

共有仮想空間での コミュニケーション支援

東京農工大学大学院 藤田欣也, 宮島俊光



遠隔コミュニケーション(2/2)

★ プライベートユーザの立場に立つと

- 余計な個人情報は伝送したくない : 個人情報保護
- 電話の時刻が見てないと手持ちぶさた : 注意の対象
(今、私の姿が見えなくなって声だけが聞こえたらどうしようか...)

会話相手の実写画像ではなくても、会話の時に
注意を向ける適切な対象が欲しい



アバターを利用し、共有仮想空間で
音声コミュニケーション

個人情報流出が避けられる
視線や表情などのノンバーバル情報が欠落



アバターノンバーバル情報の制御の必要性

遠隔コミュニケーション(1/2)

- ◆ 広帯域・常時接続ネットワークの整備
メッセンジャなど実時間コミュニケーションの普及

◆ 実写型映像・音声コミュニケーション (TV電話/ビデオ会議)

表情や身振りなどのノンバーバル情報を伝送可能
伝送したくない個人情報伝送の可能性



- オフィシャルな場所や用途なら可能
(勤務先での遠隔会議など)

- プライベートな利用は疑問

【学生の多くはメッセンジャ(テキスト)を使用しているが、TV電話機能利用者はほとんど皆無】



実写型メッセンジャの例
(Microsoft Windows Live Messenger Webサイトより)

アバターIV情報制御の例

直接法 カメラで動作を計測し制御

間接法 音声でユーザ動作を制御



視線 (GAZE Queens
大学WEBサイトより)



視線



表情
(Face Communicator
沖電気WEBサイトより)



表情



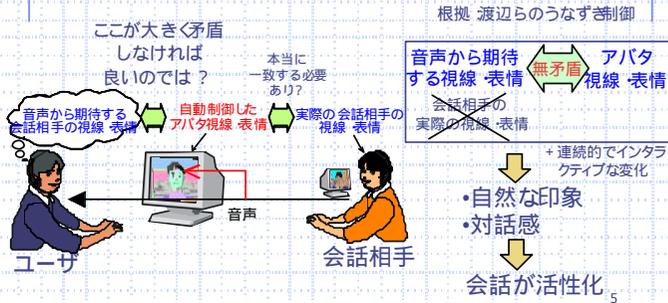
うなずき (USC:
IVA06デモ)



うなずき
(InterActor 岡山
県立大学WEBサ
イトより)

仮説 :間接法NV情報制御の有効性

- アバタの視線や表情は、必ずしも会話相手の実際の視線や表情と一致しなくても、ユーザが会話相手の音声から期待する視線・表情とアバタの視線・表情の間に大きな矛盾がなければ、連続的な視線・表情の変化は、より自然な印象を与え、会話を活性化すると期待される。



本日の内容

1. **ノンバーバル情報**とコミュニケーション
2. アバタ**視線**の自動制御
3. アバタ**表情**の自動制御
4. 共有仮想空間での、より豊かなコミュニケーションにむけて

7

研究の目標 と 想定前提条件

◆ 目標

- NV情報の自動制御による**共有仮想空間でのコミュニケーション支援**



◆ 前提条件

1. 既知の相手との会話 (マルチユーザ)
2. カジュアルな会話 (会話の活性化)
3. 自動制御を基本とし手動制御を排除しない

6

会話におけるNV情報の機能

(ノンバーバルインタフェース(黒川 彦改変))

● メタコミュニケーション機能

- 挨拶, 会話の開始...

● 心的情報の伝達

- 感情, ストレス

● バックグラウンド情報の伝達

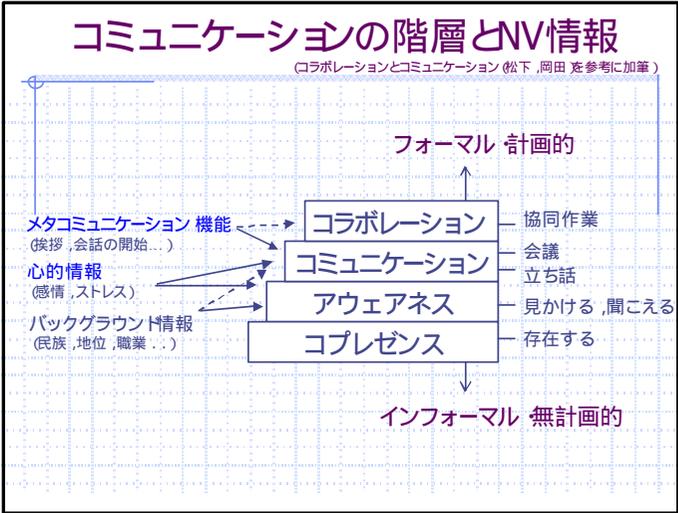
- 民族, 地位, 職業...



- 「ことば」より重要と考える研究者も多い

	ノンバーバル	バーバル
Birdwhistell(1970)	65 ~ 70%	30 ~ 35%
Katz(1983)	65 ~ 90%	10 ~ 35%
Mehrabian(1968)	93%	7%

8



身体動作の種類とコミュニケーション

調整子	発話権の授受や会話の流れを円滑にする(うなずき, 視線)	話者交替などを円滑にする
感情表出	情動に伴う表情, 身振りなど	好感情の表出はコミュニケーションに好影響
適応子	状況に適応するための動作(髪の手いじり, 貧乏揺すり)	自然な印象に好影響
標識	音声言語に翻訳可能で表象, サインとも呼ばれる	
例示子	発話内容を強調, 補足する	

斜線部分: 言語処理(意味解釈)または手動制御が必要

想定される自動制御の対象

11

ノンバーバル情報の種類

(ノンバーバルインタフェース(黒川 改変))

分類	具体例	
身体動作 (特に顔周辺)	視線, 表情, 唇の動き, 瞳孔, 身振り, 姿勢	
周辺言語	声質(アクセント, 発音...), 発声法	実現
対人距離	距離, 位置関係	実現
身体的特徴	体格, 体型...	実現
身体装飾	衣服, 装飾品, 化粧, 身体加工	実現
対人接触	抱擁, 握手, タッチ, (触感共有)	必要性低
生理現象	あくび, 発汗	必要性低

緑文字: 仮想空間音声チャットシステムで既に実現、または容易に実現可能

10

- ### 本日の内容
1. ノンバーバル情報とコミュニケーション
 2. アバタ視線の自動制御
 3. アバタ表情の自動制御
 4. 共有仮想空間での、より豊かなコミュニケーションにむけて
- 12

視線の機能

- ◆ 会話の開始合図 調整子
- ◆ **発話権の移譲と要求** 調整子
- ◆ 意思表示
- ◆ 感情表現 感情表出
- ◆ 対人関係の性質の伝達
- ◆ 関心表明

- その他のキーワード:
注視 (凝視), 視線交差, 儀礼的無関心

13

遠隔共有空間での視線制御

実写型



仮想空間型

センサやカメラを用いてユーザ視線を検出する直接法
 × 高精度の視線方向検出が必要



GAZE

発話情報に基づく疑似的な
アバタ視線制御 (間接法による話者注視)

15

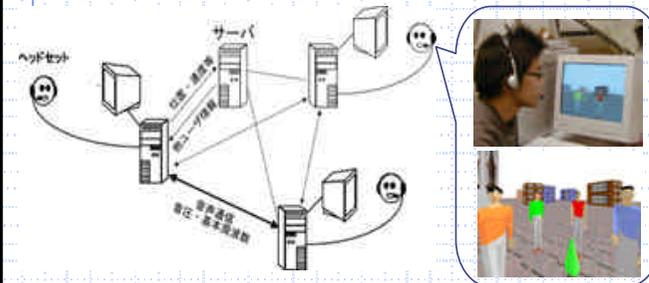
注視に関する知見

- 会話中に視線をそらせたり交差させたりする行為は、**会話の開始合図**や**発話権の授受**など**会話の調整機能を有している** : S.D.Duncan, Jr., and D.W.Fiske, "Face-to-face interaction" Research, Methods, and Theory, Hillsdale, N.J:Lawrence Erlbaum(1977).
- 視線には**表現機能**もあり 特に凝視は大きな作用を持つ : A.Kendon, "Some functions of gaze-direction in social interaction", Acta Psychologica, 26, 22-63(1967)
- 注視確率の増大が**発言を促す** : R. Vertegaal, Y. Ding, "Explaining Effects of Eye Gaze on Mediated Group Conversations: Amount or Synchronization?" Proceedings of the 2002 ACM conference on CSCW, pp. 41-48, 2002
- 長時間の持続的な注視は**支配感により悪い印象を与える可能性がある** : 深山, 大野, 武川, 濵木, 秋田, "擬人化エージェントの印象操作のための視線制御方法", 情報処理, 43(12), 3596-3606(2002)

14

マルチユーザ音声チャットシステム

- ◆ 各ユーザ情報 :サーバ管理
- ◆ 音声 :音声情報 :VoIP機能 (P2P通信)
- ◆ 音声情報を用いてアバタ視線を制御

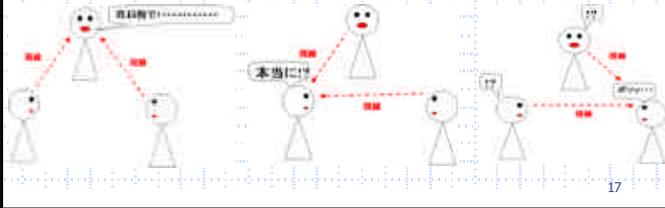


16

発話情報を利用した視線制御

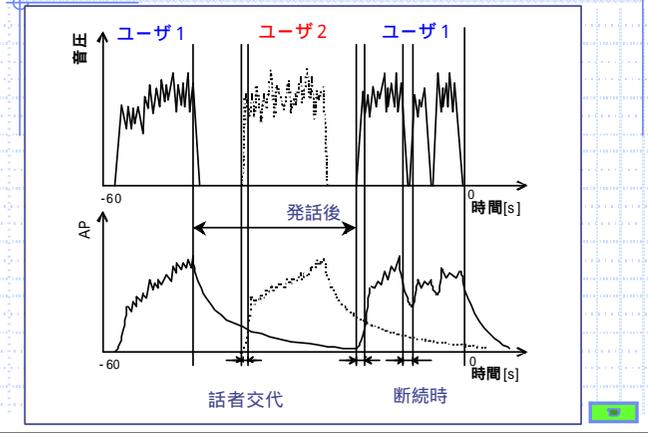
仮想空間で誰が注目を集めるか？

注目度(Appeal Point)
 長時間話している } 音圧重み付き発話量
 大きな声で話している }
 話を始めた瞬間 } 発言開始効果



17

アピール度の計算例 (話者注視効果APcのみの場合の概念図)



発話情報を利用した視線制御

注目度(AP:Appeal Point)にもとづいて注視対象を制御

AP_c 話者注視効果 $AP_c = \int_{-60}^0 \log(v(t)) \exp\left(\frac{t}{60}\right) dt$
 長時間話している
 大きな声で話している
 過去60秒間の発話量 (音量と経過時間の重み付発話時間)

AP_s 発言開始効果 $AP_s = \frac{5-(t-t_s)}{5}$
 話を始めた瞬間
 (5秒間沈黙すると発生、それ以外の時は0)



$$AP = a \times AP_c + b \times AP_s \quad (a,b \text{は実験的に} 1:2)$$

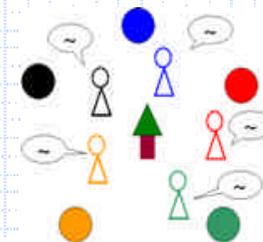
18

視線制御評価実験

会話に対する視線制御の効果を評価

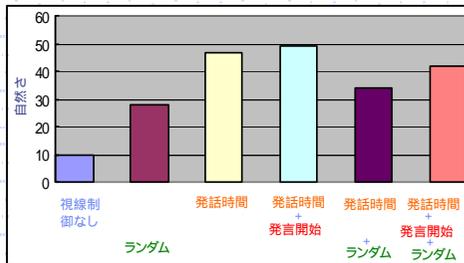
実験内容

仮想空間内で5名で会話
 日常的な話題 (被験者が話しにくいと回答したものは除外)



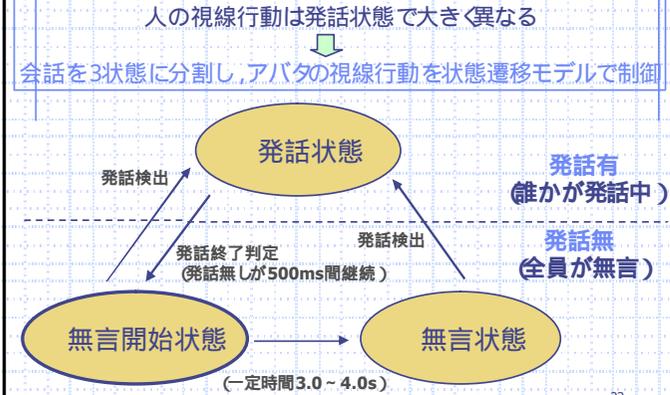
20

主観評価結果 (順序尺度)



- 視線制御が自然な印象を与えた
アバタの挙動に自律性 (他者発言時) と **インタラクティブ性** (自己発言時) を付与
- ランダムな注視対象者の変更は、却って逆効果²¹

改良型注視制御モデル概要

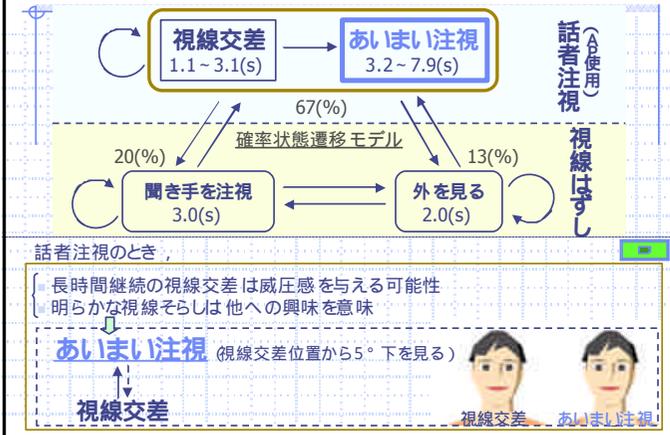


得られた知見と問題点

- ※ 知見1 発話に基づく視線制御はアバタの挙動に自律性 (他者発言時) とインタラクティブ性を与え、自然な印象につながる
- ※ 知見2 発話量や発話開始の前後関係の反映は、自然さにつながったが、ランダムな注視対象者の変更は逆効果
- ※ 問題点1 ユーザ発話中であってもランダムに他者を注視
- ※ 問題点2 話者を注視するのみで、沈黙継続時の制御なし
- ※ 問題点3 視線が持つ **発話委譲**、**促進**、**機能** の利用なし
 アバタに応用すればユーザの発話を促進できるのでは?

↓
 人間の視線行動モデルにもとづく、
 発話情報を用いたアバタ注視自動制御

発話状態制御



無言開始状態制御

実空間

人の発話終了時の会話権委譲行為, [模本, 仮]

- 発話者 **次話者を見る**
- 非話者 **話者から次話者に視線を移す (発話者の後)**

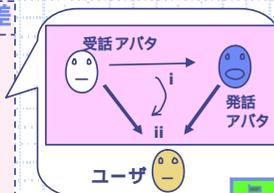
発話終了時に, アバタがユーザーを見ることでユーザーの発話促進

仮想空間

ユーザーが発話者のとき: **ユーザーと視線交差**

ユーザーが受話者のとき

- 発話アバタ: **ユーザーと視線交差**
- 受話アバタ:
 - i) 発話アバタ
 - ii) ユーザーを見る



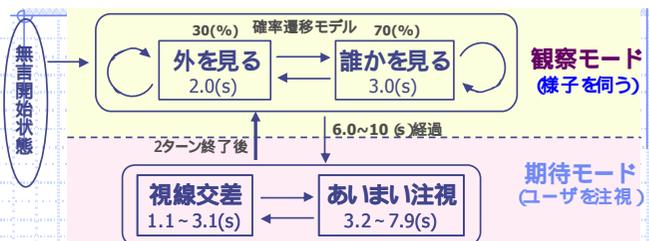
実験 ~ :各状態制御の評価

- 実験条件: 各制御の有無の組合わせ
- 評価方法:
 - リカード法による5段階の主観評価
 - 内観報告
- 被験者: 15名 (3人 × 5グループ, 男13人・女2人)
- 実験方法:
 - ヘッドセット着用で5分間の対話 (会話内容は事前に聞き取り調査し、興味のある話題を指示)
 - 条件の試行順序は順序効果を考慮



27

無言状態制御



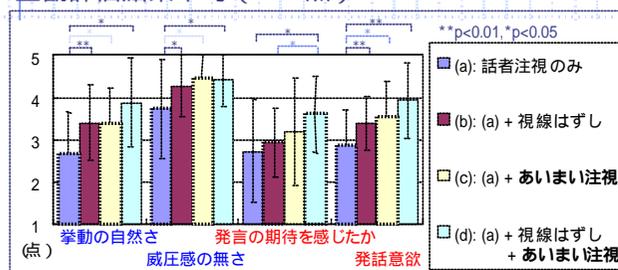
実空間では、人は無言のとき視線を下げ、他者の様子伺う
会話活性化のためには、各ユーザーの発話を促進することが望ましい

観察モード: **ランダム**に人や外を見る

期待モード: **ユーザーを注視** 発話促進

実験結果 :発話状態制御

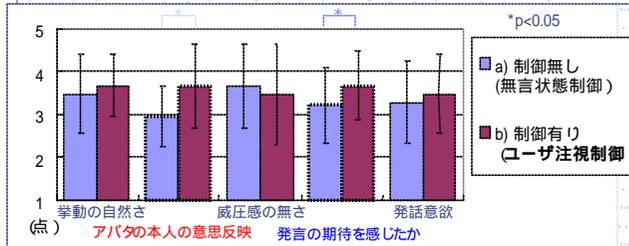
主観評価結果平均 (1~5点)



- あいまい注視** 自然さ, 威圧感の軽減, **発話促進** に有効
(あいまい注視により, 視線交差が強調された可能性)

実験結果 無言開始状態制御

主観評価結果平均 (1~5点)



ユーザ注視制御:

- ユーザの発言促進に有効
- 実際はパートナーの意思を反映していないにも関わらず、アバタの本人の意思反映が向上

アバタ視線の自動制御のまとめ

- あいまい注視や適度な視線はずしは、自然さや威圧感の軽減に加えて視線交差を強調する効果がある
- 適度なユーザ注視は発話を促進する効果がある

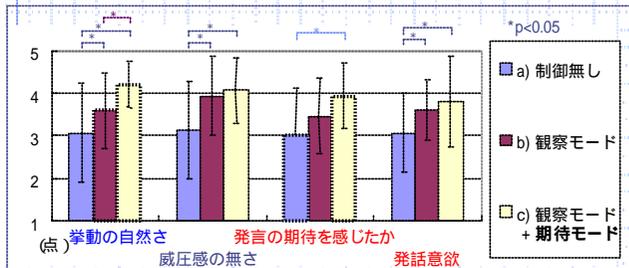
< 視線に関する今後の検討課題 >

- 意思表示, 感情表現, 対人関係の性質の伝達, 関心表明 (Joint Attention, Gaze Awareness...) など, 他の視線機能の実現方法の検討

31

実験結果 無言状態制御

主観評価結果平均 (1~5点)



期待モード (適度にユーザを注視)

自然さ, ユーザの発言促進に有効

(観察モード > 期待モード > 観察モード)

本日の内容

- ノンバーバル情報とコミュニケーション
- アバタ視線の自動制御
- アバタ表情の自動制御
- 共有仮想空間での、より豊かなコミュニケーションにむけて



このあたりでちょっと一息

32

アバタ表情制御の先行研究 と本研究



本研究

音声情報のみから、アバタの表情を自動制御したい

自動制御するアバタ表情の種類

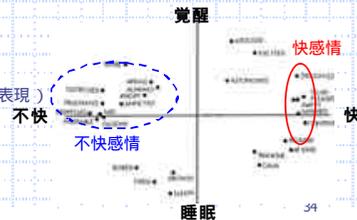
誤動作による意図しない不快感情の表出はコミュニケーション阻害の懸念、チャットでは、怒り顔や泣き顔は意志表出手段としての断続的使用が多い

自動制御対象を快感情(笑顔)のみに限定
(不快感情の表出制御は手動のみと想定)

Russellの円環モデル

(感情を覚醒-睡眠, 快-不快の2軸で表現)における

覚醒-睡眠軸のみの1軸の制御に単純化



笑顔の心理的要因とモデル化

◆本能的で不随意的笑い(快感情)感情発露

- ムード(Mood)
 - ◆比較的長時間持続する穏やかな感情
- 情動(Emotion)
 - ◆認知や生理が関わる複雑な過程

連続成分

瞬時成分

疑似意志表出成分

◆社会的な随意的笑い意志表出

- ◆(苦笑いなどの追従や肯定など)

笑顔強度
Smile Level

(志水他, 人はなぜ笑うのか 笑いの精神生理学)

連続成分 (SLc) の算出方法

快感情(喜び)と怒りで基本周波数が高くなる

(備考: 感情の判別分析から見た感情音声の特性)

(例: 会話が盛り上がるにつれて徐々に笑顔に変化)

発話者の基本周波数を利用
(自己相関を用いて推定)

基本周波数は抑揚の影響を受ける

移動平均処理

$$SLc = \frac{1}{T} \int_{-T}^0 F_{norm} dt$$

$$F_{norm} = \frac{F_t - F_{L0}}{F_{H0} - F_{L0}}$$

F_{norm} : 正規化基本周波数

T : 移動平均時間 (実験的に過去5秒間に設定)

F_{L0} : 平均基本周波数の最低値

F_{H0} : 平均基本周波数の最高値

瞬時成分 (SLr) の算出方法

会話相手の発言は時として瞬時的な情動変化を誘発
(例: 冗談による「ばか笑い」)

話者交代直後の応答初期音圧 (反応の強さ) を利用
(話者交代検出: 相手ユーザが発話して3秒以内の発話)

$$SLr = \frac{P_i - P_{s0}}{P_{L0} - P_{s0}}$$

P_i : 発話初期音圧のピーク値
 P_{s0}, P_{L0} : 平均音圧の最小値、最大値

反応は瞬時的
指数的に減衰処理 (1.5秒)

37

総笑顔強度 (SL) の算出

笑顔強度 (SL)

$$SL = a \times SLc + \text{Max}(b \times SLr, c \times SLpi)$$

会話相手の発話

自己発話

(実験的に $a=0.75, b=0.5, c=0.25$ に決定)

- 本能的で不随意の笑い: **感情発露**
 - ムード(Mood)
 - 比較的時間持続する穏やかな感情
 - 情動(Emotion)
 - 認知や生理が関わる複雑な過程
- 社会的な随意の笑い: **意志表出**
(苦笑いなどの追従や肯定など)



39

疑似意思表出成分 (SLpi) の算出方法

社会的な随意 (意志表出) の笑顔は
会話相手の発話や行動に対して発生
発話を伴わない場合も多い (例: 愛想笑いなどの追従、肯定)



自己発話に応じた笑顔表出で疑似的に意思表出を実現

(顔辺りのわずき制御と類似の考え方)

自己の発話終了直前の音圧を利用

$$SLpi = \frac{P_i - P_{s0}}{P_{L0} - P_{s0}}$$

P_i : 発話終了前音圧のピーク値
(発話終了前300 - 700ms)
 P_{s0}, P_{L0} : 最小, 最大平均音圧

38

システム構成

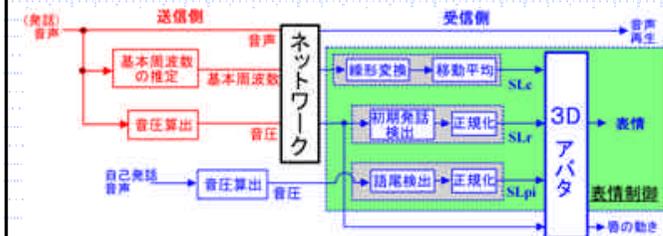
マルチユーザ・アバタ音声チャットシステム

基本周波数と音圧から笑顔強度SLを算出

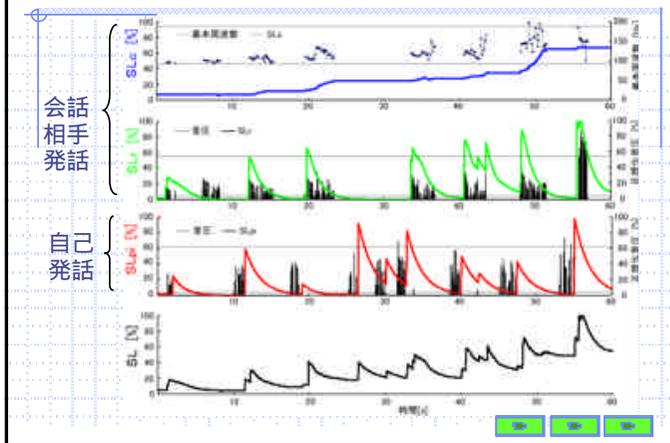
SLからFACSに従い**アバタ笑顔を制御**



笑顔 (左から50%, 50%, 100%)

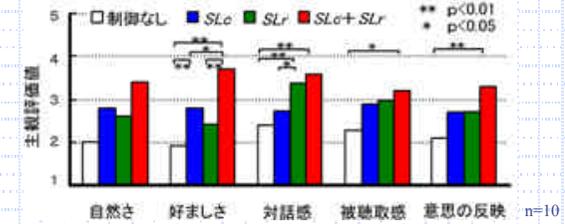


会話時の笑顔強度 (SL) の例



SLc と SLr の検証結果

主観評価結果



- **SLc + SLr の条件で最も良好**
- 自然さ, 好ましさ : **SLc > SLr**
ゆっくりした表情変化 (連続成分) は自然さに有効
- 対話感, 被聴取感 : **SLc < SLr**
反応の強さの反映 (瞬時成分) は対話感増強に有効

連続成分(SLc)と瞬時成分(SLr)の検証実験

実験条件

表情制御なし **SLc のみ**
SLr のみ **SLc+SLr**

評価項目: Likkert法による主観評価
(対話感, 被聴取感, 自然さ, 好ましさ, 意思の反映)

被験者: 10名 (2人 × 5組)

実験方法:

- 選定した日常的な話題の中から、被験者が共通して会話可能と回答した話題を選定し、4分間会話
- 4条件を連続して試行した後、聞き取り調査

42

疑似意思表出制御法 (SLpi) の検証

実験条件

SLc+SLr **SLc+SLr+SLpi**

評価項目: Likkert法による主観評価
(対話感, 被聴取感, 自然さ, 好ましさ, 意思の反映)

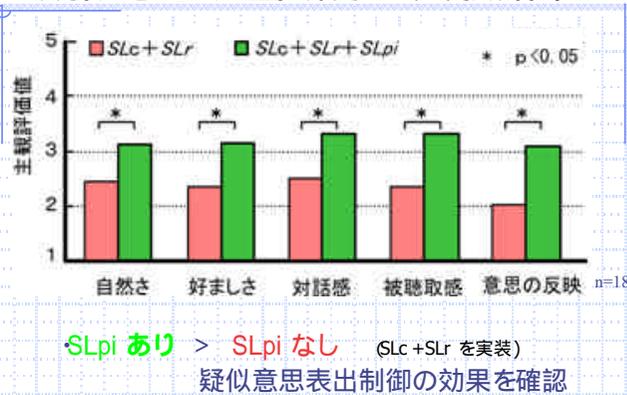
被験者: 18名 (3人 × 6組)

実験方法:

- 共通する会話可能な話題で4分間会話
- 2条件を連続して試行した後、聞き取り調査

44

疑似意思表出成分の実験結果



45

アバタ表情制御のまとめ

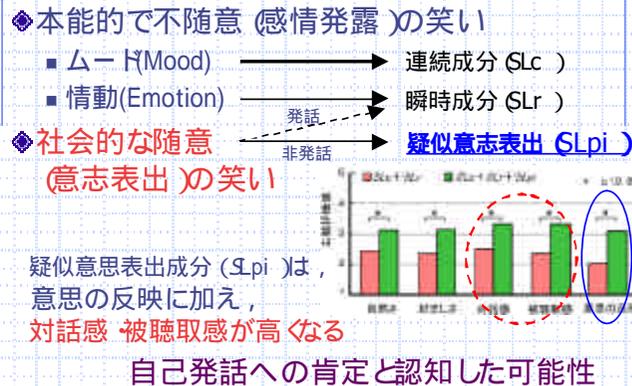
- 音声情報に基づく間接的なアバタの表情制御は、自然さなどの印象につながる
- ゆるやかな表情の変化は自然さや好ましさに、動的な表情の変化は対話感や被聴取感につながる
- 疑似的であっても、社会的意志表出の笑顔は意思の反映に加え対話感や被聴取感を高める

< 表情に関する今後の検討課題 >

- より詳細な笑顔発生要因と笑顔のモデル化 (笑顔 / 上機嫌 / 愛想笑い など)

47

考察



46

本日の内容

1. ノンバーバル情報とコミュニケーション
2. アバタ視線の自動制御
3. アバタ表情の自動制御
4. 共有仮想空間での、より豊かなコミュニケーションにむけて

48

1.身体動作以外のNV情報の制御

分類	具体例
身体動作	表情, 視線, 身振り, 唇の動き, 姿勢
周辺言語	声質(アクセント, 発音, ピッチ), 発声法
対人距離	距離, 位置関係
対人接触	抱擁, 握手, タッチ, (触感共有)
生理現象	発汗
身体的特徴	
身体装飾	

音声チャットを前提にすれば、特に工夫しなくても伝わる
(マイク感度の自動調節機能等があるので現実空間とは違って来る可能性有り)

1.身体動作以外のNV情報の制御

分類	具体例
身体動作	表情, 視線, 身振り, 唇の動き, 姿勢
周辺言語	声質(アクセント, 発音, ピッチ), 発声法
対人距離	距離, 位置関係
対人接触	抱擁, 握手, タッチ, (触感共有)
生理現象	発汗
身体的特徴	
身体装飾	

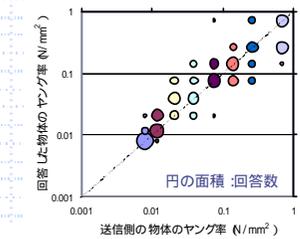
親密な間柄を除きあまり生じない
(触感の共有は話題提供の可能性)

1.身体動作以外のNV情報の制御

分類	具体例
身体動作	表情, 視線, 身振り, 唇の動き, 姿勢
周辺言語	声質(アクセント, 発音, ピッチ), 発声法
対人距離	距離, 位置関係
対人接触	抱擁, 握手, タッチ, (触感共有)
生理現象	
身体的特徴	
身体装飾	

距離や位置は関係や立場を反映
(密接距離, 個体距離, 社会距離, 公衆距離)
移動は会話の相手を制御する手段
(手動制御を基本とし, 自動制御の可能性も?)

物体の柔らかさの遠隔伝送



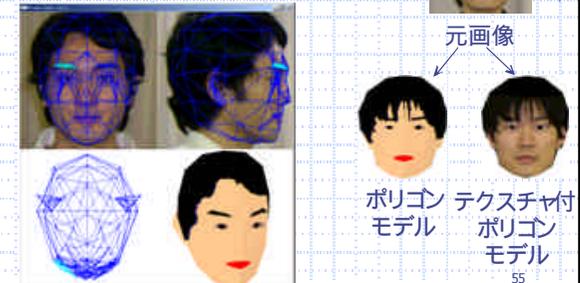
1. 身体動作以外のNV情報の制御

分類	<ul style="list-style-type: none"> 意図せずに発露 重要な手がかり 随意制御できない 伝送したくない (状態推定に基づく擬似的な間接制御?)
身体動作	
周辺言語	
対人距離	
対人接触	抱擁, 握手, タッチ, (触感共有)
生理現象	あくび, 発汗
身体的特徴	体格, 体型, 容貌, 皮膚, 頭髪
身体装飾	衣服, 装飾品, 化粧, 身体加工

2. アバタへのユーザ個性の反映

既知の友人との会話では, 会話の相手が「その人」である感覚が欲しい (特にマルチユーザ環境)

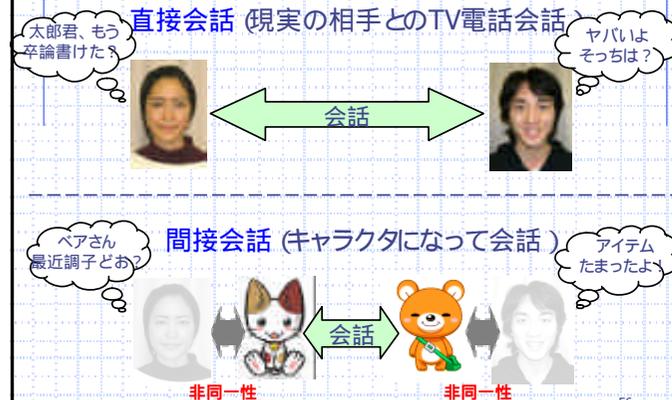
↓
アバタのパーソナライズ



1. 身体動作以外のNV情報の制御

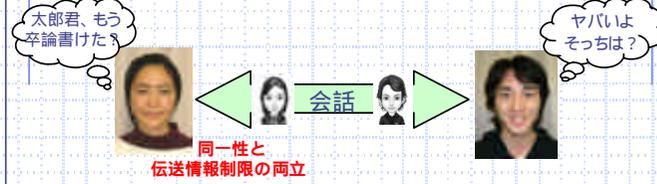
分類	<ul style="list-style-type: none"> アバタの設定次第 現実空間と同様にユーザの嗜好を表現できる × 現実空間では容易に変更できない身体特徴も変更できてしまう (意味が異なる)
身体動作	
周辺言語	
対人距離	
対人接触	
生理現象	あくび, 発汗
身体的特徴	体格, 体型, 容貌, 皮膚, 頭髪
身体装飾	衣服, 装飾品, 化粧, 身体加工

2つのネットコミュニケーション



もうひとつ(?)のネットコミュニケーション

アバタ介在直接会話 (アバタを介した現実の相手と会話)



間接会話 (キャラクタになって会話)



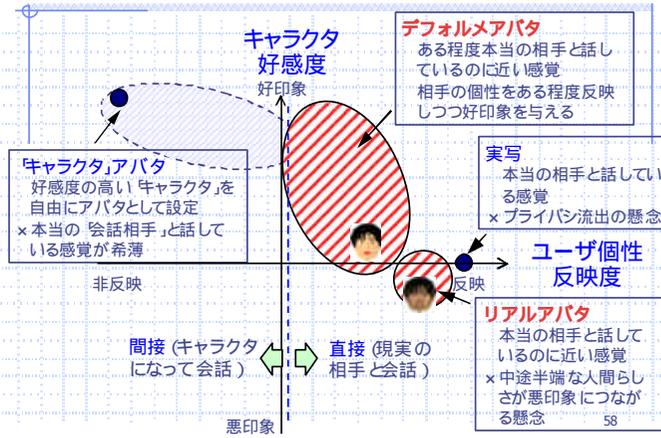
57

おわりに

- ◆ 共有仮想空間でのカジュアルなコミュニケーションは、調整子や感情表示などの身体動作を(モデルに基づいて)自動制御することで、支援できる可能性がある
- ◆ アバタのパーソナライズや、他のノンバーバル情報の制御など検討課題も多く、多方面の研究者の参画が望まれる

59

アバタへの個性の反映とコミュニケーション形態



58

研究協力者

- 宮島俊光技術職員, 田中貴紘助教
- 下地崇, 石井亮, 平家雅之, 熊谷陽をはじめとする研究室学生諸君
- 中野有紀子特任准教授

研究支援

- 文部科学省科学特別教育研究経費
- 文部科学省科学研究費補助金

60