中山 志穂\*<sup>2</sup> 渡辺 富夫\*<sup>1</sup>

Evaluation of a Video Communication System ‘E-VChat’ Using the Self-character Based on the Contactless Measurement for Virtual Face-to-face Projection

Yutaka Ishii\*<sup>1</sup>, Shiho Nakayama\*<sup>2</sup>, and Tomio Watanabe\*<sup>1</sup>

**Abstract** – We previously proposed an embodied video communication system called “E-VChat” in which a CG avatar is superimposed on the other talker’s video image in a face-to-face scene in order to comprehend the talkers’ mutual interaction during remote communication. Remote talkers can communicate interactively via their own avatar with the partner’s video image, which moves based on the talker’s motion and automatically generates motion based on the on-off pattern of the talker’s voice. This system uses a headset-type motion capture device that reflects the talker’s head movements directly using acceleration and gyro sensors. Further, we demonstrated the effectiveness of the E-VChat system through a communication experiment.

In this paper, we propose an advanced E-VChat system that uses image processing to sense the talker’s head motion without wearing sensors for practical use, and discuss the development of a prototype of this system. We demonstrate the effectiveness of the self-avatar that uses the talker’s head motion and auto-generated motion based on speech input through a communication experiment in which 12 pairs of subjects are evaluated in three separate communication modes. In one mode, only head motion is used; in the second mode, motion is auto-generated based on speech input; and in the third mode, both the head motion and the auto-generated motion based on speech input is used. In addition, we confirm through a communication experiment the effectiveness of the superimposed avatar for a face-to-face floating projection based on the partner’s face angle in comparison with a bottom-center projection of the partner’s image irrespective of the partner’s angle.

**Keywords** : Embodied Interaction, Video Communication, Avatar Mediated Communication, Head Motion Tracking, Voice-driven Embodied Character

## 1. はじめに

遠隔での実時間コミュニケーション手段として、情報インフラの向上とともにビデオカメラで撮影された実映像通信が利用される機会が増えている。現在の対話用途としては遠隔会議、家族間対話、友人とのビデオチャットなどが挙げられる。実映像は対話相手の振る舞いや状況を瞬時に把握できるだけでなく、映像を用いることで会議資料や製作物、購入対象の提示や共同確認などの目的にも有効であり、実映像を介した新しい技術提案も数多く行われている<sup>[1]~[6]</sup>。ここで実映像を介した場合の対話性に着目すると、確かに対話相手や提示対象の情報を得ることは容易に行うことができるものの、対話相手との身体的なかわりを実感することは難しい。これは対話相手との視線の不一致

や背景による分離感などの空間的な問題、また解像度や描画フレームレートなどの映像品質、遅延などの通話品質といった実時間性に関する問題が影響していると考えられる。

著者らはこれらのうち空間的な問題に着目し、対話相手の実映像に対して自己アバターとしてのCGキャラクタを重畳合成することで、対話者すなわち相手と自己の対面対話感を仮想的に実現するシステムを開発し、解決を試みてきた<sup>[7],[8]</sup>。まず対話者に装着した磁気センサを用いて、自己の上半身の身体動作を高精度で再現した自己アバタを対面合成したシステムを開発し、この提案手法が有効であることをコミュニケーション実験により確認した<sup>[7]</sup>。次に小型軽量化及び可搬化のため、ヘッドセット型頭部動作検出デバイスと音声から自動生成した身体動作を組み合わせた自己アバタを開発し、対話相手の実映像に対面合成した Enhanced-VideoChat (E-VChat) システムのプロトタイプを構築するとともに、観客キャラクタの複数提示などへ応用展開するなどシステムの有効性を確認した<sup>[8]</sup>。ここで

\*1: 岡山県立大学 情報工学部

\*2: 岡山県立大学大学院 情報系工学研究科

\*1: Faculty of Computer Science and Systems Engineering, Okayama Prefectural University

\*2: Graduate School of Systems Engineering, Okayama Prefectural University

従来研究では、提案システムに対して相手映像のみ提示したシステム、または対話者自身の映像を Picture-In-Picture 手法により合成提示したシステムとの比較を行っており、対話音声に基づく身体動作自動生成モデルのみに基づく CG キャラクタを自己アバタとして対面合成したシステムと比較検討されていなかった。これは、対話者の動作計測に接触型のセンサを用いたシステム開発が行われており、センサの着脱による中断等により、対話を継続しながら各システムを連続的に比較評価することができなかつたためである。音声による自動生成モデルを用いた CG キャラクタ利用など、対話者の動作計測に因らない非接触で利用できるシステムとの比較検討は、本提案手法の応用展開に向けて重要な評価指標の一つである。

そこで本研究では、画像処理によりデバイスを装着することなく対話者の頭部動作を計測して自己アバタに反映させる実映像対話システムを開発し、音声から自動生成した身体動作に基づくキャラクタ、対話者の頭部動作のみ反映させたキャラクタ、さらにそれぞれを組み合わせたキャラクタを用いたシステムを評価した。友人同士を対象とした遠隔での自由対話によるコミュニケーション実験を行い、システム評価を行った。12組 24人に対する7段階リッカート尺度のアンケートを用いた官能評価結果により、システムの有効性が示された。

また実際のビデオチャット利用場面では、対話相手が自己を斜めや横方向から撮影している場面も多く、自己アバタを固定した配置では対応できない。そのため対話相手の撮影方向を検出し、自己アバタを相手の顔が向いている先へ対面合成する必要がある。さらに仮想空間において身体的エージェントの身体方向・登場位置が印象に影響を与える指摘がなされており、自己アバタの配置によりシステム評価が異なる可能性がある<sup>[9]</sup>。よって画面上の対話相手と自己アバタの向きが印象に与える影響を検討するため、画像処理計測による対面合成手法を提案し、斜め前方から撮影された相手映像に対してアバタを対面合成した配置と、相手の向きにかかわらずアバタを画面下方中央に固定した配置による比較評価を行った。システムを用いた自由対話における一対比較により、提案手法の有効性を確認した。

## 2. E-VChat:自己アバタ対面合成型実映像対話システム

### 2.1 コンセプト

人は対話相手の発話に対してうなずきや身振りなどの身体動作で応答反応を示し、また自己の思いを円滑に、効果的に伝えるために身体全体で表現しながら発

話を行っており、互いの発話音声と身体動作は密接にかかわり合っている。対面対話では直接的に対話相手と情報共有できるため、対話者相互の身体動作あるいは音声などに含まれる身体性が自然に同調し、互いの身体リズムや距離感などの関係をとらえながらコミュニケーションを行うことができる。遠隔においてもこのような対面対話状態を実現する手法として、対話相手のビデオ映像と自己のアバタを利用して仮想的に対面合成することが本システムの大きな特徴である(図1)。相手のビデオ映像に対して対面している感覚を実現するためには、相手映像に対して提示するアバタの大きさ、配置、方向などに加え、自分自身の身体とのかかわりなど、表現される身体的関係が重要となる。

E-VChatシステムでは、センサ装着による拘束感や画面上の違和感、可搬性に対する問題を解決するために、代役のCGキャラクタ動作の半自動化により身体動作を補完的に表現し、相手映像に重畳合成する。つまり対話者の頭部動作を忠実に再現するとともに、マイクから入力された音声に基づいて生成された、人同士の対面対話における音声と身体動作の相関関係からモデル化された身体動作を行うキャラクタを用いて、身体的なかわりをよりインタラクティブに支援するシステム構成とする。交渉や協議など細心の注意が必要な即応性の高い対話や、内容の信頼性が重視されるキャラクタを介した情報発信などの用途においては、自動生成された身体動作だけでなく、対話者自身の意思を直接的に示す必要がある場合がある。また先行研究において、自己アバタを介したコミュニケーションにおいて頭部動作が重要な役割を果たすことが確認されている<sup>[10]</sup>。そのため頭部動作をキャラクタに連動させ、肯定・否定の意思表示を反映可能にすることで、よりキャラクタを介したコミュニケーション特性を生かしたシステム構築を行うことができる。また頭部のみ計測する理由として、計測のキャリブレーションを不要とし、かつ将来的な携帯デバイス利用時の近距離撮影範囲の想定も挙げられる。

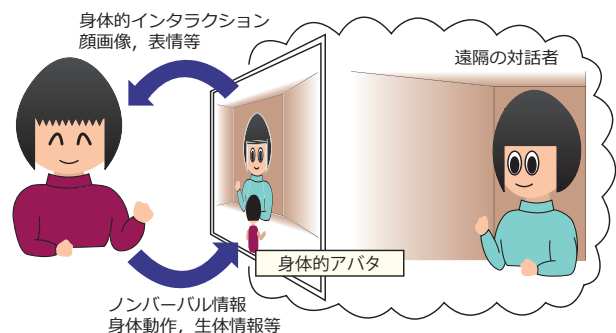


図1 E-VChatシステムのコンセプト  
Fig.1 Concept of E-VChat system.

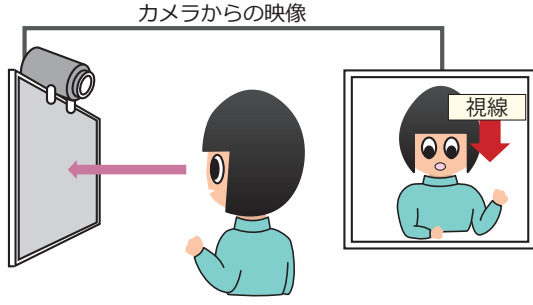


図2 カメラ位置による対話者視線

Fig.2 Talker's gaze lines based on camera positions.

ここでキャラクタ重畳合成の構成として、図2に示すようにカメラをモニタ上部に置くことで、カメラは自然とモニタを見る対話者を見下ろす角度となる。従来は大きな問題となっていた対話相手の視線のずれを利用し、キャラクタを画面の下方中央に配置することで、正面から撮影される対話相手が自己の代役キャラクタと対話しているように見える。本システムの特徴は、キャラクタを観察するのはあくまで対話者自身だけであることにある。画面上対話相手が自己の代役キャラクタを観察しているように見えるが、実際には対話相手も対話相手自身の代役キャラクタを観察しており、キャラクタの振る舞いは対話相手に伝達されない。そのためキャラクタによる効果は直接的には対話者自身にのみ起こりうるが、結果的にその効果は対話者相互のインタラクションに影響を与えられられる。

## 2.2 音声に基づく自動生成によるキャラクタ動作モデル

音声に基づくキャラクタの身体動作の自動生成モデルは、これまでに著者らが提案してきた音声の ON-OFF の時系列からうなずき、瞬き、身体動作のタイミングを推定する MA(Moving-Average) モデルを使用する<sup>[11]</sup>。音声データは 16bit 22.050kHz でサンプリングし、閾値で二値化するとともに、音節間の短時間の無音区間による発話の断片化を除去するために 133msec でハングオーバー処理を施している。

この音声に基づく身体動作は「話し手」及び「聞き手」の動作モデルによって生成される。まず「聞き手」のインタラクションモデルとして音声の ON-OFF パターンに基づく身体的引き込みモデルを導入している(図3)。このモデルは、マクロ層(Macro Stage)とマイクロ層(Micro Stage)からなる階層モデルを用いてうなずきの予測を行う。マクロ層では音声の呼気段落区分での ON-OFF 区間からなるユニット区間にうなずきの開始が存在するかを  $[i-1]$  ユニット以前のユニット時間率  $R(i)$  (ユニット区間での ON 区間の占める割合、(1) 式の線形結合で表される (2) 式の MA モデルを用いて予測する。予測値  $M_u(i)$  がある閾値を越

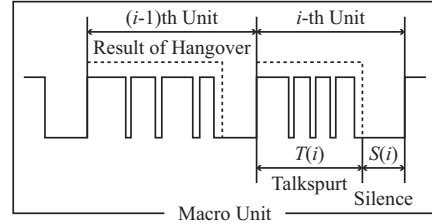
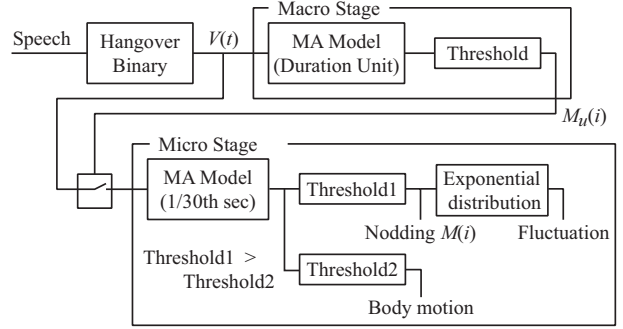


図3 対話者音声の ON-OFF パターンに基づく動作生成モデル

Fig.3 Motion-generation model based on the on-off pattern of talker's speech.

えて、うなずきが存在すると予測された場合には、処理はマイクロ層に移る。マイクロ層では音声の ON-OFF データ (30Hz, 60個) を入力とし、(3) 式を用いて MA モデルでうなずきの開始時点を推定する。予測値が閾値 (Threshold1) を越えた場合にキャラクタがうなずき動作を行う。瞬きについては、うなずきと同時に瞬きをさせ、うなずきを基点として指数分布させている。また、動物キャラクタ特有の動作として、瞬きと同じタイミングで耳を動作させている。身体動作についてもこの予測値に基づき、うなずきよりも低い閾値 (Threshold2) で各部位 (頭部, 胸部, 右肘, 左肘) のうちいずれかを選択して動作させることでうなずきと関連付けている。

$$R(i) = \frac{T(i)}{T(i) + S(i)} \quad (1)$$

$$M_u(i) = \sum_{j=1}^J a(j) R(i-j) + u(i) \quad (2)$$

$a(j)$ : linear prediction coefficient

$T(i)$ : talkspurt duration in the  $i$ -th duration unit

$S(i)$ : silence duration in the  $i$ -th duration unit

$u(i)$ : noise

$$M(i) = \sum_{j=1}^K b(j) V(i-j) + w(i) \quad (3)$$

$b(j)$ : linear prediction coefficient

$V(i)$ : voice  
 $w(i)$ : noise

また「話し手」は、対面での自由対話を計測して対話者自身の発話音声の ON-OFF パターンに基づいて、入力音声に対する頭部動作および腕や胴体などを動作させることで発話音声と関連付けた。

### 3. 非接触計測によるシステム開発

序章で述べた通り先行研究においては、従来の自己アバタは対話者自身の上半身の動作を忠実に再現した CG キャラクタ VirtualActor, および対話者自身の頭部のみ動作を再現し、両腕や胴体を含む動作は対話音声から動作を自動生成することで補完的に表現したキャラクタを用いてシステムを構築してきた。一方でジャイロセンサや加速度センサを用いた独自のセンサデバイスを装着して使用させていたため、特定のデバイス使用や、デバイスを装着することによる拘束感など、システム評価や実用性の面で様々な問題があった。そこで本研究では画像処理による非接触計測によるシステムを開発した。構築したシステム構成において対話者間は 1Gbps のイーサネット接続している。映像は Kinect for Windows (L6M-00005) で撮影し、パソコンに USB 接続する。Kinect センサを用いて、対話者の顔の 6 軸情報 (3 軸角度情報, 3 軸位置情報) を 33ms 毎に取得し、位置情報の前後方向については赤外線センサにより計測した深度情報を取得する。取得した対話者の 3 軸顔角度情報から自己アバタへの反映量を算出し、自己の頭部動作がアバタ頭部に直接反映されるように動作させる。

図 4 に上下左右の顔画像検出範囲を示す。両目・鼻・

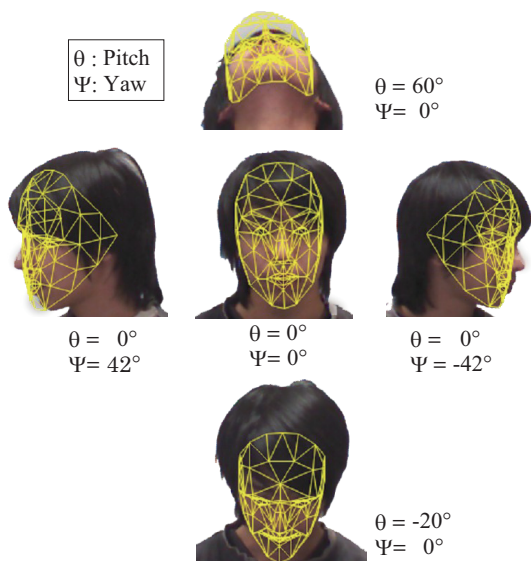


図 4 顔画像検出範囲

Fig. 4 Detection range of the face angle.

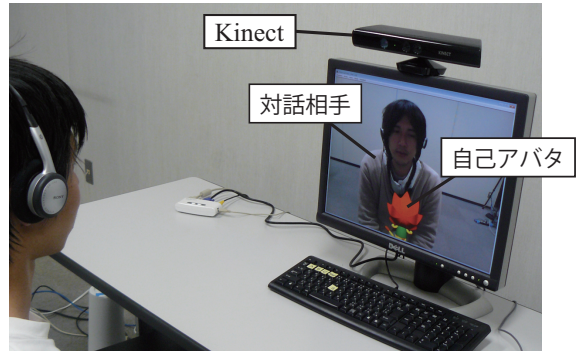


図 5 E-VChat を用いた対話場面例

Fig. 5 Example of communication scene using the E-VChat.

口などの特徴点を追跡するため、その特徴点が Kinect の撮影範囲に収まっている必要がある。本システムでは、Pitch 角が上約  $60^\circ$ 、下約  $-20^\circ$ 、Yaw 角が左右約  $42^\circ$  となっている。構築したシステムのプロトタイプを用いた対話場面例を図 5 に示す。

### 4. 自己アバタ動作比較

#### 4.1 コミュニケーション実験概要

開発したシステム評価を目的として、遠隔での日常対話による評価実験を行った。比較対象とした動作モードは、A:対話者の頭部動作のみ反映、B:頭部動作を反映させず、音声から動作を自動生成、C:頭部動作反映に加え、音声から動作を自動生成、の 3 モードとした。今回非接触による頭部動作計測を実現したことで、被験者がデバイスの着脱無く連続的に各システムを評価することが可能となったため、従来検討されていなかった音声に基づく自動生成モデルとの直接比較を行っている。被験者は 20~23 歳の男女学生で、同性の友人関係である者を 2 人 1 組とし 12 組 24 人に実験を行った。対話内容は自由対話とした。実験に使用した部屋の配置を図 6 に示す。

実験前に、実験者が各モードの違いと操作方法を説明した後、被験者に実際にシステムを 30 秒程度使用してシステムを試させた上で実験を行った。モードの

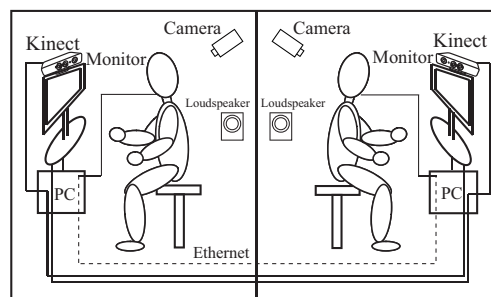


図 6 実験環境

Fig. 6 Experimental setup.

切り替えや操作は実験者が室外からスピーカ経由で指示した。被験者にはアンケート用紙をあらかじめ配布し、評価項目ごとに記入させた。ただし、各モードを使って対話している途中は、対話に集中させるためアンケート用紙に触らないよう指示した。

実験はまず、A~Cのモードのうち2つのモードを用いて、各モード2分間ずつ使用させた後、対比較によって総合的な判断で良かった方を選択させた。モードの提示順序はカウンタバランスを取り、3モードの比較で3(=3C<sub>2</sub>)回の試行を行った。次に、各モードに対し身体的コミュニケーション支援の観点から定めた6項目について7段階評価(中立0)を行った。6項目は、著者らの目指すコミュニケーションシステム評価としての「楽しさ」、身体的なかかわりの実感についての「一体感」、対話の円滑さ・促進についての「対話しやすさ」、空間的な問題解決に対する「安心感」、直感的なシステムに対する印象としての「好み」、総合的な評価としての「システムを使用したいか」とした。

また、自己アバタとして実験で使用するキャラクタは図7の選択画面を用いて被験者自身に選ばせ、キャラクタが自己の代役であることを対話者に認識させた。



図7 アバタ選択画面  
Fig.7 Avatar selection window.



図8 正面から見たキャラクタ  
Fig.8 Characters from the frontal viewpoint.

実験中は選択した同一のキャラクタを使用した。実験のために用意した7種類のキャラクタ(人型(男・女)、動物型(ライオン,羊,犬,兔),ロボット型)を図8に示す。

#### 4.2 実験結果

一対比較評価の結果を表1に示す。この結果を定量的に評価するために Bradley-Terry モデルを想定し、強さを最尤推定した結果を図9に示す。Cの頭部動作反映に加え音声に基づく動作を行うモードが最も高く評価されていることが分かる。また、Aの頭部動作のみを行うモードとBの音声に基づく動作のみを行うモードは同等の評価となった。

次に、7段階評価の結果を図10に示す。全ての項目でCが肯定的に評価されていることが分かる。Wilcoxonの符号順位検定を行った結果、AとCの間では全ての項目において有意水準1%で有意差が認められた。また、BとCの間では「一体感」「安心感」「好み」「システムを使用したいか」の項目において有意水準1%で有意差が認められ、「楽しさ」の項目においては有意差

表1 一対比較結果

Table 1 Result of pair comparison in the communication experiment.

	A	B	C	計
A		11	3	14
B	13		1	14
C	21	23		44

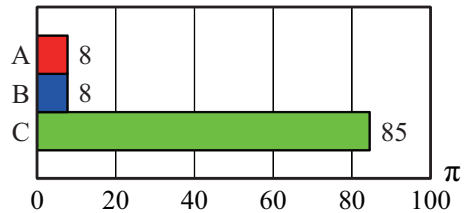


図9 Bradley-Terry モデルに基づく好ましさ  $\pi$   
Fig.9 Preference  $\pi$  based on the Bradley-Terry model.

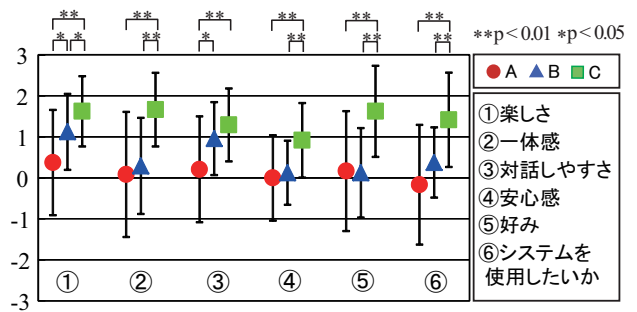


図10 7段階評価  
Fig.10 Seven points bipolar rating.

表2 キャラクタ選択結果  
Table 2 Result of avatar selections.

	Total	Boy	Girl	Lion	Sheep	Dog	Rabbit	Robot
All	24	0	1	5	7	0	6	5
Male	12	0	1	4	2	0	3	2
Female	12	0	0	1	5	0	2	3



準5%で有意差が認められた。AとBの間では「楽しさ」「対話しやすさ」の項目において有意水準5%で有意差が認められた。

対話者による各キャラクタの選択回数を表2に示す。男女ともに、人型よりも動物型やロボット型のキャラクタが好まれる傾向にあった。選択されたキャラクタと対話内容の関連については、対話ペア12組中で選択されたキャラクタが発端となったと考えられる動物の話題を取り上げた組が2組(羊, 兎)観察された。またシステムに関する内容が話題となった組が5組で、その他ほとんどが日常生活や趣味に関する内容であり、キャラクタとの関連は観察されなかった。

#### 4.3 考察

実験において、Cの「頭部動作反映に加えて音声に基づく身体動作」を行う手法が全ての項目において高く評価された。このことから、頭部動作反映と音声に基づく動作の両方が対話者に対して肯定的に働いていると考えられる。ここで既述の通り、従来研究では対話者の動作計測に接触型のセンサを用いたシステム開発が行われており、Bの音声のみに基づいて身体動作を自動生成するシステムとの比較検討を行うことができなかった。しかし本研究で非接触計測によるシステムを新たに開発したことで比較検証することが可能となり、コミュニケーション実験によって対話者動作を反映した効果を明確に示した。

また、Aの「頭部動作反映」とBの「音声に基づく身体動作」の比較から、自己アバタのうなずきや身振り・手振りといった対話音声から自動生成された身体動作が対話者に楽しさを与え、対話を促進させていることが再確認された。被験者の自由記述においても「音声に基づく身振り手振りの動作があると話しやすかった」などの意見が複数見られており、上記を裏付けるコメントが得られている。

キャラクタに関する自由記述では「自己キャラクタの表情が見たかった」というコメントが多く見られた。また、「自分のキャラクタが相手にどう見えているか気になる」などのコメントもあった。本システムでは相手側の画面には自分のキャラクタは表示されてい

いにも拘わらずこのようなコメントが見られたことから、対話者は自己キャラクタと対話相手の関係性を意識していると考えられる。

また、「人の顔を見るとキャラクタが視界に入っていない」「キャラクタにあまり目が行かなかった」などの意見も複数得られた。その一方で、「キャラクタがいることで話しやすかった」という反対の意見も多く得られた。これらの差については、対話相手を実映像であることにより、映像の情報量の多さや抽象度の高いキャラクタとの組み合わせによる整合性等により生じた可能性があり、自己アバタに対する運動主体感やエンパシー性<sup>[12]</sup>なども含めてさらに詳細に分析評価する必要がある。

#### 5. 対話者顔方向検出に基づく自己アバタ提示応用

E-VChatシステムにおいて、対話相手がモニタの上にカメラを設置して自己を正面から撮影する場合を想定し、自己アバタを画面下方中央に固定して表示していた。しかし、実際のビデオチャット利用場面では、対話者が自己を斜めや横方向から撮影する場面も少なくない。画像処理による動作計測手法を用いることで対話相手の撮影方向を検出することが可能となり、自己アバタを相手の顔が向いている先へ対面合成することで、互いのインタラクションが捉えやすい状態を作ることができると考えられる。

##### 5.1 アバタ配置手法の検討

前章で開発した対話者顔方向検出に基づくE-VChatシステムでは対話者の顔の位置や向きを計測することを可能としており、これらの情報を用いることで対話相手の視線の先を推定することが可能である。そこで、対話相手が自己を不特定の方向から撮影している場合においても自己アバタを相手の顔向きの先に対面合成する手法を検討する。

Kinectセンサを用いて対話相手の顔の位置や向きを計測して対話相手の視線の先を推定し、推定位置に自己アバタを移動させ、画面上の対話相手と対面するように配置する。これにより、対話相手が正面を向いていない場合においてもアバタを介して自己と相手とのインタラクションを把握しやすい状況をつくることができる。自己アバタの移動には、対話相手の顔の3軸角度情報と位置情報を用いる。対話相手の顔の3軸角度情報の内、Yaw角に対応して自己アバタをy軸回転させ、同時にx軸移動を行う。また、顔の位置情報から対話相手と対面するために適切な位置を算出しアバタを移動させる。移動は一定の範囲で画面下部を平行移動し、その範囲外は対話相手を基準に楕円軌道を描くよう移動させる。

相手視線の先へ自己アバタを移動させる手法につい

ては、次の3つのモードが選択可能である。

- A：アバタを常に適切な場所へ移動
- B：対話者自身がキー入力した際に移動
- C：一定時間同一方向を向いている場合に移動

Aの手法は自己アバタを常に相手視線の先に配置することができ、首振りなどの瞬間的な動作にも反応する。常に相手映像にアバタを対面合成することができるが、アバタが常に動き回ることになるため、その動作が対話を阻害することも考えられる。Bの手法では対話者自身が好きなときにアバタを移動させることができるが、操作自体が対話の妨げとなることが考えられる。よって対話相手が一定時間同一方向を向いている場合のみアバタを移動させるCの手法が、対話相手が主に向いている先へ自然にアバタを移動させることができると考える。

### 5.2 アバタ配置評価

E-VChatシステムを用いた実映像対話において、正面以外を向いている対話相手映像に自己アバタを対面合成することは互いのインタラクション把握を助けるのに有用であると考えられる。しかし、自己アバタを画面の中央から移動させることは、キャラクタが自己の代役であるという認識を薄れさせてしまい対話者に好まれない可能性も考えられる。

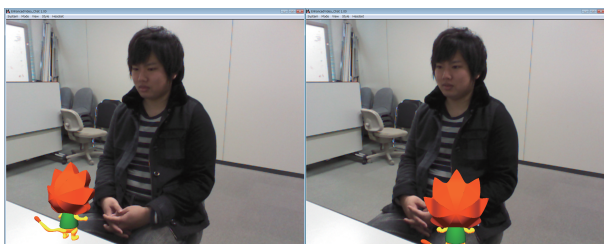
そこで、対話相手の映像に自己アバタを対面合成することの有効性を確認するために、前章の被験者を対象に、自己アバタの配置手法について評価実験を行った。対話内容は前章同様に自由対話とした。

実験では次の2つの自己アバタ配置モードを用意した。

- $\alpha$ ：自己アバタを相手の視線の先に配置
- $\beta$ ：自己アバタを画面下方中央に配置

各配置の提示例を図11に示す。

実験では、2つのモードに対して2分間ずつ使用させた後、総合的に良かった方を選択させた。一対比較評価の結果を図12に示す。 $\alpha$ の「自己アバタを相手の視線の先に配置する手法」が96%（24人中23人）の被験者に選択された。



$\alpha$ : 自己アバタを相手の視線の先に配置  $\beta$ : 自己アバタを画面下方中央に配置

図11 アバタ配置例

Fig. 11 Examples of talker's avatar arrangements.

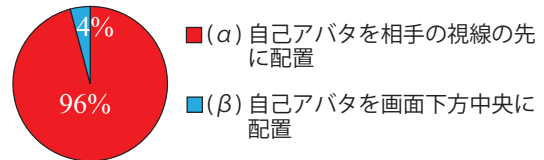


図12 アバタ配置方法の一対比較

Fig. 12 Result of paired comparison for the avatar's arrangements.

### 5.3 考察

評価実験では、非接触計測による特性を活かして対話相手の顔方向を推定することで、撮影角度に因らず自己アバタを相手の視線の先に配置する手法が高く評価された。自由記述においても「キャラクタが相手の視線の先にいる方が一体感があつた」など、肯定的な意見が多く見られた。このことから、E-VChatを用いた実映像対話においては、本研究で提案している相手映像に自己アバタを対面合成する手法が有効であると考えられる。

## 6. おわりに

本論文では、非接触による画像処理計測により、自己の頭部動作を自己アバタに反映するE-VChatシステムを開発し、音声から自動生成した身体動作に基づくキャラクタ、対話者の頭部動作のみ反映させたキャラクタ、さらに両者を組み合わせたキャラクタを用いたシステムを評価した。コミュニケーション実験により、頭部動作反映に加えて対話音声から動作を自動生成して補完的に表現するモードが、身体的コミュニケーション支援の観点から定めた6項目「楽しさ」、「一体感」、「対話しやすさ」、「安心感」、「好み」、「システムを使用したいか」の全てにおいて高く評価された。

また、E-VChatにおいて対話相手が不特定の方向から自己を撮影している場合においても、自己アバタを相手の視線の先に自動で配置することでインタラクション把握を支援する手法を提案した。さらに自己アバタ配置位置について比較検討するために評価実験を行い、提案した相手視線の先に自己アバタを対面合成する手法が有効であることを示した。

近年では画像処理による身体動作計測デバイスが一般的に利用可能となりつつあり、カメラによる動画撮影機能を搭載した携帯端末などを用いた本システムの応用展開が期待できる。

謝辞

本研究の一部は科学研究費(22300045, 24118707)の助成を受けたものである。

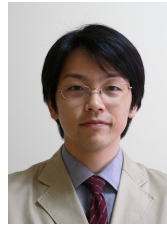
参考文献

- [1] 森川 治, 橋本 佐由理, 前迫 孝憲: 仮想的な抱擁を取り入れた遠隔カウンセリングシステム, 日本バーチャルリアリティ学会論文誌, Vol.14, No.1, pp.3-10 (2009).
- [2] 磯 和之, 伊達 宗和, 高田 英明, 安藤 康子, 松浦 宣彦: 視線の向きを表現可能な 2 画面積層表示を用いたテレビ会議システムの提案, 情報処理学会論文誌, Vol.52, No.3, pp.1224-1233 (2011).
- [3] Ishii, R., Ozawa, S., Mukouchi, T., and Matsuura, N.: MoPaCo: Pseudo 3D Video Communication System, Human Interface and the Management of Information, Symposium on Human Interface, HCII 2011, LNCS 6772, pp. 131-140 (2011).
- [4] Murata, T., Hattori, M., and Shibuya Y.: Visual Feedback to Reduce Influence of Delay on Video Chatting, Human Interface and the Management of Information, Symposium on Human Interface, HCII 2011, LNCS 6772, pp. 157-164 (2011).
- [5] Rae, I., Takayama, L., and Mutlu, B.: The Influence of Height in Robot-Mediated Communication, Proc. of the 8th ACM/IEEE International Conference on Human-Robot Interaction, pp.1-8 (2013).
- [6] 前田 和甫, 長谷川 洗平, 小森 政嗣: キネクトセンサーを用いた 3 次元ビデオ対話システム, ヒューマンインタフェース学会研究報告集, Vol.14, No.3, pp.183-186 (2012).
- [7] 石井 裕, 渡辺 富夫: VirtualActor を対面合成した身体的ビデオコミュニケーションシステム, ヒューマンインタフェース学会論文誌, Vol.5, No.2, pp.225-234 (2003).
- [8] 石井 裕, 高田 友寛, 渡辺 富夫: E-VChat: 頭部動作運動型音声駆動身体引き込みキャラクタを対面合成した実映像対話システム, ヒューマンインタフェース学会論文誌, Vol.14, No.4, pp.467-476 (2012).
- [9] 鈴木 聡, 森島 泰則, 中村 美代子, 槻館 尚武, 武田 英明: 身体化エージェントの身体方向・登場位置がユーザに与える影響, 知能と情報, Vol.20, No.4, pp. 513-525 (2008).
- [10] 石井 裕, 渡辺 富夫: 聞き手の VirtualActor の頭部動作を矛盾的に止めた身体的コミュニケーションの合成的解析, ヒューマンインタフェース学会論文誌, Vol.4 No.3, pp.9-16 (2002).
- [11] 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).
- [12] 竹内 勇剛, 上杉 繁, 寺田 和憲, 片上 大輔: インタラクシヨンのミニマムデザイン, ヒューマンインタフェース学会論文誌, Vol.15, No.1, pp.1-14 (2013).

(2013 年 10 月 18 日受付, 2014 年 3 月 5 日再受付)

著者紹介

石井 裕 (正会員)



1998 年岡山県立大学情報工学部卒。2003 年同大学院情報系工学研究科博士後期課程修了。博士(工学)。同年科学技術振興事業団戦略的創造研究推進事業(CREST)研究員。同年神戸大学情報基盤センター助教。2011 年岡山県立大学情報工学部助教。2013 年同准教授。主にヒューマンインタラクシヨンの研究に従事。2000 年ヒューマンインタフェース学会学術奨励賞, 2002 年・2004 年ヒューマンインタフェース学会論文賞, 2003 年情報処理学会大会奨励賞等受賞。情報処理学会, 日本認知科学会, 日本人間工学会, 日本子ども学会, 計測自動制御学会各会員。

中山 志穂 (学生会員)



2012 年岡山県立大学情報工学部卒。2014 年岡山県立大学大学院情報系工学研究科博士前期課程修了。主にヒューマンコミュニケーション, 身体的インタラクシヨンの研究に従事。

渡辺 富夫 (正会員)



1983 年東京大学大学院工学系研究科産業機械工学専攻博士課程修了(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。1993 年岡山県立大学情報工学部情報システム工学科教授。1992 年~1993 年米国ブラウン大学ブラドリ・リサーチセンター客員研究員。ヒューマンインタラクシヨンの研究に従事。1998 年・2003 年・2012 年 IEEE RO-MAN, the Best Paper Award, 2001 年・2002 年・2004 年・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, 情報処理学会, 計測自動制御学会, 日本バーチャルリアリティ学会, 日本子ども学会, 日本赤ちゃん学会, 日本ロボット学会, IEEE 各会員。