

情動表現を即時表示するタイピング駆動型 身体的引き込みキャラクタチャットシステムの開発

Development of a Typing-Driven Embodied Entrainment Character Chat System with Immediate Display of Affective Expression

岩佐厚郎^{1*} 小原瑞希¹ 渡辺富夫² 石井裕²
Atsurou Iwasa,¹ Mizuki Kohara,¹ Tomio Watanabe,² Yutaka Ishii²

¹ 岡山県立大学大学院 情報系工学研究科

¹ Graduate School of Systems Engineering, Okayama Prefectural University

² 岡山県立大学 情報工学部

² Faculty of Computer Science and Systems Engineering, Okayama Prefectural University

Abstract: In recent years, communication of text on Internet is popular. In face-to-face verbal communication, a listener's movements such as nodding and body motions are interactively synchronized with a speaker's speech. This entrainment is an important role in embodied interaction and communication. We have developed a typing-driven embodied entrainment character chat system called "InterChat" which generates automatically the bodily movement from the rhythm of typing, and have shown the effectiveness of the system for supporting chat communication. Conventional InterChat displays affective expression from the meaning of words to input. However, time lag occurs between message and affective expression. In this paper, we propose two methods displaying affective expression from the meaning of words to input without delay. And, we develop an advanced typing-driven embodied entrainment character chat system by using the methods.

1 はじめに

人の対面コミュニケーションでは、言葉による言語情報だけでなく、音声に対するうなずき、瞬き、身振り手振りといった身体動作などの非言語情報を相互に同調して、対話者が相互に引き込み合うことで円滑なコミュニケーションを行っている。この身体的引き込みが対話時の一体感を生み、対話相手とのかわわりを実感させている [1]。

しかし近年では情報技術の発達により、趣味や仕事など、多くの場面でパソコンが用いられ、キーボードを介した入力手段は一般的なものとなった。加えて、ネットワークの発展に伴い、ブログやチャット、SNS といった、テキストベースのコミュニケーションが盛んに使用されるようになった [2]。中でもチャットは、その利便性やリアルタイム性から、オンラインゲーム等にも導入され、自分のアバタを通して複数の人と同じ時空間を共有したコミュニケーションが実現されている。このようにチャットの利用者は多く、今後もチャットを用

いたコミュニケーション手段がますます増えると見込まれる。

この現状と身体的インタラクションの重要性から、著者らはこれまでに、タイピングのリズムを音声に見立て、タイピング情報から話し手と聞き手のコミュニケーション動作を自動生成するタイピング駆動型身体的引き込みキャラクタを用いて、コミュニケーション支援を行うチャットシステム InterChat を開発している [3]。さらに、入力メッセージの同期表示を可能にし、テキスト情報から対応するキャラクタ動作を適用する技術を開発し、コミュニケーション支援の有効性を示してきた [4][5]。しかしながら、従来の InterChat では、メッセージの送信後にテキストに対応したキャラクタ動作が画面提示され、テキストとテキスト情報から対応するキャラクタ動作の提示の間にずれが生じていた。また送信メッセージに意味のある単語が複数含まれていた場合、最初に検出された単語に対応する動作のみが提示されていた。

そこで本研究では、従来の InterChat において、テキストに対応したキャラクタ動作の提示時間の差を縮め、入力に合わせて即時表示する手法を考案し、より高いコミュニケーション支援効果を目指した新たなタ

*連絡先： 岡山県立大学
〒719-1197 岡山県総社市窪木 111
E-mail: 011_iwasa@hint.cse.oka-pu.ac.jp

タイピング駆動型身体的引き込みキャラクタチャットシステムを開発している。

2 コンセプト

本システムのコンセプトを図1に示す。本システムは、参加者が代役キャラクタを通じて仮想空間に入り込み、チャットを行う3Dキャラクタチャットシステムである。仮想空間上のキャラクタは参加者のタイピングに基づき、話し手および聞き手の動作をすることでかかわり合い、相互に引き込み合うコミュニケーション場を生成する。タイピング時、自分のキャラクタは身振りなどの話し手の身体動作を行うとともに入力中のメッセージを吹き出しとして発言し、相手のキャラクタはうなずきなどの聞き手の身体動作を行う。また、テキスト内にある特定の文字列を認識し、キャラクタにその文字列に対応した動作を行わせている。これによって、使用者の思いを表現することができる。

本システムでは、入力メッセージに合わせたタイミングで文字列に対応した動作を同期して提示する。これによって、より自然な会話の動作をしている印象を与え、参加者とキャラクタとの一体感や場の盛り上がりが高まる。

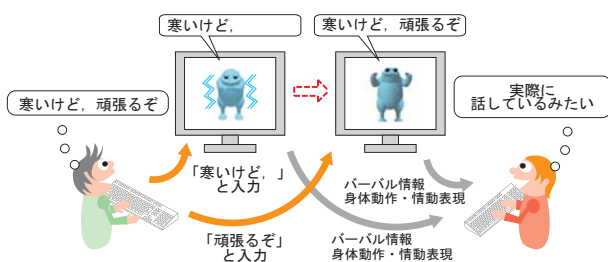


図 1: コンセプト

3 身体的コミュニケーション技術

著者らはこれまでに、人の対面コミュニケーション時の身体的リズムの引き込み現象に着目して、会話音声のみを入力としてキャラクタの豊かなコミュニケーション動作を自動生成するインタロボット技術iRT(InterRobot Technology)を開発している[6]。iRTはコミュニケーション時の発話音声と身体動作との関係をモデル化することで発話音声からコミュニケーション動作を自動生成し、身体的リズムの引き込みによりインタラク션을円滑にして、コミュニケーションを支援する技術である。

iRTを用いて、うなずきや身振り、手振りなどの身体性を物理メディアであるロボット(InterRobot)や電子メディアのキャラクタ(InterActor)などの情報機

械に導入することで、遠隔地においても互いの身体性を共有し対話者相互に相手とのかかわりを知覚できる身体的コミュニケーションシステムを開発し、コミュニケーション支援での有効性を示してきた。

タイピングのキー入力データを音声のデータに見立ててiRTの入力とすることで、タイピングからもInterActorのようなコミュニケーション動作の生成が可能であり、キャラクタを介したチャットコミュニケーションにおいても、iRTのメカニズムを導入することで、一体感のあるチャットコミュニケーションを実現できる。

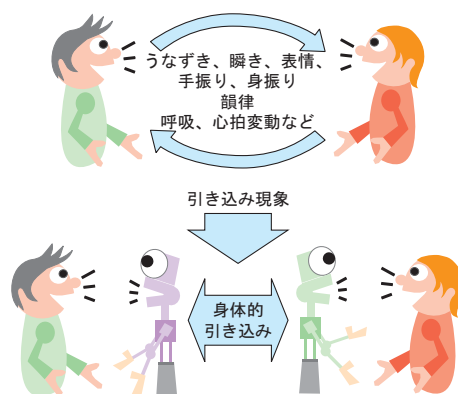


図 2: 身体的コミュニケーション

4 InterChat

4.1 InterChatの概略

iRTを拡張し、タイピングからキャラクタの身体動作を生成することで、身体的リズムの共有による引き込みによって対面コミュニケーションでの一体感を文字ベースのコミュニケーションで実現することが可能になる。

InterChatの概略図を図3に示す。InterChatは、仮想空間内にチャット参加者の代役となるキャラクタを配置し、互いのキャラクタの身体動作によって参加者間で身体的リズムを共有しながらチャットを行うコミュニケーションシステムである。

システムは参加者毎にPC(Microsoft Windows)、ディスプレイ、キーボードによって構成され、互いにネットワークで接続されている。PC同士の接続はMicrosoft DirectPlayを用いてpeer-to-peerの接続を行うことで、使用者間で直接タイピング情報を送受信している。仮想空間はMicrosoft DirectX9.0bを用いて作成した。キャラクタはテーブルを囲むように向い合せて配置することで複数の参加者でコミュニケーションができる。使用者自身の画面上では、タイピング時、自己の代役キャラクタの上に入力中のメッセージ及び変換中の文字列

を吹き出しとして表示することで話し手であることを明示し、そのキャラクタがタイピングのリズムに合わせた身振り・手振りを交えて話しているように動作する。代役キャラクタが身振り・手振りを伴って情報を伝えることで、効果的に情報を伝達することが可能である。同時に、他の参加者の代役キャラクタが聞き手としてのコミュニケーション動作を生成し提示することで、参加者は仮想空間内でのインタラクションを確認することができる。参加者が相互にインタラクションを把握することで、円滑なコミュニケーションが実現される。

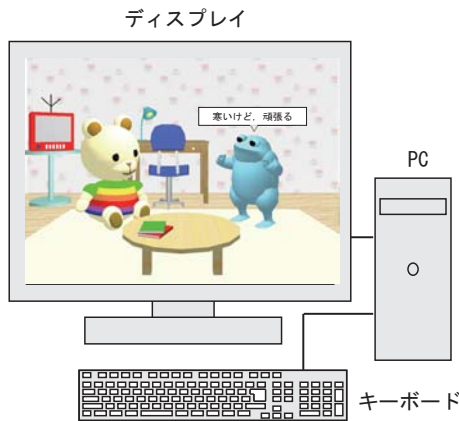


図 3: InterChat の概略図

4.2 聞き手の動作生成モデル

InterChat の聞き手のインタラクションモデルを図 4 に示す。InterChat の聞き手動作モデルでは、タイピングを音声入力に見立て、30Hz で取得したキー入力の ON-OFF データにハングオーバー処理 (各 ON 区間を伸ばし、定めた時間以下の OFF を ON に置換する操作) を施したタイピングデータを iRT の音声入力に対応させ、予測値が閾値を越えた場合にキャラクタにうなずき反応を行わせている。

うなずきの予測モデルはマクロ層とミクロ層からなる階層モデルである。マクロ層ではキー入力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、(2) 式) の線形結合で表される MA (Moving-Average) モデルを用いて予測する ((1) 式)。予測値 $M_u(i)$ がある閾値を越えて、うなずきが存在すると予測された場合には、処理はミクロ層に移る。ミクロ層ではキー入力の 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。予測値が閾値を越えた場合にはうなずきさせる。

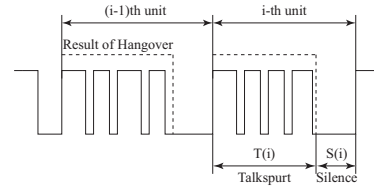
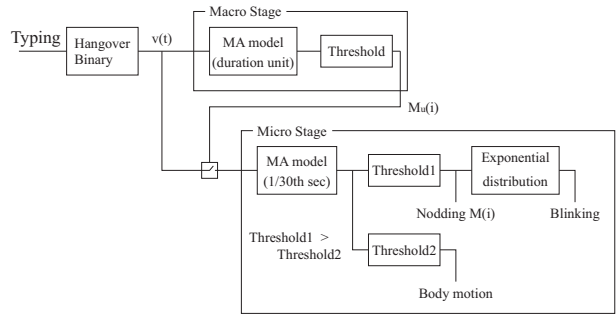


図 4: InterChat の聞き手のインタラクションモデル

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

$$R(i) = \frac{T(i)}{T(i) + S(i)} \quad (2)$$

$a(j)$: 予測係数

$T(i)$: i 番目ユニットでの ON 区間

$S(i)$: i 番目ユニットでの OFF 区間

$u(i)$: 雑音

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

$b(j)$: 予測係数

$V(i)$: タイピングデータ

$w(i)$: 雑音

さらに同モデルにより、うなずきの予測値から、うなずきよりも低い閾値でキャラクタの各部位 (頭部、胸部、右肘、左肘) のうちいくつかを選択して動作させている。瞬きについては、対面コミュニケーション時における瞬き特性に基づいてうなずきと同時に瞬きさせ、それを基点として指数分布させる。

また、タイピングの ON-OFF データを取得する際、押下キーの種類を判別することが可能なため、変換やスペース挿入時に使用する Space キーや、メッセージ送信時に使用する Enter キーを判別し、その行為を認識する。これらの行為はメッセージの区切りで行われるため、うなずき反応に適したタイミングである。このように、タイピングのリズムに基づいた動作に加え、利用者の行為に対する反応動作を合わせて行わせることで、より豊かなコミュニケーション動作を実現する。

4.3 話し手の動作生成モデル

話し手の動作モデルについても同様に iRT を適用し、タイピングの ON-OFF データから身体全体の動作を予測し、キャラクタの各部位を動作させる。

また、日常会話における身体動作が話の盛り上がりに伴って大きくなることから [7]、タイピングのスピードをチャットの盛り上がりとしてキャラクタの腕部動作に反映させた。ある時点におけるタイピングの盛り上がりを示すタイピング活性量 $u(i)$ を、(4) 式のように定義し、キー入力を短い間隔で連続して打つほど大きな値をとるようにした。このタイピング活性量が、数段階に定めた閾値を越える毎に、腕部の変位量を大きくする。図 5 は実際にタイピングを行った際の $u(i)$ の変化量である。継続したタイピングに $u(i)$ が対応しているのがわかる。

$$u(i) = \sum_{j=1}^K (K + 1 - j)x(i - j) \quad (4)$$

K : タイピング活性量定数

$x(i)$: タイピングデータ

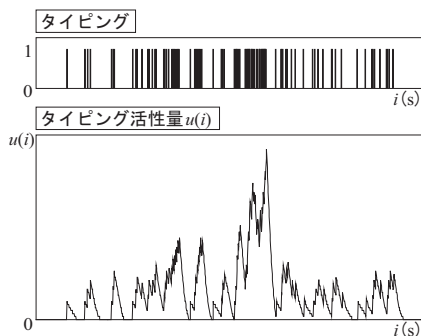


図 5: タイピング活性量

4.4 情動表現の動作生成

InterChat ではテキスト情報を用いるため、メッセージの内容を認識し、言語情報からも人の動作・情動表現を生成することが可能であり、さらなるコミュニケーション支援効果が期待できる。そこで、聞き手、話し手の動作に加え、メッセージ内にある特定の文字列を認識し、対応した意味動作をキャラクタの身体動作に関連付けている。チャットにおいて文字認識は容易であり、タイピング情報による身体動作を自然に行いながら即座にキャラクタの動作に反映させることが可能である。例えば「おはよう」や「こんにちは」といった、特定の文字列を認識し、対応した動作を生成できる。さらに、動作に加えて、感情や状態を表すシンボ

	意味動作	シンボル表示	意味動作 + シンボル表示
動作			
文字列	やあ バイバイ	好き ハート	寒い ブルブル

図 6: シンボルマークの表示

ルマークを表示することで、使用者の情動を効果的に表現することができる(図 6)。

文字列に対応した動作はデータベース化しており、追加や削除などの管理を容易に行うことが可能である。これらの文字列は感情表現辞典 [8] を参考に選出した。現在、データベースには、1000 種以上の文字列とそれに対応した 110 種類の動作が登録されている。動作・情動表現の例を図 7 に示す。

文字列	こんにちは すみません	やあ バイバイ	なるほど 思いついた	寒い 凍える
動作				
文字列	悩む 照れる	いいえ 違います	好き かわいい	驚く びっくり
動作				

図 7: 文字列に対応した動作の例

4.5 対話時の画面表示

タイピングの ON-OFF データやメッセージ内容は、ネットワークでつながったほかの参加者と常時共有されており、すべての参加者が、入力のリズムや内容といった入力情報とキャラクタの身体動作の時間的関係を保って確認することが可能となる。本システムでは、先行研究により一体感や場の盛り上がりに対して有効性が示された「メッセージを入力する度に送信し、ほかの参加者の画面に入力中のメッセージを吹き出しとして同期表示する」手法を採用している。

メッセージの送信側が入力中の文字と受信者側に表示される画面の関係の例を図 8 に示す。図では送信者のタイプした変換前の文章が、受信者側のディスプレイに代役キャラクタの台詞として表れる様子を示している。

状態	入力	変換	確定	入力	確定	送信
送信側 入力文字	がんばる	頑張る	頑張る	頑張る!!	頑張る!!	頑張る!!
受信側 表示画面						

図 8: メッセージ表示手法

状態	入力	入力	入力	変換	確定	入力	入力	入力	変換	確定	送信
表示文字	さむ	さむい	さむいけど、	寒いけど、	寒いけど、	寒いけど、	寒いけど、	寒いけど、	寒いけど、	寒いけど、	寒いけど、
身体動作	通常										
	確定時										
	入力時										

図 9: 動作提示手法

5 情動表現の即時表示手法

文字列に対応した動作を提示する方法を発展させるために、動作の提示タイミングに着目した新たな動作の提示手法を提案する。従来の InterChat での動作提示手法として使用されていた、メッセージの送信時に文字列に対応した動作を提示する提示手法を「通常動作」とし、それに加えて「確定時動作」、「入力時動作」の2つの提示手法を考案した。これらの3つの動作提示手法の比較を図9に示す。

確定時動作では、入力中のメッセージが確定される度にそれまで入力していた文章の認識を行い、特定の文字列が文中にあった場合はその文字列に対応した動作を提示する。「寒いけど、頑張るぞ」といった文章を入力する場合に、「寒いけど、」まで入力して確定し、その後「頑張るぞ」と入力して送信するようにメッセージの送信までに文章を複数回に分けて入力することで、1つの文章中で複数回の動作提示を行うことが可能となる。文章を確定するまで動作を行わないので、打ち間違いにより意図しない動作が提示される事を抑えつつ、情動表現の提示を従来よりも早めることが可能である。

入力時動作では、メッセージが入力される度にそれまで入力していた文章の認識を行い、特定の文字列が文中にあった場合はその文字列に対応した動作を提示する。確定時動作での例のように1つの文を複数回に分けて入力しなくても、1つの文章中で複数回の動作提示を行うことが可能となっている。メッセージと情動表現の提示時間の差は3つの提示手法の中では最も小さくなり、自然な会話の動きに近づけることができる。

6 おわりに

本研究では、タイピング駆動型身体的引き込みキャラクターチャットシステム InterChat における、文字列に対応した情動表現の提示手法に着目し、情動表現を即時表示する新たな動作提示手法の考案を行った。さ

らに、今回考案した動作提示手法を取り入れ、従来の InterChat との選択を可能とした新たな InterChat を開発した。

本システムはデータベースに登録されている文字列とのマッチングにより情動表現の提示を行うが、データベースの内容と使用している動作の提示手法によっては、文章の内容から期待される動作が正しく提示されなかったり、逆に会話の内容とは関係の無い動作が誤って提示されてしまう可能性がある。このような誤動作が起こらず、最も効果的に動作するようにデータベースの内容を見直し、これらのシステム改良を行う必要があると考えられる。

謝辞

本研究の一部は、JSPS 科研費 26280077 の助成を受けたものである。

参考文献

- [1] 渡辺 富夫: 身体的コミュニケーションにおける引き込みと身体性 - 心が通う身体的コミュニケーションシステム E-COSMIC の開発を通して、ベビーサイエンス, Vol.2, pp.4-12 (2003)
- [2] 例えば Second Life (Linden Research, Inc.) や LINE PLAY (LINE Corporation) など
- [3] 服部 憲治, 渡辺 富夫, 山本 倫也: タイピング駆動型身体引き込みキャラクタチャットシステム InterChat, ヒューマンインタフェース学会論文誌, Vol.15, No.4, pp.389-398 (2013).
- [4] 畔地 耕太, 渡辺 富夫, 山本 倫也: 入力情報を同期表示するタイピング駆動型身体引き込みキャラクタチャットシステム, ヒューマンインタフェースシンポジウム 2009 DVD 論文集, pp.793-796 (2009)

- [5] 野條 諒, 渡辺 富夫, 檀原 龍正: タイピング駆動型身体引き込みキャラクタチャットシステムにおける入力情報同期表示手法, ヒューマンインタフェースシンポジウム 2010 DVD 論文集, pp.1063-1066 (2010)
- [6] Tomio Watanabe: Human-entrained Embodied Interaction and Communication Technology, Emotional Engineering, Springer, pp.161-177 (2011)
- [7] 渡辺 富夫, 大久保 雅史, 小川 浩基: 発話音声に基づく身体的インタラクションロボットシステム, 日本機械学会論文集 (C編), Vol.66, No.648, pp.251-258 (2000)
- [8] 中村 明, 感情表現辞典, 東京堂出版, 1993.