

EinführungMors : 非随伴的・非自立的モダリティ の追加による遠隔音声会話拡張の試み

加藤千佳^{†1} 小倉加奈代^{†1} 西本一志^{†1}

今日、計算機とネットワークの発達により、多様なモダリティを使用したコミュニケーションメディアが実現されている。しかし、多くのモダリティを利用可能なメディアは、豊富な情報を伝えられる利点を持つ反面、伝えたくない情報まで伝わってしまう問題を有する。一方、少ないモダリティしか利用できないメディアに関しては、それとは逆の得失となる。このため、筆者らは両者の中間的特性を持つメディアが有用と考え、新たなコミュニケーションの試みとして随伴性と自立性のいずれも有しない副次的モダリティを組み込んだ遠隔音声会話拡張アプリケーション EinführungMors を試作した。基礎検討として行った実験から、EinführungMors の副次的モダリティを使ったコミュニケーションは、状況依存になると同時に、受信側の解釈に強く依存するものとなることが分かった。

EinführungMors : Augmenting Remote Voice Communication by Adding A Non-concomitant and Non-self-contained Modality

CHIKA KATO^{†1} KANAYO OGURA^{†1} KAZUSHI NISHIMOTO^{†1}

Various multi-modal communication media have been developed so far based on development of computers and networks. However, the media that allow users to use many modalities have a merit that the users can convey rich information while they have a demerit that even information that the users do not want to communicate is also transmitted. The media that allow users to use a few modalities have opposite features. Therefore, we assume that a remote communication medium that has intermediate feature of the conventional media is useful, and we developed a remote voice communication medium named EinführungMors, which is equipped with a non-concomitant and non-self-contained sub-modality, as an attempt of new communication. We carried out basic user studies and found that communications using EinführungMors strongly depends on the situation and on the receiver's interpretation.

1. はじめに

今日、我々の身の周りには相手と会話するためのコミュニケーションメディアが多数存在する。文明の発達によってテキストベースの手紙から音声による電話、そして音声と映像によるテレビ電話へとマルチモーダル化が進んだことにより、遠隔地間コミュニケーションは、次第にリアルな対面会話に近づきつつある。しかし、モダリティ数が増えると、“伝えたい情報”を多く伝えることができる反面、“伝える必要がない情報”まで伝わってしまう。例えばビデオ電話であれば、音声会話のように会話内容に関係しない作業や動作をすることは困難で、興味がない話であっても話を聞いている姿勢を保ち、相手の気分を害さないように演じなければいけない。一方、モダリティ数が少ない場合、“伝える必要がない情報”が伝わりにくい反面、“伝えたい情報”が伝わりづらことがある。例えば興味がない話を相手がしており会話に飽きた場合、対面対話であれば表情などでその旨を緩やかに伝えられるが、電話では言葉で伝えるしかないため、ともすると相手の気持ちを害してしまう結果となりがちである。このように、情報量の多さゆえの問題と少なさゆえの問題がそれぞれある。この問題を解決するには、2つの情報量の中間に位置づけられる遠隔音声会話におけるモダリティ拡張メディアが有用であると考える。

図1に、一般的なコミュニケーションメディアの構造を示す。コミュニケーションメディアには、メインモダリティと副次的モダリティから成り立つメディアと、メインモダリティのみから成り立つメディアがある。ここでメインモダリティとは、そのコミュニケーションメディアにおける主たる情報伝達手段として機能するモダリティである。また副次的モダリティとは、基本的にはメインモダリティと同時に伝達され、メインモダリティが伝達する情報に追加的な情報を与えるモダリティである。副次的モダリティは、さらに随伴性と自立性の2つの性質の有無によって分類できる。随伴性とは、副次的モダリティがどのように発生するかに基づく性質である。たとえば音声の場合、発話に伴い必ずパラ言語が発生するので、パラ言語は随伴性がある副次的モダリティである。自立性とは、そのモダリティ単独で（おおむね社会的にコンセンサスが得られた）意味の伝達が可能かどうかに基づく性

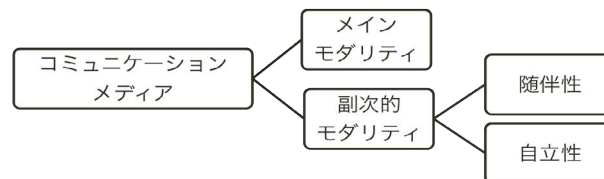


図1 コミュニケーションメディアの構造
Figure 1 Structure of Communication Media .

^{†1} 北陸先端科学技術大学院大学
Japan Institute of Science and Technology

表 1 各メディアのモダリティの構造と性質

Table 1 Structure and properties of modalities of each medium.

	メイン モダリティ	副次的 モダリティ		
			随伴性	自立性
音声電話	言語	パラ言語	○	△
メール	言語	—	—	—
TangibleChat	言語	打鍵振動	○	×
HandShakeSystem	言語	パラ言語	○	△
		握手	×	○
FeelLight	—	1 bit	—	×
EinführungMors	言語	パラ言語	○	△
		手指動作	×	×

質である。たとえば、対面対話中に同時に握手をする場合があるが、握手はそれ単独でも友好の意を伝えることが明確であるため、自立性があるモダリティであると見なせる。一方パラ言語や表情はおおむね社会的コンセンサスがあるものの、個人的差異や文化的差異があるなど、やや曖昧性が高いため、自立性は有るもののそのレベルは低いと考えられる。

表 1 に、いくつかの既存のコミュニケーションメディアに関して、上記の考え方にに基づき構造を分析した例を示す。

音声電話は、言語をメインモダリティとし、パラ言語を副次的モダリティとする構造を持つメディアである。パラ言語は随伴性を有し、やや低い自立性を有するモダリティである。メールは、言語をメインモダリティとし、副次的モダリティを基本的には伴わないメディアである。

Tangible Chat [2]は、言語をメインモダリティとし、テキスト入力時にキーボードを打鍵する行為によって生じる振動を副次的モダリティとするメディアである。打鍵振動を音声対話におけるパラ言語に相当するモダリティとして扱っているが、打鍵振動はそれ単独ではほとんど意味をなさないため、自立性は無いモダリティであると見なせる。

Handshake Telephone System [3]は、音声電話にロボットハンドを介した遠隔握手機能を追加したものであり、握手も副次的モダリティとして機能する。握手は、音声対話に伴って必ず生じるものではないので随伴性は無いが、それ単独で意味をなすので、自立性があるモダリティであると見なせる。

FeelLight [1] は、メインモダリティのみで構成される構造を持つメディアである。FeelLight は、送信側で LED が内蔵されたボタンを押下すると、受信側に設置された同じデバイスの LED の色が変わるという、ボタンの ON・OFF のみを伝える 1 bit 通信メディアである。このような情報には一般的意味が認められないため、これをメインモダリティとみなす

よりは、メインモダリティ

を持たず、自立性がない副次的モダリティのみで構成されるメディアであると見なす方が自然であろう。

以上のように、現在さまざまな構造を有するコミュニケーションメディアが存在するが、それらは原則としてメインモダリティを有し、副次的モダリティを伴う場合、その副次的モダリティは、随伴性と自立性の少なくともいずれか一方を持つものとなっている。自立性があるモダリティによって伝達される情報は当然意味をなす。また、随伴性があるモダリティは、それ単独では明確な意味をなさないものの、関連するメインモダリティと協調することによって具体的な意味を形成する。このように、従来の副次的モダリティは、随伴性か自立性のいずれかの性質を持つことによってなんらかの具体的な意味を伝達していた。このような副次的モダリティの特性が、最初に述べたようなモダリティの多少によって生じる問題の大きな要因となっているのではないかと、筆者らは考えた。

そこで本研究では、新たなコミュニケーションの試みとして、随伴性と自立性のいずれも有しない副次的モダリティを組み込んだコミュニケーションメディアを構築し、その影響を検証する。筆者らの知る限り、このようなコミュニケーションメディアに関する研究事例は、これまでのところ存在しない。随伴性も自立性も無い副次的モダリティによって伝えられる意味は非常に曖昧で、状況依存的になると同時に、受信側の解釈に強く依存するものとなるであろう。これによって、従来問題となった、副次的モダリティが少なすぎることによって「意味が伝わらない問題」と、副次的モダリティが多すぎることによって「意味が伝わりすぎる問題」の両方を解決できることを期待している。

以下本稿では、遠隔音声会話メディアに、随伴性と自立性を有しないジェスチャを副次的モダリティとして追加した双方向通信アプリケーションを提案する。試作したプロトタイプメディアを用いたユーザスタディを実施し、提案手法のコミュニケーションへの影響と有用性を検証する。

2. 関連研究

FeelLight は、きわめて単純な情報の交信によるミニマムなコミュニケーションの可能性を示しており、「相手の存在を感じる」といったように、情報の受け手がきわめて単純な信号を意味のあるメッセージとして解釈することを示している [1]。しかし FeelLight 単体では受け取る情報が少なすぎることから、相手の感情などの複雑な意味付けをすることは困難であるとも言える。Tangible Chat は、実験結果から感情の伝達による対話内容の活性化が見られている [2]。また Handshake Telephone System は、相手の存在感を感じさせる事への有効性を示している。しかし、通話中の握手は不自然な行為であり一部の被験者は違和感を感じている [3]。そのため通話においてより自然な行為を入力手段とするべきであると考えられる。

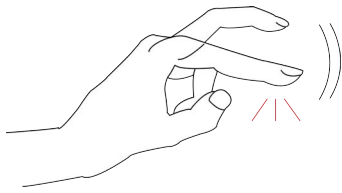


図 2 扱う手指動作
Figure 2 Finger action.

このように、それぞれのメディアにおいて様々に得失はあるものの、これらの先行研究は、非言語情報によってモダリティを僅かに拡張することの有用性を示唆している。

3. 提案手法

随伴性も自立性も持たない副次的モダリティ拡張の手段として手指動作に着目した。まず扱うジェスチャについて述べた上で提案手法である手指動作について述べる。

3.1 ジェスチャ

ジェスチャには、他者へある信号伝達のみを目的とした動作である“一次的ジェスチャ”と、受け手（動作を見ている人）によって偶発的な意味を付与される“偶発的ジェスチャ”とがある[4]。“偶発的ジェスチャ”は、感情・気分・意志をジェスチャの動作者が意図しない意味（感情・気分・意思など）を、受け手が作動的に推測するものである。例えば頬杖をつきながら相手の話を聞いている場合、動作者はただ頭を支えるための行為だとしても、見ている相手は話に飽きたのだろうといったネガティブな推測をするようなことである。

3.2 手指動作

ジェスチャによって相手に何かを伝える際は、意図的に伝達者が信号を送るだけでなく、受け手に信号を推測可能にすることが重要であるため、意図的に信号送信する“一次的ジェスチャ”のみでなく“偶発的ジェスチャ”も取り入れるべきであると考え。本研究では両方のジェスチャを取り入れるために、意図的な信号と、通話中に行われる自然な無意識的ジェスチャとしての信号との両方が受け手に伝達されるようにする。そこで、とりわけテンポやリズムなどによって表現に幅を持たすことが可能な手指動作に着目した。何かを反復的にさすったり弧を描いたりさまざまな手指動作が存在するが、第1段階として本稿では、手首を固定し指を上下に動かす手指動作を扱う（図2）。

4. EinfühlungMors : 手指動作入出力アプリケーション

手指動作を音声通話と併用する副次的モダリティとし、どのような意図があるかの意味づけや解釈を送り手と受け手に委ねたコミュニケーションメディアとして EinfühlungMors を試作した。なお Einfühlung とは、哲学用語で自己投入の意味である。

EinfühlungMors は、タッチパネル付き端末上でタッチパネルと叩く手指動作を入力とし、相手の端末で音を鳴らすこと



図 3 EinfühlungMors 使用イメージ図
Figure 3 Dialog image with using EinfühlungMors.

によって相手の手指動作を知らせるアプリケーションである（図3）。端末入力完了情報と相手のタッチ情報受信情報をテキストフィードバックによって提示される。

本アプリケーションはクライアントとサーバから構成される。クライアントは Android アプリケーションである。クライアントとサーバは TCP/IP 通信している。

5. 実験

「手指動作が伝達情報としてどのような意味で行われるのか」と「意味をもたない手指動作が行われるのか」について調査するために、提案アプリを用いた実験を行った。実験1の結果をもとに、信号送信と受信信号の意味の相違を調査するため実験2を行った。

5.1 実験1概要

著者らが所属する大学院の学生5名（男子4名、女子1名）を被験者とし、自由に提案アプリケーションを使ってもらいながら非対面で実験者と通話してもらった。通話の様子をビデオ録画し、実験後にビデオを被験者とともに見ながら聞き取り調査を行った。実験の手順を図4に示す。実験は、実験者1名と被験者1名の2名1組のペアで行い、提案アプリケーションを使用しながら、はじめに信号の送受信を確認し試用してもらった後、(1)雑談→(2)聞こえないフリ→(3)共感できないであろう話→(4)雑談から構成された9分30秒の音声会話をを行った。実験中の手指動作は、(4)の雑談パートの開始から2分経過後に実験者が一定のリズムで手指動作を行う以外は、被験者も実験者も適宜自由に手指動作を行った。被験者は、会話内容を事前に知らされていない。パート(1)と(3)の主な話者は実験者であり、最後の(4)雑談パートでは被験者に話をするように求めた。話題や話す内容は被験者によって異なった。聞き取り調査では「どのような意味

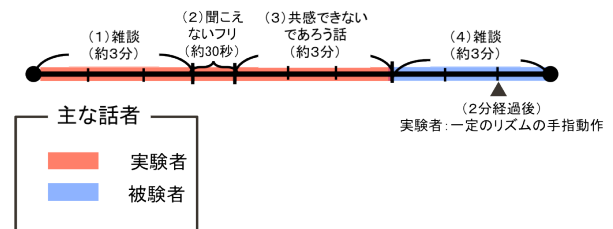


図 4 実験1の手順と主な話者

Figure 4 Procedure of Experiment 1 and the main speaker of each part.

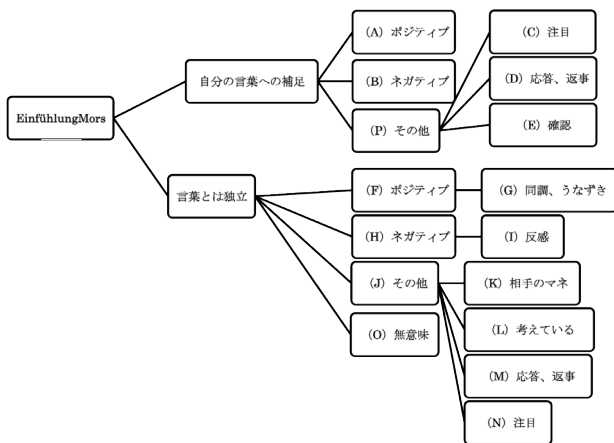


図 5 実験 2 アンケートの選択肢

Figure 5 Options of the questionnaire in experiment 2.

で手指動作を行ったか」という質問に対し、あらかじめ用意した意味カテゴリからの選択による回答をしてもらった。また、(2) 聞こえないふりとパート (4) の開始後 2 分経過から行った、実験者から送られた信号をどのような意味で捉えたかの質問にも同様に回答してもらった。

5.2 実験 2 概要

筆者らが所属する大学院の学生 6 名 (男子 5 名, 女子 1 名) を被験者とし、被験者 2 名 1 組のペアを 3 組構成した。ペアは互いに顔見知り同士で対面会話頻度が毎日、週 2~3 回、月 1 回とペアによってそれぞれ異なる。各ペアの被験者はそれぞれ別々の部屋に入り、自由に提案アプリケーションを使いながら会話をした。実験者は、カメラを介して別室からその会話の様子を観察・録画した。実験後、録画したビデオを被験者と共に見ながら聞き取り調査を行った。

今回の実験では、被験者ペアのうちの 1 名 (以下、被験者 A とする) に話の主導権を握ってもらい、(1) 雑談 (3 分)、(2) 相手が共感できないと考えられる話 (3 分)、(3) 雑談 (3 分) の音声会話の順で、それぞれ 1 分間ずつ空けて会話してもらった。もう 1 名の被験者 (以下、被験者 B とする) は、どのような会話を行うかを知らされていなかった。パート (1) と (2) の主な話者は被験者 A であり、パート (3) の雑談では被験者 B に話をするように求めるように指示した。実験中の手指動作は自由に行ってもらい、やり方について特に指示はしなかった。実験後の調査では、「どのような意味で手指動作信号を送信したか」、「受信信号をどのような意味で捉えたか」という質問に対し、予備実験 1 の結果をもとに用意した選択肢 (図 5) から回答してもらった。また被験者ペア同士の関係を知るために、ペアと出会ってからの期間、対面会話頻度、通話頻度、相手についての理解度のアンケートを行った。

6. 結果

6.1 実験 1 の結果

まず、手指動作信号を送信する場合の意味づけについて検

討する。信号送信の回数は 27~218 回と、被験者によって差が見られた。送信信号の意味づけには 3~7 種類と差があったが、平均的には 5.1 種類あったことからさまざまな意味をもたせて信号を送信することが分かった。被験者によって送信する信号の意味にはばらつきがあったが、パート (2) の「聞こえないふり」の際は、全ての被験者が「応答・確認」の意味で信号の送信を行った。また、全ての被験者が意味を持たない手指動作による信号送信を行った。また、相槌を発しながら送信するような言葉の内容に則した補足意味で送信する場合と、言葉を発さずに信号送信のみで相槌を打つような言葉とは独立して用いる場合があることがわかった。興味が無いので話を変えてほしいといったようなネガティブなことを感じた際には、3 名の被験者が言葉には出さずに、信号にその意味を込めて実験者に伝えようとしていた。

次に手指動作信号を受信した場合の意味づけについて検討する。パート (2) の「聞こえないふり」では、全ての被験者が「応答・確認」の意味で信号を受け取っていた。パート (4) の雑談時の、2 分経過後に実験者が送った一定のリズムでの送信信号に対しては、1 名の被験者のみが「反感」の意で捉えたが、他の 3 名は「無意味」と捉え、残り 1 名は信号に気付かなかった。

6.2 実験 2 の結果

表 2 に、被験者が送信信号および受信信号に対してどのような意味づけを行ったかに関して調査した結果を示す。表中のアルファベットは、図 5 に示す選択肢に対応している。信号の送受信は、ペアによって合計 49~377 回と幅があった。送信信号への意味付けは 2~9 種類、平均で 5.5 種類であった一方受信信号への意味付けは 2~6 種類、平均で 3.6 種類であった。また、送信信号の種類が多い被験者ほど受信信号に多くの意味付けをする傾向にあった (表 2)。送受信信号の意味の一致率は最も低いペアで 0%、最も高いペアで 36.3%、平均 24.1%であった (表 3)。送信した信号への意味づけとして

表 2 被験者による送受信信号の意味分類

Table 2 Classification of meanings of the transmitted/received signals by the subjects.

ペア	被験者	送信信号		受信信号	
		言葉への補足	言葉とは独立	言葉への補足	言葉とは独立
1	a	E	O	E	O, M, K
	b	D, E, P	M, O, G, K, L	E	O, L
2	c	A, C	H	—	I, M, H
	d	B	J, B, O	—	I, J
3	e	A, B, C, P, E	M, G, O, F	P	L, G, O, J
	f	C, E	M, O, H	E	M, O, H

表 3 送受信信号の意味一致率と一致した意味の種類

Table 3 Matching rate of meanings between transmitted and received signals and categories of the concordance meanings.

ペア	意味一致率	一致した意味の種類	
		独立	相手のマネ, 無意味, 応答・返事
1	36.3%	独立	相手のマネ, 無意味, 応答・返事
		補足	確認
2	0%	—	
3	36.1%	独立	無意味, 応答・返事
		補足	確認

最も発生頻度が高かったのは手遊びや手持ち無沙汰を意味する“無意味”であった。一方、受信した信号への意味づけとして最も発生頻度が高かったのは“無意味”と、言葉を補うための“確認”の意味であった。表4に、言葉への補足としての送信信号と言葉とは独立して送信された信号に関し、送信側と受信側でどの程度意味が一致していたかの割合を示す。言葉への補足の意味で送受信が一致したのは“確認”で、ペア1が72%、ペア3が84%であった。言葉と独立した意味で送受信が一致したのはペア1、ペア3ともに“無意味”、“応答、返事”で、ペア1ではさらに“相手のマネ”であった。受信信号と送信信号の意味一致率は、ペア1では“相手のマネ”が、ペア3では“確認”が最も高かった。全体で両ペアとも“確認”の一致率が最も高かった。

被験者ペア同士の関係に関するアンケート結果を表5に示す。この結果から、特にペア2間では対面会話頻度、通話頻度、相手に関する理解度が低く、出会ってからの期間も浅いことが分かった。

また、口頭調査から送信者が信号送信した時、受信者は前後の会話の状況から相手の心理や状況を考え、受信信号に意味付けして理解しているという意見があった。また、ある被験者は言葉とは独立した意味のみで送信受信を行っていた。また、全体的に実験終了に近づくにつれ信号送信頻度が減る傾向にあった。

7. 考察

実験結果から、ユーザは簡単な意味であれば受信信号の意味を理解可能なことが考えられる。言葉に伴って信号を受信した際、その言葉をもとに相手の心理や状況を信号に意味付けることが考えられる。また、言葉とは独立した信号を受信した際は、前後の状況、タイミング、会話内容から意味付けを行うことがわかった。受信者が想定しない時に受け取った信号をネガティブな意味に捉えた被験者もいたことから、特にタイミングが大きく関係していると考えられる。以上から、提案アプリケーションは状況依存的になると同時に、受信側の解釈に強く依存するものであることが言える。しかし、今

表 4 言葉と独立した送受信信号と言葉への補足とした送受信信号の意味一致率

Table 4 Concordance rate of meanings of transmitted/received signals that were transmitted independent from verbal messages and supplement to the words.

ペア	言葉への補足	言葉とは独立		
		確認	無意味	応答、返事 相手のマネ
1	72%	21%	66%	100%
2	0%	0%	0%	0%
3	84%	53%	66%	0%

表 5 ペア同士の関係に関するアンケート回答一覧

Table 5 Results of the questionnaire on the relationship between the subjects in each pair.

ペア	被験者	相手についての理解度 (1~10)	出会ってから の期間	対面会話頻度	通話頻度
1	a	7	7~9ヶ月	毎日	毎日
	b	6			
2	c	2	4~6ヶ月	2~3回/週	1回/月
	d	3			
3	e	3	1年4ヶ月 ~1年6ヶ月	1回/月	1回/月 以下
	f	7			

回の実験ではほとんどの送信信号の意味と受信信号の意味の一致が見られなかった。これは被験者ペアの普段のコミュニケーション頻度や2人の出会ってからの期間、相手に関する理解度が関係しうると考えられる。さらに、今回は1回だけの使用だったが、提案アプリケーションを何度も使用することにより、ペア間で信号の意味をパターン化して意味を創発していく可能性も考えられる。そのため、今後は長期的な実験観察が必要となると考えられる。さらに、副次的モダリティが少なすぎることによって「意味が伝わらない問題」と、副次的モダリティが多すぎることによって「意味が伝わりすぎる問題」の両方を解決できるかどうかに関し、遠隔音声のみの会話と提案アプリケーションを用いた遠隔音声会話の比較実験を行う必要がある。

実験終了に近づくにつれ信号送信頻度が減る傾向は、特にタスクを与えた被験者A側に目立って見られた。これはタス

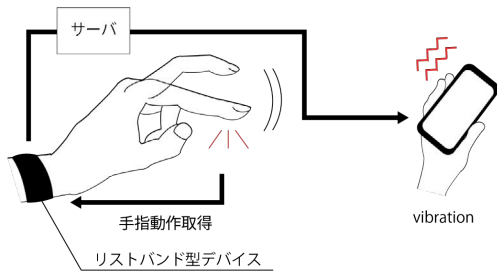


図 6 今後のシステムイメージ図

Figure 6 Future system diagram image.

クに集中し提案アプリケーションの存在を忘れるため、すなわち随伴性の欠如によるものだと考えられる。また、聴覚出力が通話への集中力を妨害するという意見や、信号受信出力が視覚と聴覚によるものだったため直感的でなく意味を受け取りにくいという意見があった。さらに、ビデオ観察からデバイス上以外での手指動作が見られた。それらを踏まえ今後は場所の制約をなくすため加速度センサを搭載したリストバンド型デバイスを開発する。そして、リストバンド型デバイスにより手指動作を取得し携帯電話への振動によって出力するためのシステムの開発を目指す (図 6)。

8. おわりに

本稿では、マルチモーダル化が進んだコミュニケーションメディアによる情報量の多さゆえの問題と少なさゆえの問題を解消させる一助として、遠隔音声会話メディアに随伴性と自立性のいずれも有しない手指動作を副次的モダリティとして追加した遠隔音声会話拡張アプリケーション

EinführungMors を提案した。

基礎的検討として、提案アプリケーションを用いた実験を 2 つ行った。結果として、手指動作にさまざまな意味をもたせて信号送信することが分かり、また信号受信の際はなんらかの意味付けを行って受け取ることが分かった。さらに 2 つ目の実験から、ユーザは簡単な意味であれば送信信号意味を理解可能であることがわかった。言葉を伴う信号を受信した際には、その言葉をもとに相手の心理や状況を信号に意味付けし、言葉とは独立した信号を受信した際は前後の状況、タイミング、会話内容から意味付けを行うことがわかった。

今後は、遠隔音声のみの会話と提案アプリケーションを用いた遠隔音声会話の比較実験を行って、随伴性と自立性の両方を有しない副次的モダリティの効果をさらに詳細に検討したい。また、聴覚出力による会話の妨害などの問題やデバイス上以外での手指動作が見受けられたので、加速度センサを搭載したリストバンド型デバイスを開発することも課題とし今後研究を進める。

参考文献

- 1) Kenji Suzuki, Shuji Hashimoto: FeelLight: A Communication Device for Distant Nonverbal Exchange, ETP' 04 Proceedings of the 2004 ACM SIGMM workshop on Effective telepresence, pp.40 - 44 (2004).
- 2) Kazushige OUCHI and Shuji HASHIMOTO: Handshake Telephone System to Communicate with Voice and Force, IEEE International Workshop on Robot and Human Communication, pp.466 - 471(1997).
- 3) 山田裕子, 平野貴幸, 西本一志: TangibleChat :打鍵振動の伝達によるキーボードチャットにおける対話状況アウェアネス伝達の試み, 情報処理学会論文誌, vol. 44, No5, pp.1392 - 1403(2003)
- 4) Desmond Moris, 藤田統訳 (2007). マンウォッチング, 株式会社小学館.