

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 における情動表現提示タイミングの評価

Evaluation of Affective Expression Presentation Timing in a Typing-Driven Embodied Entrainment Character Chat System

岩佐厚郎^{1*} 服部憲治¹ 渡辺富夫² 石井裕²
Atsurou Iwasa,¹ Kenji Hattori,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 the text through the 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. The conventional InterChat displays affective expression from the meaning of words to input. However, the time lag occurs between the message and affective expression. In this paper, we develop an advanced typing-driven embodied entrainment character chat system by using two methods to reduce the time lag. The effectiveness of the system is demonstrated by a sensory evaluation experiment conducted in chat communication scenario.

1 はじめに

人の対面コミュニケーションでは、音声による言語情報だけではなく音声に対するうなずきや瞬き、身振り、手振りといった身体動作などの非言語情報が相互に同調して、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。対話時においてこの身体的引き込みが対話者同士の一体感を生み、対話相手とのか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ており、ブログやチャット、SNS といった遠隔地間での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。その中でもチャットは、利便性やリアルタイム性から Second Life [2] や LINE PLAY [3] 等のオンラインゲームなどに導入されており、自身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。

この現状とコミュニケーションにおける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ることで、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 [4]。さらに、入力メッセージを同期表示する手法、テキスト情報から対応するキャラクタ動作を提示する手法を開発し、チャットコミュニケーション支援の有効性を示してきた [5]。しかしながら、従来の InterChat では、テキストに対応したキャラクタ動作がメッセージの送信後に提示されており、テキストの表示との間にずれが生じ、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しか提示されない、などの問題点があった。そこで、テキストの表示と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た [6]。本研究では、このシステムを用いてチャットコミュ

*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

ニケーション実験を行い、好ましいテキストに対応したキャラクタ動作の提示タイミングを評価することで、コミュニケーション支援における開発システムの有効性を示している。

2 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 (Inter-Robot Technology) を開発している。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ションを円滑にして、コミュニケーションを支援する技術である。この 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、遠隔地間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いる [7]。

3 キャラクタチャットシステム

3.1 コンセプト

本研究で用いるシステムのコンセプトを図 1 に示す。InterChat 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、手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、メッセージの送信時、テキスト内にある特定の文字列を認識し、キャラクタに文字列に対応した動作を行わせている。これによって、使用者の細かな感情をキャラクタに表現させている。

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。これにより、自然な会話の動作をしている印象を与え、参加者と代役キャラクタとの一体感や場の盛り上がりが高まり、円滑なコミュニケーションが実現される。

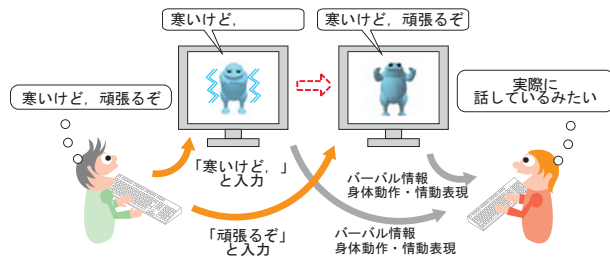


図 1: コンセプト

3.2 InterChat の概略

InterChat の概略図を図 2 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かい合わせて配置することで、複数の参加者がコミュニケーションを行うことができる。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およ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。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把握することができる。



図 2: InterChat の概略図

3.3 聞き手の動作生成モデル

InterChat では、タイピング中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る。InterChat の聞き手のインタラクションモデルを図 3 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かせる。

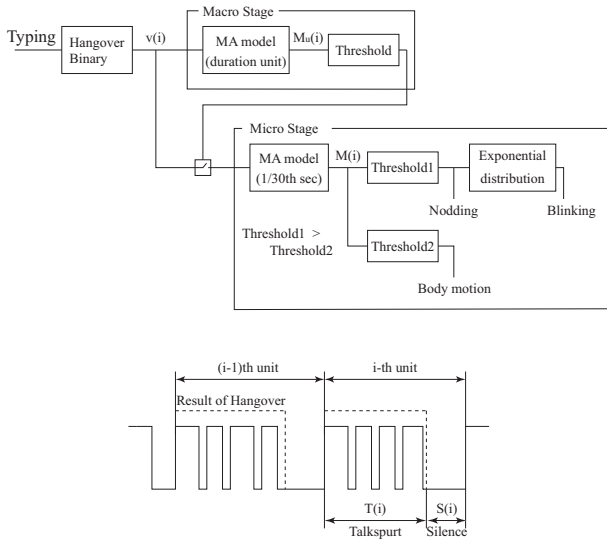


図 3: InterChat の聞き手のインタラクションモデル

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

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

$a(j)$: 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、胴部、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、押した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、その行為を認識している。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と考え、予測値が閾値を越えた場合と同様にキャラクタにうなずき反応を行わせている。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している。

3.4 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 [8]、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 4 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

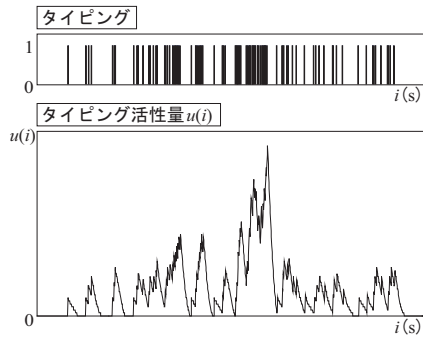


図 4: タイピング活性量

3.5 情動表現の動作生成

InterChat では聞き手、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することができる。さらに、動作に加えて、感情や状態を表すシンボ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。

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。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。これらの文字列は感情表現辞典 [9] を参考に選出した。InterChat での動作・情動表現の例を図 5 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 5: 文字列に対応した動作の例

3.6 対話時の画面表示

メッセージ入力時、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他の参加者と常時共有されている。これにより、全ての参加者が、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關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の有効性が示された「メッセー

ジを入力する度に送信し、他の参加者の画面に吹き出しとして同期表示する」手法を採用している。メッセージ送信側の入力状態と表示される画面の關係の例を図 6 に示す。共有されたメッセージは各参加者の PC に表示され、全てのシステム使用者が同様の入力状態を確認することができる。InterChat の使用風景を図 7 に示す。

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 6: メッセージ表示手法



図 7: InterChat の使用風景

3.7 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、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「通常動作」に加え、新たに「確定時動作」、「入力時動作」の 2 つの提示手法を開発した。

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。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。これらの 3 つの動作提示手法の比較を図 8 に示す。

4 チャットコミュニケーション実験

4.1 実験方法

本研究では、好ましい情動表現の提示タイミングを検証するために、被験者 2 人 1 組で別々の部屋に分かれ、「A:通常動作」、「B:確定時動作」、「C:入力時動作」の 3 つのモードを使用してチャットコミュニケーションを行った。被験者には最初に、システム全体と各モー

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 8: 動作提示手法

ドの特徴を説明した。その後別々の部屋に分かれ、システムに慣れるために各モードを3分ずつ使用させてから、2つの評価法を用いて比較させた。最後に実験に関する感想を自由記述形式で記入させた。被験者は18~24歳の男女学生12組24人(同性同士)で、普段からパソコンを使用しておりタイピングにある程度慣れた学生である。

評価法 1: 一対比較

A~Cのモードからランダムに2つを一対比較させ、これを計3(=3C2)回繰り返した。まず、1つのモードでチャットをした後に別のモードによるチャットを行い、2モードを使用した後に「総合的によい」という観点からどちらがよかったかアンケート用紙に記入させた。対話者同士は同じモードを使用してチャットし、1つの比較につき対話時間を4分(1つのモードにつき2分)とした。

評価法 2: 7段階評価

A~Cのモードを「楽しさ」、「好み」、「チャットしやすさ」、「使用したい」、「一体感」、「思いが通う」、「盛り上がる」の7項目について評価させた。各モードをランダムな順番で使用し、その後、各モードがどの程度その項目に当てはまるかを7段階(中立0)でアンケート用紙に記入させた。対話者同士は同じモードを使用してチャットし、1つのモードにつき対話時間を3分とした。

4.2 実験結果

一対比較の結果を表1に示す。表中の数字は各行のモードを選択した回数を表している。

表 1: 一対比較の結果

	A	B	C	合計
A	-	11	7	18
B	13	-	10	23
C	17	14	-	31

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 [10] を想定した。

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j} \tag{5}$$

$$\sum_i \pi_i = const.(= 100) \tag{6}$$

π_i : i の強さの量

P_{ij} : i が j に勝つ確率

これにより、一対比較に基づく評価を一義的に定めることができる。結果から最尤推定した π の値の比較を図9に示す。Cの入力時動作が最も高く評価され、次いでBの確定時動作、Aの通常動作の順に評価された。

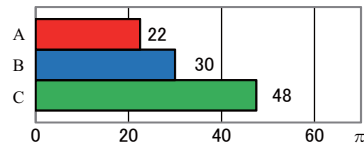


図 9: 各モードの強さ

7段階評価の平均値および標準偏差の結果と Friedman の分散分析および多重比較として Wilcoxon の符号順位検定を行った結果を図10に示す。Cの入力時動作は全ての項目で高く評価され、「一体感」、「思いが通う」の項目では、Aの通常動作に比べて有意水準5%で有意差が認められた。また、Bの確定時動作と比べた

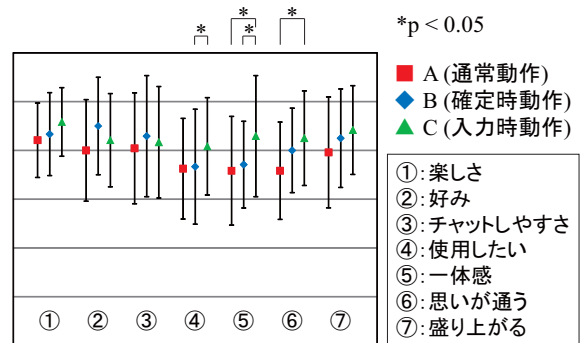


図 10: 7段階評価

場合も「使用したい」、「一体感」の項目において有意水準 5%で有意差が認められた。

また、自由記述式のアンケートで得られた意見を表 2 に示す。

表 2: システムに対する意見

- ・相手との一体感があって楽しかった
- ・動きがあると、感情が伝わりやすいと思った
- ・思っていた動作と違うと混乱してしまう
- ・一部のアクションが若干変に感じる
- ・C だと入力を間違えた時でも動いてしまう
- ・一文打ちきってから変換するので、C や A の方がよかった
- ・Enter を押してから動作した方 (A) が使いやすいと感じた
- ・C だと自分の動作が確認しづらかった

4.3 考察

実験の結果、一対比較において C の入力時動作は他のモードに比べて高く評価された。7 段階評価においても、「使用したい」、「一体感」、「思いが通う」の項目で高く評価された。これは、メッセージと情動表現の提示時間の差を小さくすることで、使用者とキャラクターとの一体感が向上し、より豊かな感情表現が可能となった結果だと考えられる。B の確定時動作は、C の入力時動作ほど高い評価は得られなかったものの、7 段階評価の「好み」、「チャットしやすさ」の項目において高く評価された。これは、文章を確定するまで動作を行わせないことで、打ち間違いなどにより意図しない動作が提示されることを抑えられたためだと考えられる。これらの結果により、情動表現を即時表示する本システムはチャットコミュニケーションの支援に有効であることがわかる。

また、自由記述アンケートから得られた意見において、C の入力時動作に対して「入力を間違えた時でも動いてしまう」「自分の動作が確認しづらかった」など、否定的な意見が得られた。しかしながら、一対比較および 7 段階評価において C の入力時動作が高く評価されたことから、情動表現の提示を行うキャラクターチャットシステムにおいては、メッセージの表示と情動表現の提示時間の差を小さくすることが好ましいと考えられる。一方で、会話のコンテキストにおける使用者の意図と動作の一致率とシステム評価の関係についてはさらに詳細に検討する必要がある。

5 結論

本研究では、情動表現を即時表示する新たな動作提示手法を取り入れた InterChat を用いてチャットコミュニケーション実験を行い、提示タイミングの評価を行っ

た。その結果、入力の度に文章の認識を行い動作を提示する手法が高く評価され、コミュニケーション支援におけるシステムの有効性を示した。

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, ペーパーサイエンス, Vol.2, pp.4-12 (2003)
- [2] 浅枝大志: ウェブ仮想社会「セカンドライフ」 - ネットビジネスの新大陸 -, アスキー新書 (2007)
- [3] LINE Corporation: <http://lp.play.line.me/>
- [4] 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ーチャットシステム 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).
- [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ー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] 岩佐 厚郎, 渡辺 富夫, 石井 裕: 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ーチャットシステムの開発, HAI シンポジウム 2015, pp.331-336 (2015)
- [7] Tomio Watanabe: Human-entrained Embodied Interaction and Communication Technology, Emotional Engineering, Springer, pp.161-177 (2011)
- [8] 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)
- [9] 中村 明: 感情表現辞典, 東京堂出版 (1993)
- [10] 広津 千尋: 実験データの解析 - 分散分析を超えて -, 共立出版 (1992)