

講演者キャラクタに眼球動作モデルを付与した 音声駆動型複数身体引き込みシステム*

瀬島 吉裕^{*1}, 渡辺 富夫^{*2}, 神代 充^{*3}, 長 篤志^{*4}

A Speech-driven Embodied Group Entrained System with an Eyeball Movement Model for Lecturer Character

Yoshihiro SEJIMA^{*1}, Tomio WATANABE, Mitsuru JINDAI, and Atsushi OSA

^{*1} Graduate School of Science and Engineering, Yamaguchi University
2-16-1 Tokiwadai, Ube, Yamaguchi, 755-8611, Japan

We have already developed a speech-driven embodied group entrained communication system called “SAKURA” for activating group interaction and communication in which speech-driven CG characters called InterActors with both functions of speaker and listener by generating communicative actions and motions are entrained to one another as a teacher and some students in the same virtual classroom. In this paper, the eyeball movement of lecturer in virtual group communication is analyzed by using an embodied communication system with a line-of-sight measurement device. On the basis of the analysis, we propose an eyeball movement model, which consists of a saccade model and a model of lecturer’s gazing to audiences called “group gaze model.” Then, an advanced communication system in which the proposed model is applied to SAKURA is developed for enhancing group interaction and communication. This system generates lecturer’s eyeball movement based on the proposed model by using only speech input. By using the system, we perform experiments for the effects of the proposed model by sensory evaluation. The results demonstrate that the system with the model is effective in group interaction and communication.

Key Words : Human Interface, Man-Machine-Communication, Emotion, Kansei and Embodiment, Human Engineering, Virtual Reality

1. は じ め に

近年の情報機械の進展や情報インフラの整備により, skypeなどの音声を手軽に利用したサービスが充実している⁽¹⁾. さらに, メタバースやオンラインゲームなど, ネットワーク上の3Dバーチャル空間でCGキャラクタを介したコミュニケーションシステムが実用化されている⁽²⁾. これらのシステムでは, CGキャラクタの身体動作や視線情報などがキー入力による予め用意された表現に限定されており, 人と人との行っているような語りかけに対するうなずきや身振りなどの身体的リズムの引き込みによる身体性の共有は考慮されていない. 本来, 人と人との対面で行うコミュニケーションでは, 単に言葉によるバーバル情報だけでなく, うなずきや身振り・手振り, 視線情報などの言葉によらないノンバーバル情報が相互に引き込み合うことで, 相互の身体性が共有され, 円滑なコミュニケーションが行われている⁽³⁾. この身体性の共有が一体感を生み, 人とのかわりを実感させている.

著者らはこれまでに, 身体的リズムの引き込みに着目して, 発話音声から話し手および聞き手の身体動作を自動生成し, 対話者相互の身体性を共有することでコミュニケーションを支援する音声駆動型身体引き込みキャラクタ InterActor を開発している⁽⁴⁾. さらに, この InterActor を複数体用いて, 仮想教室内に集団コミュニケーション

* 原稿受付 2012年4月2日

^{*1} 正員, 山口大学大学院 理工学研究科 情報・デザイン工学系専攻 (〒755-8611 山口県宇部市常盤台 2-16-1)

^{*2} 正員, 岡山県立大学 情報工学部 情報システム工学科 (〒719-1197 岡山県総社市窪木 111)

^{*3} 正員, 岡山県立大学 情報工学部 スポーツシステム工学科 (〒719-1197 岡山県総社市窪木 111)

^{*4} 山口大学大学院 理工学研究科 情報・デザイン工学系専攻 (〒755-8611 山口県宇部市常盤台 2-16-1)

ン場を生成する音声駆動型複数身体引き込みコミュニケーションシステム SAKURA を開発し、その有効性を示している⁽⁵⁾。このような CG キャラクタを介した集団コミュニケーションでは、講演者の身振り・手振りだけでなく、注視時間や視線交差・視線はずしなどの視線情報が重要な役割を果たしている⁽⁶⁾⁽⁷⁾。例えば、井上らは複合現実感技術を用いて複数の CG キャラクタによる遠隔会議システムを開発し、CG キャラクタの位置と向きによって人間の視線行動に影響を及ぼすことを示している⁽⁸⁾。また、藤田らは、対話者の音声情報や発話状態に応じて 3 者間での CG キャラクタの視線情報を制御するシステムを開発し、視線交差や視線はずしを適度な割合で付加することで、相互の発話が促されることを示している⁽⁹⁾。これらのシステムでは、集団コミュニケーションにおける視線情報を実現することを目的としているが、集団コミュニケーションを対象としている以上、集団コミュニケーションにおける人間の視線行動特性に基づくインタフェースを設計・開発する必要がある。そのため、集団コミュニケーション時の行動特性に基づく眼球動作を発話音声から自動生成することで、円滑な遠隔コミュニケーション支援が期待される。

そこで本研究では、集団コミュニケーション時における講演者の眼球動作特性を解析し、その解析結果に基づいて講演者キャラクタの眼球動作を自動生成する眼球動作モデルを提案している。さらに、眼球動作モデルをこれまで開発した SAKURA に適用し、発話音声のみから講演者の身振り・手振りなどの身体動作に加え眼球動作を自動生成する音声駆動型複数身体引き込みシステムを開発している。本システムを用いて遠隔コミュニケーション実験を行い、官能評価により提案モデルおよび開発したシステムの有効性を示している。

2. 集団コミュニケーションにおける視線情報解析実験

本研究では、講義を想定した集団コミュニケーション実験により、集団コミュニケーション時における講演者の眼球動作特性を解析する。

2.1 実験システム

実験システムの概略を図 1 に示す。本実験では、3 台同一のディスプレイ (IO DATA LCD-AD203G) を用意し、被験者と左右のディスプレイとの距離関係が各辺 2m の正三角形で構成されるようにディスプレイを配置した。ディスプレイには、聞き手役となる CG キャラクタ InterActor⁽⁴⁾ を正面に表示し、画面解像度は 1600×1200 とした。InterActor の描画には、Microsoft DirectX 9.0 SDK を使い、Windows XP を搭載した PC (CPU:Corei7 2.93GHz, メモリ:3GB, グラフィック:NVIDIA Geforce GTS250) を使用した。PC から出力された映像をディスプレイ分配器 (ELECOM VSP-A2) に入力することで、3 台のディスプレイに同一かつ同期した映像を提示した。このときの描画フレームレートは 30fps であった。対話者のうなずきや首振りなどの上半身の動きを、対話者の頭部、背部、両手首に取り付けた 4 個の磁気センサ (POLHEMUS FASTRAK) から、位置及び角度情報を 30Hz でサンプリングした。

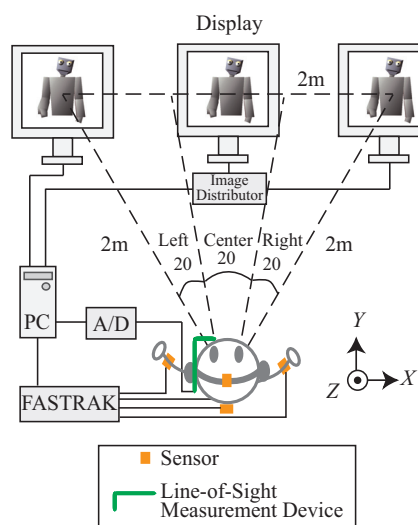


Fig. 1 Setup of the experimental system.



Fig. 2 Example of the communication scene.

眼球動作計測にはこれまでに開発した視線計測デバイス⁽¹⁰⁾を用い、A/D コンバータ (CANOPUS ADV110) によりデジタルデータに変換して PC に入力した。入力画像に対して、予め用意した眼球の角膜部分のテンプレートをを用いてパターンマッチングによる画像処理を行い、眼球の回転角度データを 30Hz でサンプリングした。音声データは、マイク付きヘッドホン (SONY DR-260DP) を用いて 16bit 11kHz でサンプリングした。取得した各種データは、リアルタイムにハードディスクに記録した。

2.2 実験方法

本実験では、講義形式の集団コミュニケーションを想定し、被験者が聞き手役の InterActor に対して語りかける実験を行った (図 2)。実験手順は、まず実験前に眼球動作のキャリブレーションを行い、その後システムを 1 分間試用させた。次に、講義を想定して、被験者が 3 体の InterActor に対して 5 分間語りかける実験を行った。このときの話題は指定せず、日常の出来事やサークル活動などの被験者が話しやすい内容とした。被験者には、集団を見渡すように 3 体の InterActor に対して等しく語りかけるように指示した。InterActor には、被験者の発話音声に基づいてうなずきや身振りなどの身体反応を行わせ、3 体とも同じ反応をさせた。被験者は 22 歳～24 歳までの男子学生 10 人であった。

2.3 集団コミュニケーションにおける視線情報の解析

まず、3 体の InterActor を注視している時間的な割合の解析を行った。解析では、頭部動作データを用い、図 1 に示すように右・中央・左をそれぞれ 20 度として 3 分割し、頭部動作データがその領域内に存在するのであれば注視しているものとした。頭部動作の時系列データとして、Z 軸方向の回転の一例を図 3 に示す。図より、被験者は主に中央 (-10～10 度) を注視している割合が多いことがわかる。また、左右方向に首を振っている回数としては、右側・左側にあまり差がなく、それぞれの InterActor を見ていることがわかる。全被験者の注視している時間的な割合の結果を表 1 に示す。表 1 より、中央を見ている割合が約 60% と最も高く、続いて左側 (-30～-10 度) が約 27%，右側 (10～30 度) が約 13% となり、左右それぞれを見る回数はあまり変わらないが、左側が右側に比べて約 2 倍ほど多く注視していることが示された。これは、人間の空間認知において、右側よりも左側の方が強く注意を惹く (左側優位性) ことが報告されていることから^{(11)~(13)}、本研究においても被験者は右側に比べ左側を注視する割合が多くなったものと考えられる。

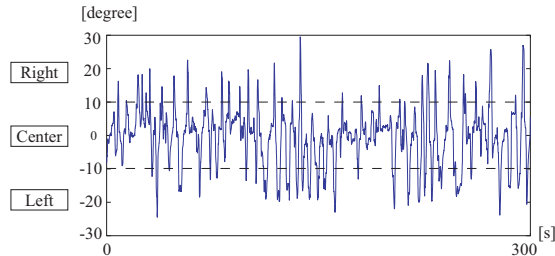


Fig. 3 Example of time change of head movement.

Table 1 Rate of gaze duration in the experimental scene.

	Left	Center	Right
Average	27.6	59.5	12.9
Standard deviation	20.0	24.9	8.7

次に、被験者の頭部動作と眼球動作に着目して解析を行った。なお、被験者は 10 人であったが、2 人分の眼球動作データに情報欠損が生じていたため、8 人分のデータを対象として解析した。頭部動作データ $x(i)$ は、被験者の頭部につけた磁気センサによって得られた Z 軸方向における角度データ $p(i)$ の前後の差 $[p(i+1) - p(i-1)]$ から求まる角速度をその時点での動きの変化量として平滑化した。眼球動作データ $y(i)$ は、視線計測デバイスによって得られた X 軸方向の角度データ $c(i)$ の前後の差 $[c(i+1) - c(i-1)]$ をその時点での動きの変化量として平滑化した。各被験者の頭部動作と眼球動作を次式の相互相関関数 $C(\tau)$ で評価した。

$$C(\tau) = \frac{\sum_{i=1}^{n-\tau} \{x(i) - \mu_x\} \{y(i+\tau) - \mu_y\}}{\sqrt{\sum_{i=1}^n \{x(i) - \mu_x\}^2} \sqrt{\sum_{i=1}^n \{y(i) - \mu_y\}^2}} \quad (1)$$

μ_x, μ_y : x, y の平均値, n : データ数, i : 描画フレーム, τ : 時間遅れ

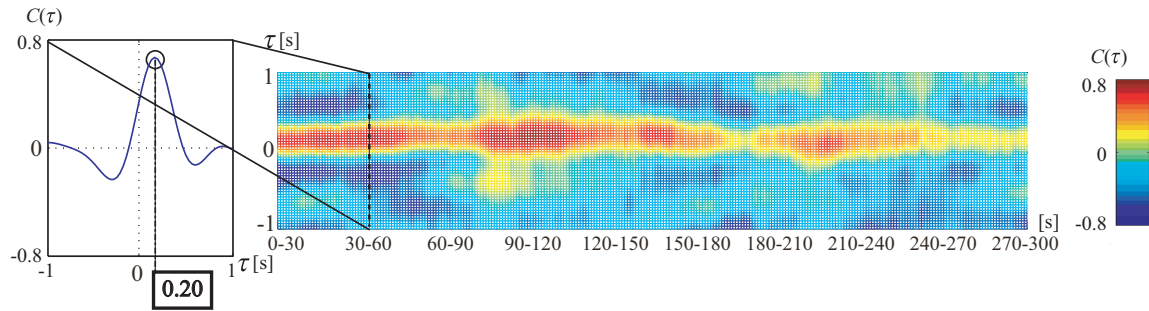


Fig. 4 Example of time change of cross-correlation $C(\tau)$ s between a lecturer's head movement and his own eyeball movement.

ここでは、分析区間を 30 秒として τ を前後最大 1 秒のずれ時間で評価した。頭部動作の左右方向（首振り方向）において、相互相関関数 $C(\tau)$ の分析区間を 1 秒ずつ移動させた時系列変化の一例を図 4 に示す。図より、全分析区間において時間遅れ τ が約 0.20 秒付近に正の相関が観察され、とくに実験開始後の 30 秒から 60 秒の区間においては、 $\tau = 0.20$ で最も強い正の相関が確認できる。これは、眼球動作が頭部動作に対して同方向に 0.20 秒の遅れを伴って動作していることを示している。被験者 8 人における全分析区間の眼球動作の平均遅れ時間は 0.196 ± 0.06 秒であった。この知見は、先行研究⁽¹⁰⁾でのアバタとの対面コミュニケーションを行った際に観察された滑動性眼球運動⁽¹⁴⁾（潜時：0.13 秒程度）や網膜像の安定化を図る代償性眼球運動⁽¹⁵⁾（潜時：0.10 秒以内）とは異なる運動であることがわかる。一方、注視点を素早く変化させる場合に生じるサッケード⁽¹⁴⁾（潜時：0.20 秒程度）と本研究で確認された約 0.20 秒の遅れが非常に類似している。また、実験時の被験者の頭部動作は、左右方向（首振り方向）において大きく変位していることから、眼球は主にサッケード運動を行っているものと考えられる。

以上の結果より、集団コミュニケーション時では、被験者は約 60% の割合で中央を最も多く注視しており、左右それぞれを見る回数はあまり変わらないが、時間的な割合が左側約 27%、右側約 13% の割合で約 2 倍ほど多く注視していることが示された。さらに、注視方向を変更する際には、注視点を素早く移動させるためサッケード運動を行っていることが示された。

3. 眼球動作モデル

講演者の眼球動作解析から、集団コミュニケーションでは眼球は頭部動作に対して 0.20 秒の遅れを伴って動作し、さらに約 60% の割合で中央を注視しており、ついで左側約 27%、右側約 13% の割合で注視する結果が得られた。そこで本研究では、これら講演者の眼球動作特性に基づいて、集団を適度に注視するように講演者キャラクターの眼球動作を自動生成する眼球動作モデルを提案する。この眼球動作モデルは、講演者キャラクターの頭部動作に対して 0.20 秒の遅れを伴って眼球を動作させるサッケードモデルと講演者の集団注視時間特性を基にした集団注視モデルから構成されている。以下にこれらのモデルの概略を示す。

(1) サッケードモデル

サッケードモデルは、集団コミュニケーション時の眼球動作解析結果に基づき、講演者の頭部動作に対して 0.20 秒の遅れを付加して同方向に動作させる。まず、バーチャル空間内における講演者キャラクターの頭部動作方向の角度を算出する。次に、頭部が動作し始めた 0.20 秒後（6 フレーム後）に講演者キャラクターの眼球方向に頭部動作角度のずれ量を加え、遅延動作を含めた講演者キャラクターの眼球回転角度を算出する（図 5）。

(2) 集団注視モデル

講演者の集団注視時間特性の解析結果から、集団コミュニケーションでは全対話時間の約 60% の割合で中央の聞き手を注視し、ついで左側約 27%、右側約 13% の割合で注視していた。そこで、提案する集団注視モデルは、集団注視時間特性解析結果に基づいて、図 6 に示すように集団を注視する確率が中央を 60%、左側を 27%、右側を 13% となるように注視動作を指数分布的に生成する。集団に対してそれぞれ異なる割合で注視することで、継続的な注視を抑制するとともに効果的な視線が生成されるものと期待される。

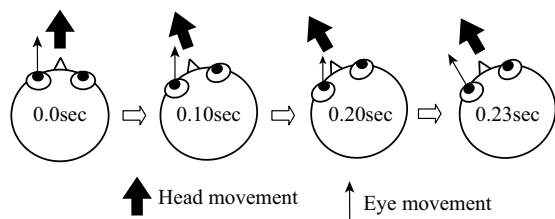


Fig. 5 Example of the saccade model.

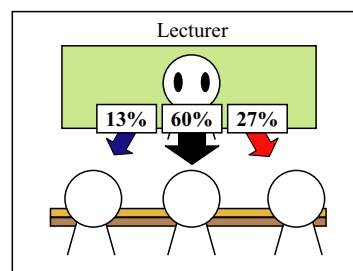


Fig. 6 Group gaze model.

4. 講演者キャラクタに眼球動作モデルを付与した音声駆動型複数身体引き込みシステム

提案した眼球動作モデルをこれまでに開発してきた音声駆動型複数身体引き込みコミュニケーションシステム SAKURA⁽⁵⁾に付与した音声駆動型複数身体引き込みシステムを開発した(図7)。本システムは、Microsoft WindowsXPをOSとしてサウンドカードを搭載するPC(Hewlett-Packard hp workstation xw4200: Pentium4 2.8GHz, メモリ1GB, グラフィックボード NVIDIA Quadro FX3400), マイク(IBM ANC-300)などの音声入力装置から構成される。音声は、16bit 11kHzでサンプリングしている。プログラムは、Microsoft DirectX SDK (Summer 2003)を用いて記述し、CGの描画にはMicrosoft DirectX9.0bを利用した。このときの描画フレームレートは30fpsである。

本システムでは、同一仮想教室に講演者役のInterActorと複数の聴講者役InterActorを配置し、講演者の発話音声に基づいて講演者役InterActorが話し手として振舞い、聴講者役のInterActorが聞き手としてうなずきや身振りなどの身体動作を行うことで、かかわりの場を生成する。さらに、眼球動作モデルを適応することにより、講演者役InterActorが聴講者に対して自然な視線を送り、一体感のある集団コミュニケーション場を生成する。このとき、講演者役InterActorの眼球遅延動作(サッケード)および注視動作は、うなずき反応モデル⁽¹⁶⁾によって生成された頭部動作から自動生成している。

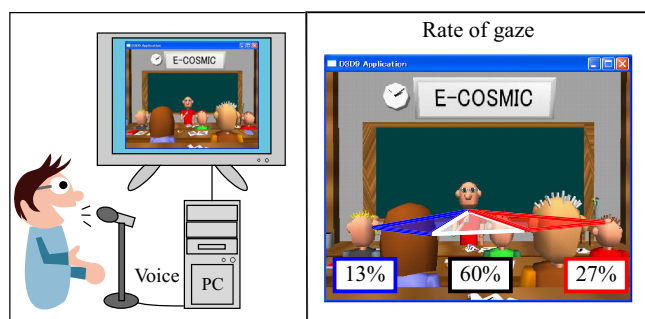


Fig. 7 Configuration of the developed system.



Fig. 8 Example of the experimental scene.

5. 眼球動作モデル評価実験

本研究で提案する眼球動作モデルは、サッケードモデルと集団注視モデルから構成される。本章では、まずサッケードモデルを評価し、次に集団注視モデルにおける注視時間の割合を評価する。

5.1 サッケードモデルの評価

本節では、まずサッケードモデルを評価する。サッケードモデルは、講演者の頭部動作に対して0.20秒の遅れを伴って眼球を動作させるモデルである。そこで、講演者キャラクタの眼球動作遅延時間を0.00秒、0.20秒、0.40秒と変化させて10人の男子学生に評価させた。その結果、10人のうち8人が遅延時間0.20秒を高く評価した。これは、0.20秒が人間の眼球動作特性⁽¹⁴⁾と類似していることから、高く評価されたものと考えられる。また、その他の遅延時間が低く評価された理由として、0.00秒では機械的な運動として認識され、0.40秒では遅延時間が誇張されてしまうためと考えられる。これらのことから、本研究では眼球動作の遅延時間を0.20秒として採用した。

5.2 集団注視モデルの評価

5.2.1 実験概略

本実験では、予め作成した動画をディスプレイに提示し、それを被験者に視聴させた後に一対比較と7段階評価を行わせた（図8）。被験者には、ディスプレイに対して正面に着席させ、一人ずつ動画を視聴させた。動画は、消費税について講義している2分間の音声ファイルを入力として、講演者キャラクタの振る舞いが異なる3種類のモードでシステムを動作させ、各モードごとにディスプレイに表示される映像をビデオカメラにより撮影し使用した。3種類の比較モードは、図9に示すように（A）講演者キャラクタが周囲を注視しないモード（中央注視）、（B）従来手法⁽¹⁷⁾⁽¹⁸⁾である講演者キャラクタが集団を均等に注視するモード（均等注視）、（C）講演者キャラクタが中央を60%、左側27%、右側13%の割合で注視するモード（集団注視モデル）とした。なお、眼球動作の遅延時間は0.20秒として動作させた。被験者は19歳～24歳までの男性15人、女性15人の計30人であった。

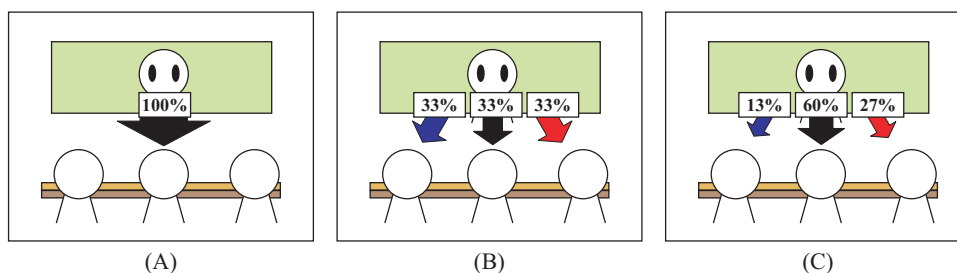


Fig. 9 Outline of each mode in the communication experiment.

5.2.2 実験手順

まず3種類のモードをそれぞれ視聴させた。その後、3種類のモードから2種類のモードをランダムに選定し、2種類のモードについてそれぞれ視聴させ、「総合的に良い」観点から一対比較させた。次に、各モードを視聴させ「(a) 聴講しやすさ」、「(b) (講演における) 視線の自然さ」、「(c) 講演者からの視線」、「(d) 一体感」、「(e) 臨場感」、「(f) 生命感」、「(g) 好み」の7項目について7段階（中立0）で官能評価させた。各モードの提示順は順序効果を考慮しランダムとした。

5.2.3 実験結果

一対比較の結果を表2に示す。表中の数字は各列に対し各行の表示をより好ましいと答えた被験者の数を表している。この結果から表示の好ましさを定量的に評価するために、以下に示すBradley-Terryモデルを想定した⁽¹⁹⁾。

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j} \quad (2)$$

$$\sum_i \pi_i = \text{const.} (= 100) \quad (3)$$

π_i : i の強さの量, P_{ij} : i が j に勝つ確率

Table 2 Result of paired comparison.

	(A)	(B)	(C)	Total
(A)		5	2	7
(B)	25		7	32
(C)	28	23		51

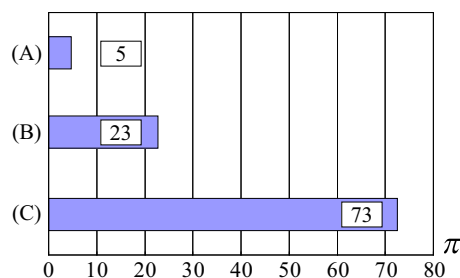


Fig. 10 Comparison of π .

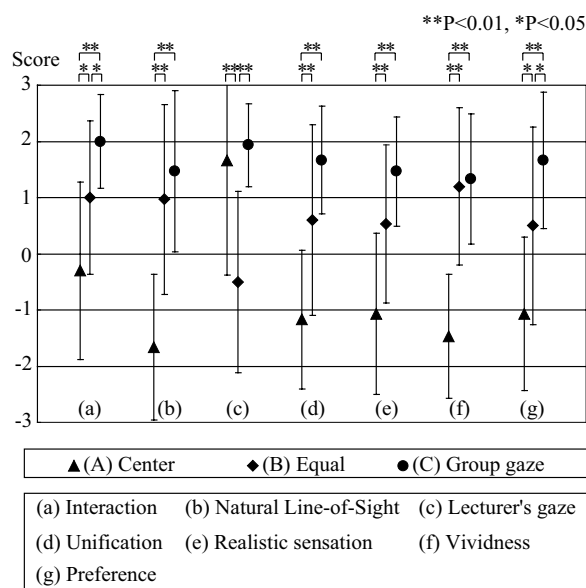


Fig. 11 Seven points bipolar rating.

π_i は 3 種類のモードにおけるモードの好ましさ (3 モードの合計 100) を表し、このモデルを想定することにより、一対比較に基づく好ましさを定量的に定めることができる。Bradley-Terry モデルより推定されたモードの好ましさ π を図 10 に示す。このモデルの整合性を検討するために、有意水準 5% で適合度検定及び尤度比検定を行った結果、モデルは棄却されず、 π の妥当性が保証された。図 10 より、提案モデルであるモード (C) が他のモードに比べて極めて高く評価された。次に高く評価されたのが、従来手法である集団全体を注視するモード (B) であり、最も低く評価されたのは集団を注視しないモード (A) であった。

また、7 段階評価の結果を図 11 に示す。図には各項目の平均値とその標準偏差を示している。Friedman の分散分析法及び Nemenyi の多重比較を行った結果、提案モデルであるモード (C) は、モード (A) に対して「(a) 聴講しやすさ」、「(b) (講演における) 視線の自然さ」、「(d) 一体感」、「(e) 臨場感」、「(f) 生命感」、「(g) 好み」の項目で有意水準 1% の有意差が認められた。また、モード (C) とモード (B) と比較した結果、「(c) 講演者からの視線」の項目で有意水準 1%、「(a) 聴講しやすさ」、「(g) 好み」の項目で有意水準 5% の有意差が認められた。これは、講演者の被験者への注視割合が適度に多くなることで講演者への信頼感が向上することが報告されていることから⁽²⁰⁾、本研究においても注視時間の割合が適度に多い提案モデルが高く評価されたものと考えられる。

これらのことから、講演者が集団を注視する場合は、集団全体を均等に見渡すのではなく、集団注視時間特性の解析結果と同様の中央を 60% の割合で注視させることで、聴講しやすさや講演者からの視線が高まることが示された。

6. 集団注視時間の評価実験

前章では、講演者キャラクタが中央を 60% の割合で注視することで、聴講しやすさや講演者からの視線が高まるなど、コミュニケーション効果が示された。本章では、残りの注視時間 40% に対して、左右の注視割合を変化させた場合のコミュニケーション効果を検討する。

6.1 実験概略

実験手順は前章と同様のものとした。本実験では、中央を注視する割合を 60% に固定し、左右の注視割合を変化させた 3 種類の比較モードを選定した (図 12)。比較モードは、提案モデルである (C) 講演者キャラクタが中央を 60%、左側 27%、右側 13% の割合で注視するモード (集団注視モデル)、(D) 講演者キャラクタが中央を 60%、左側 13%、右側 27% の割合で注視するモード (左右反転)、(E) 講演者キャラクタが中央を 60%、左側 20%、右側 20% の割合で注視するモード (左右平等) として実験を行った。各モードの提示順は順序効果を考慮しランダムとした。被験者は前章の実験とは異なる 19~24 歳までの男性 15 人、女性 15 人の計 30 人であった。

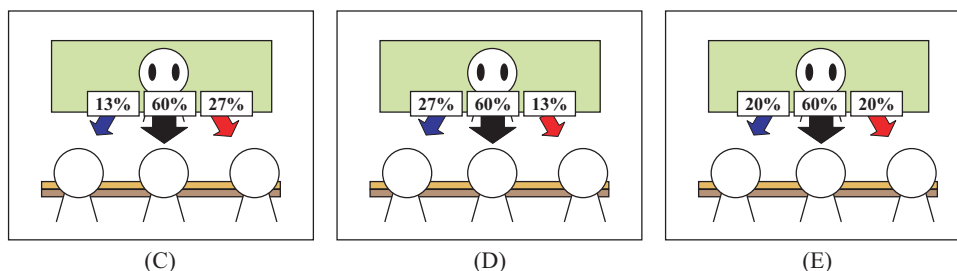


Fig. 12 Outline of each mode in the experiment.

6.2 実験結果

一対比較の結果を表 3 に示す．この結果から Bradley-Terry モデルでのモードの好ましさ π を図 13 に示す．有意水準 5% で適合度検定及び尤度比検定を行った結果，モデルは棄却されず， π の妥当性が保証された．図 13 より，提案モデルであるモード (C) が最も高く評価され，次いでモード (E)，(D) の順となった．

次に，7 段階評価の結果を図 14 に示す．図には各項目の平均値とその標準偏差を示している．Friedman の分散分析法及び Nemenyi の多重比較を行った結果，「一体感」の項目においてモード (C) とモード (D)，およびモード (E) とモード (D) との間に有意水準 5% の有意差が認められた．これは，人間の空間認知が左側優位^{(11)~(13)}であるのに対して，モード (D) は右側優位として講演者キャラクタに振舞わせたため，講演者キャラクタを含む仮想教室に違和感を感じたものと考えられる．

以上の結果から，講演者が集団を注視する場合，中央を 60 % 注視し，左右の注視割合については，左側を多く注視する，あるいは均等に注視することで，一体感が向上するコミュニケーション効果が示された．このことから，提案する眼球動作モデルおよび開発したシステムの有効性が示された．

Table 3 Result of paired comparison.

	(C)	(D)	(E)	Total
(C)		24	14	38
(D)	6		10	16
(E)	16	20		36

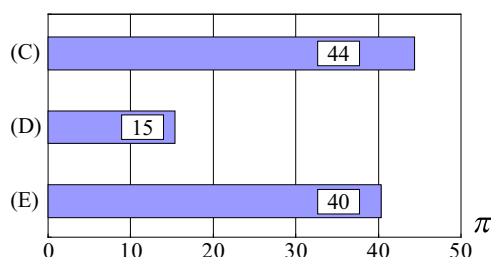


Fig. 13 Comparison of π .

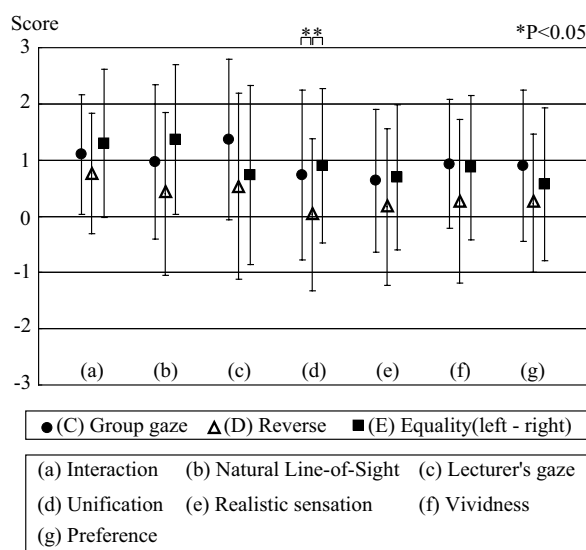


Fig. 14 Seven points bipolar rating.

7. お わ り に

本研究では、集団コミュニケーション時における講演者の眼球動作特性を解析し、その解析結果に基づいてサッケード運動を自動生成するサッケードモデルと講演者の集団注視時間特性を基にした集団注視モデルから構成される眼球動作モデルを提案した。さらに、音声駆動型身体引き込みキャラクタ InterActor を用いて仮想教室内に講演者と複数の聴講者を配置した身体的コミュニケーションシステムに眼球動作モデルを適用した音声駆動型複数身体引き込みシステムを開発した。開発したシステムを用いてコミュニケーション実験を行い、官能評価を行った。その結果、人間の集団注視特性を有する集団注視モデルを用いることで、聴講しやすさや一体感の増大に繋がるなど提案モデルおよび開発したシステムの有効性を示した。

今後は、情動変動に伴う眼球動作の検討や講演者の注視情報に応じて聴講者の振舞いを変化させるなど、講演者の視線情報を用いてコミュニケーション場の制御を行い、人を引き込む身体性メディア場の生成・制御技術を様々な観点からのアプローチにより確立していく予定である。

謝 辞

本研究の一部は、科学技術振興機構 (JST) の戦略的創造研究推進事業 (CREST) 研究領域「デジタルメディア作品の制作を支援する基盤技術」における「人を引き込む身体性メディア場の生成・制御技術」プロジェクトの支援による。また、本研究を行うにあたり、システム評価・分析に協力いただいた圖師悠香理氏に謝意を表します。

文 献

- (1) Skype Limited, “skype”, skype, <http://www.skype.com/> (参照日 2012 年 3 月 15 日) .
- (2) CyberAgent, Inc., “アメーバピグ”, アメーバピグ, <http://pigg.ameba.jp/> (参照日 2012 年 3 月 15 日) .
- (3) Condon, W. S. and Sander, L. W., “Neonate Movement is Synchronized with Adult Speech”, *Science*, Vol. 183, (1974), pp. 99-101.
- (4) 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 (2004), pp. 43-60.
- (5) 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 ECOSMIC の開発を通して - ”, ベビーサイエンス, Vol. 2, (2003), pp. 4-12.
- (6) Kendon, A., “Some Functions of Gaze-direction in Social Interaction”, *Aca Psychologica*, Vol. 26, (1967), pp. 22-63.
- (7) 磯友輝子, 木村昌紀, 桜木亜季子, 大坊郁夫, “視線行動が印象形成に及ぼす影響 - 3 者間会話場面における非言語的行動の果たす役割 - ”, 対人社会心理学研究, No. 4 (2004), pp. 83-91.
- (8) 井上智雄, “実対人距離を調節可能な複合現実分散会議システム”, 情報処理学会論文誌, Vol. 50, No. 1 (2009), pp. 246-253.
- (9) 石井亮, 宮島俊光, 藤田欣也, “アバタ音声チャットシステムにおける会話促進のための注視制御”, ヒューマンインタフェース学会論文誌, Vol. 10, No. 1 (2008), pp. 87-94.
- (10) 瀬島吉裕, 渡辺富夫, 神代充, “音声駆動型身体引き込みキャラクタに眼球動作モデルを付与した身体的コミュニケーションシステム”, 日本機械学会論文集 C 編, Vol. 76, No. 762 (2010), pp. 340-350.
- (11) 加藤孝義, “空間感覚の心理学 左が好き? 右が好き?”, (2001), pp. 37-47, 新曜社.
- (12) クリス・マクナス, “非対称の起源 ~ 偶然か、必然か ~ ”, (2006), pp. 300-349, 講談社.
- (13) 根平邦人, “生物界の左と右”, (1998), pp. 103-106, 共立出版.
- (14) 福島順子, “総説 精神分裂病の眼球運動障害”, 北海道大学医療技術短期大学部紀要, Vol. 2, (1989), pp. 1-13.
- (15) Ohno, T., “What Can Be Learned From Eye Movement?: Understanding Higher Cognitive Processes From Eye Movement Analysis”, *Transactions of the Japanese Cognitive Science Society*, Vol. 9, No. 4 (2002), pp. 565-579.
- (16) 檀原龍正, 渡辺富夫, 大久保雅史, “音声駆動型身体引き込みキャラクタ InterActor が発話音声に与える効果”, 日本機械学会論文集 C 編, Vol. 71, No. 712 (2005), pp. 152-159.

- (17) 荒巻基文, “プレゼンテーションの技術”, (2010), pp. 39-42, 産業能率大学出版部.
- (18) 別所栄吾, “あなたの話は、なぜ伝わらないのか?”, (2011), pp. 188-190, 日本経済新聞出版社.
- (19) 竹内啓, “現象と行動のなかの統計数理”, (1978), pp. 133-148, 新曜社.
- (20) 小野浩亨, 天根哲治, “教授場面における教師の視線行動が生徒による教師態度の認知に及ぼす影響”, 日本教育心理学会総会発表論文集, No. 28 (1986), pp. 494-495.