

# 人間の視線を模擬したモデルを持つ リアクティブバーチャルヒューマンの構築

三武 裕玄\*      長谷川 晶一\*      小池 康晴\*      佐藤 誠\*

東京工業大学\*

## 概要

人間とバーチャルヒューマンとのインタラクションにおいて、相手の視線はリアリティの向上のために重要な要素である。本研究では、物理シミュレーションベースのバーチャルヒューマンを対象とし、人間らしい自然な眼の動きを与えるため、人間の視覚系を模擬したモデルを構築した。具体的には、物理シミュレータの持つ物体の位置情報をもとに、人間の視覚的注意を模擬した手法によって注視点を決定し、眼球運動の制御モデルによって眼球を動作させた。

## Constraction of Reactive Virtual Human with Gaze System Mimiced Human Ocular System

Hironori MITAKE\*      Shoichi HASEGAWA\*      Yasuharu KOIKE\*

Makoto SATO\*

Tokyo Institute of Technology\*

## Abstract

The gaze movement is important factor to improve reality of a virtual human, while an interaction between a human and the virtual human. We constructed a model mimiced human ocular system into phisical simulation based virtual human system, due to create natural and humanlike gaze movement. We propose a system which gets visual information from phisical simulator, determines gaze direction with a method mimiced human visual attention, moves eyes using a model of oculomotor control.

## 1 はじめに

本稿では、3次元仮想空間内のインタラクティブな人間のモデル、すなわちバーチャルヒューマン(以降 VH と略)の自然な視線の実現方法について述べる。

VHの技術はゲームや映画の製作等に近年使用されており、特にゲーム等インタラクティブ性の高いコンテンツでは、ユーザの入力に対してリア

ルタイムで反応することが求められる。一方で、インターフェースの発展により、ユーザからの入力のバリエーションはますます増大しており、多様な入力に対して多様な反応を行うことも求められる。

現在主流の動作生成法に、あらかじめ作り込んだ動作パターンを保持しておき、再生するという方法がある。しかしこの方法では入力のバリエーションが増加するにしたがってコンテンツ製作者

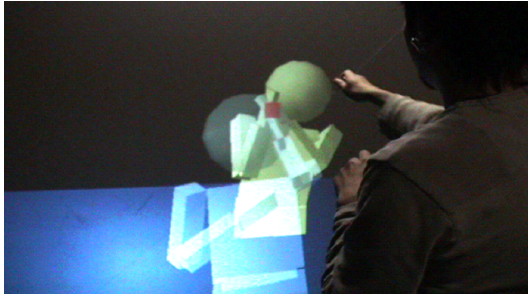


図1 バーチャルボクシング

の労力も急激に増加してしまう。

そこで、自律的にリアルな動作を生成する VH が提案されている。実現方法の一つとして、人間の身体・感覚・運動をシミュレートする事で人間と同様の動きや反応を行う VH を生成する手法があり、長谷川らが構築を行っている [1]。

ところで、人間において、視線の果たす役割は重要なものである。

例えば、長谷川らが構築している VH[1] では、アプリケーションとして人間対 VH によるボクシングゲームを構築している。ボクシングのように人間が向かい合って互いに攻撃しようとする状況で考えられる視線には次のようなものがある；

- 相手がパンチしてきた時、自然と視線が相手の手へ向く
- 相手を攻撃する直前、より正確に目標を捉えるため、視線が攻撃予定位置を向く
- 相手の視線を読むため、極力相手の顔を見ようとする。

また、ボクシングの試合中であるという条件を考えなければ、次のような視線も考えることができる；

- 相手の前で手を振ると、手の方を見る
- 相手の視野外から手を伸ばすと、視野に入った時点で気付いて振り向く

これら視線の移動は、その視線を観察している者に、相手が周囲の状況を認識している事を感じさせ、相手の意図を読み取る手がかりを与える。例えば相手の前で手を振って視線が変化すれば、

相手が自分の手の動きに気が付いたと分かるし、攻撃直前に攻撃予定位置を向く視線からは、相手が攻撃しようとしている意図を読み取ることができる。これを人間と VH とのインタラクションにおいても実現することで、我々に VH を「感覚を持って行動している存在」であると感じさせ、VH のリアリティを向上させることが可能であると考え。

以上より、本研究では、人間の視覚系を模擬したモデルの構築、眼球運動制御のモデルの実装を行う事で、人間と同様の動きをする視線を持った VH の構築を試みた。

## 2 関連研究

力覚及び視覚によるインタラクションが可能なバーチャルヒューマンに関して、長谷川ら [1] や鄭ら [3] が研究を行っている。[1] では人間の身体・感覚・運動のをデル化しシミュレーションを行う方法によって、[3] ではモーションキャプチャで取得した人間の動作をデータベース化し、それらを状況に応じて加工・再生することによって構築を行っている。これらの VH には、視線は構築されていない。

[4] では視線を持つ VH の構築が行われているが、VH を取り巻く環境に対する視線の移動をアニメーションとして生成する事を目的としており、VH とのインタラクションは実現できていない。

視線を持ち、インタラクションの可能な VH としては、対話エージェントを対象とした研究が行われている。[5] では対話インタラクションにおける視線の重要性を VH を使用して評価しており、[6] では対話インタラクションの可能な、視線を持つ VH の構築を行っている。これらの研究で対象とされているのは音声による対話であり、ユーザの動作に対して視線を生成するといった事は行われていない。

[8] では、対話インタラクションにおけるユーザの動作に対して、視線移動を行う VH を構築している。カメラでユーザの動作を取得し、動きの大きい部分に対して視線を向けるといった反射的

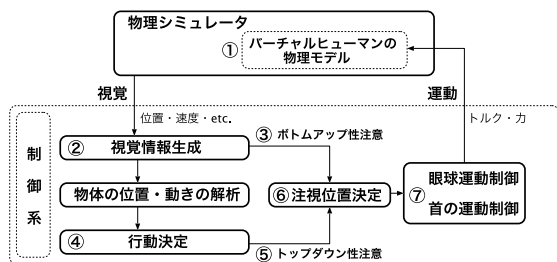


図2 提案システムの構成

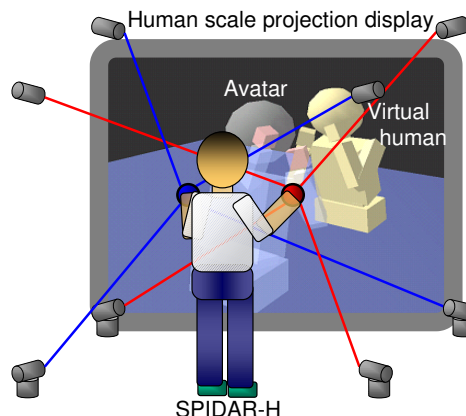


図3 バーチャルボクシングのインタラクション環境

な視線の生成を行っている。

本研究では、ユーザの動きを入力とした VH とのインタラクションを対象とし、視線の生成を行った。ユーザの動きに対する反射的な視線だけでなく、VH の意図を反映した視線の生成も行い、ユーザが VH の意図を視線から感じられる事、及び、反射的な視線・意図的な視線それぞれの優先度を調整することによって様々な性質の VH を表現可能にする事を目指した。

### 3 提案システム

#### 3.1 システムの機能

人間の視線は視覚的注意の機構によって制御される。視覚的注意は二種類に大別することができる。一つは視覚像から直接誘発され、「つい注意を向けてしまう」ような性質のボトムアップ性注意、もう一つは意識的に向けるトップダウン性注意である。

1 節後半で述べたような視線移動も、

#### ボトムアップ

- パンチする手を自然と見る視線
- 目の前で振られる手を見る視線
- 突然視界に入って来た手を見る視線

#### トップダウン

- 攻撃前に攻撃予定位置を向く視線
- 極力相手の顔を見ようとする視線

のように分類することができる。よって、VH に対してボトムアップ・トップダウンの両注意機構を用意することで、人間の視線移動に近い動きを持った視線が実現できると考える。

また、ボトムアップ性注意は視覚像に基づいた

め、VH に「視覚像」の概念を組み込む必要がある。視覚像は、人間の場合環境を視覚で捕えて得る物であり、「『どこに』『何が』あるのか」の情報を含む。ここで、VH が存在するのは物理シミュレーションされた世界であるため、物理シミュレータから「『どこに』『何が』あるのか」の情報を取得することで、視覚像の代用とする。

以上の考えに基づき、次のようなシステムを提案する。

- VH は、物理シミュレータから情報を取得し「視覚情報」とする。
- 視覚情報に基づき、ボトムアップ性注意を生成する。
- VH の行動計画に基づき、トップダウン性注意を生成する。
- 両注意の結果を統合し、注視点を決定する。
- 視線が注視点へ向かうよう、眼球運動と頭部運動を連携して制御する。

#### 3.2 システム構成

次に、提案システムの具体的な構成について述べる。

本システムは、ユーザとのリアルタイムでのインタラクションを可能にするため、長谷川らの構築したバーチャルボクシング [1] をベースとして構築した。VH の物理モデル・行動計画・腕の運動出力についてはバーチャルボクシングの物を使用している。

バーチャルボクシングでは、VH及びユーザの操作するアバタが物理シミュレーションされた世界に存在する。ユーザは力覚インターフェースを用いてアバタを操作し、VHとのインタラクションを行う(図3)。これはユーザの動きを入力としたインタラクションであると言える。

システムの構成を図2に示す。システムは、物理シミュレータ、及び、VHの制御系から成る。VHの制御系は、物理シミュレータから情報を取得する感覚系、物理シミュレータ内の身体に運動出力を行う運動系から成る。「感覚系」としては、今回は視線の構築が目的であるため、視覚系を模擬した構築を行った。また「運動系」としては、視線の変化をもたらす主要な要素である眼球の運動、及び頭部運動の制御を構築した。

図2における処理の手順は次の通りである。番号は図中の番号に対応する。

- ① 物理シミュレータによってVH及びアバタの運動がシミュレートされる。
- ② 物理シミュレータから、VHの視野に含まれる全ての物体の位置情報を取得し、それを列挙したリストを作成する。(このリストがVHの「視覚情報」である。)
- ③ 「視覚情報」を処理し、特徴的な動きを行っている物体を見つける。この段階で、「ボトムアップ性注意」の位置が決定される。
- ④ 視覚情報やその他の情報を統合し、VHの行動を決定する。
- ⑤ 行動決定にしたがって、「トップダウン性注意」の位置を決定する。
- ⑥ ボトムアップ性注意情報、及びトップダウン性注意情報を統合し、最終的な注視位置を決定する。
- ⑦ 決定された注視位置を注視するよう、眼球と首に与えるトルクを制御し、VHの視線を移動させる。

### 3.3 システム構成要素

以下では、システムを構成する各要素についての詳細を述べる。

表1 VHの眼球の仕様

直径	22mm
水平方向可動域	±55度
垂直方向可動域	±45度
頭頂からの垂直距離	125.5mm
瞳孔間幅	61.9mm

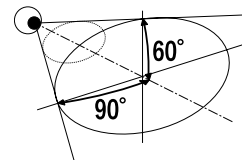


図4 眼球一つに対する視野

#### 3.3.1 眼球の物理モデル

VHの物理モデルに、眼球に相当する剛体及び関節を搭載した。VHの眼球の仕様は表1に示す通りである。眼球の大きさ及び形状は、文献[10]を参考にした。眼球の位置はAIST人体寸法データベース[11]より成人男性の平均値を使用した。

#### 3.3.2 視覚情報生成

物理シミュレータは、シミュレート対象となる全ての剛体の位置・姿勢・速度・角速度などの情報をシーングラフという形態で保持している。

まず、視覚情報生成部は、シーングラフ中の各剛体について、VHの視野(図4)の内部に、剛体の中心が含まれるかどうかをチェックする。含まれる場合、その剛体は「可視」とする。

次に、可視であった剛体全てについて、位置・姿勢・速度・角速度の情報を取得し、リストを生成する。このリストをその時刻での「VHの視覚情報」とする。

#### 3.3.3 視覚的注意生成

視覚情報に基づき、視線移動の目標位置を決定する。人間の視覚的注意のメカニズムを模擬し、構築した。

■ **ボトムアップ性注意** ボトムアップ性注意は、映像中で特徴的な形状・色・動きなどを有する位置に自然と誘発される注意である。

今回はユーザの動きを入力としたインタラク

ションを対象としており、ユーザの「動き」にしたがって誘発されるボトムアップ性注意を構築した。

3次元空間における剛体の動きを網膜のような平面に投影した際の、特徴となる主な量として「並行移動」「拡大縮小」「回転」の三種類がある。大脳の視覚野の中でも動きに反応する細胞が多く見られる部位として知られるMT野・MST野に、これらの要素に反応する細胞が存在する事が分かっている。そこで本研究では、ボトムアップ性注意を生成する際、これらの量が特徴的な値を示すような位置が注意を引き付けると考える。

以上から、可視な剛体 $i$ について

$$A_i = \left| v_i - (v_i \cdot e) \frac{e}{|e|} \right| + \left| (v_i \cdot e) \frac{e}{|e|} \right| + \left| (L_i \cdot e) \frac{e}{|e|} \right| \quad (1)$$

( $e$ : 視線方向,  $v_i$ : 剛体 $i$ の速度,  
 $L_i$ : 剛体 $i$ の角速度)

の値を計算し、剛体 $i$ が「注意を引き付ける量」とする。この値を全ての可視な剛体について計算し、最大値を実現する剛体をボトムアップ性注意の対象とした。

■ **トップダウン性注意** トップダウン性注意は、意識的に特定の位置に向けるような注意である。このような視線移動は意識して行う「行動」であり、VHの行動計画が関わって注視位置が決定されると考えた。

本システムでベースとしたバーチャルボクシングでは、VHの行動は「特定箇所を攻撃(パンチ)」「相手の攻撃を防御」「待機」の三種類であり、行動計画部によってこのいずれかが選択される。

そこで、選択された行動に対し次のようなルールを適用する。

行動	注視点	注意を引き付ける量
攻撃 ⇒	攻撃目標	$A_{attack}$
防御 ⇒	防御目標	$A_{guard}$
待機 ⇒	相手の顔	$A_{face}$

各ルールには、ボトムアップ性注意の時と同様の「注意を引き付ける量:  $A_{topdown}$ 」を付与する。この値はそれぞれ定数とした。

これにより、VHの行動が決定されると同時に、対応して注視すべき位置が一つ決定される。この位置をトップダウン性注意の対象とした。

### 3.3.4 注視位置の決定

ボトムアップ性注意の位置・トップダウン性注意の位置のうち、どちらが最終的な注視位置として選択されるかは、ボトムアップ・トップダウンの両注意のどちらがより注意を引き付けるかによる。

ボトムアップ性注意の対象として選択された剛体の注意を引き付ける量:  $A_i$  と、選択されたトップダウン性注意に付与された注意を引き付ける量:  $A_{topdown}$  を比較し、より値の大きい方を最終的な注視点として選択するものとした。

### 3.3.5 眼球運動制御

ヒトの眼球運動は、主に、サッケード及び追従運動の二種類に分類することができる。この二種類の眼球運動の制御について、実装を行った。

■ **サッケード** 3.3.3節で述べた方法によって決定される注視位置が、前回の位置から大きく変更されている場合、視線を高速に跳躍させるサッケードが起こる。

サッケード運動は、式(2)で表現し、さらに最大角速度を制限することによって近似した。具体的には、偏角  $A[rad]$  のサッケードを行う場合、サッケード開始から時刻  $t[sec]$  が経過した時点における眼球の位置  $\theta_{eye}(t)$  を次式によって決定する。

$$f(t) = (6(t/T)^5 - 15(t/T)^4 + 10(t/T)^3) \quad (2)$$

$$A' = L/T'(0.5T) \quad (3)$$

$$t_f = ((A - A')/L + 0.5T) \quad (4)$$

$Af'(0.5T) < L$  の場合

$$\theta_{eye}(t) = Af(t) \quad (5)$$

$Af'(0.5T) \geq L$  の場合

$$\theta_{eye}(t) = \begin{cases} A'f(t) & (0 \leq t < 0.5T) \\ Lt + 0.5A' - 0.5LT & (0.5T \leq t < t_f) \\ A'f(t - (A - A')/L) + (A - A') & (t_f \leq t < (t_f + 0.5T)) \end{cases} \quad (6)$$

ただし

$L$ (眼球の最大速度): 8.7266[rad/s]

$T$ (サッケードに要する最短時間): 0.05[s]

$L$ ,  $T$  の値は, [10] 及び [12] による. 決定された位置  $\theta_{eye}$  に対し, 眼球に与えるトルクを PD 制御することにより, 運動を実現する.

■ **追従運動** 注視位置が前回の位置から大きく変更されない場合は, サッケードではなく追従運動制御により眼球運動の制御を行う.

追従運動の制御モデルとしては, 張らによる両眼眼球運動制御モデル [13] の実装を行った.

これは人間の脳における眼球運動制御機構をモデル化しており, 両眼眼球運動に特有な輻輳・共役運動, また, 頭部の運動を補償する前庭動眼反射を再現する.

両眼運動制御器によって出力された眼球の目標位置  $\theta_{eye}$  に対し, 眼球に与えるトルクを PD 制御することにより, 運動を実現する.

### 3.3.6 頭部運動制御

注視位置が眼球の可動範囲外であった場合, 注視位置を眼球の可動範囲内に収めるよう, 頭部運動が誘発される.

■ **頭部運動の開始条件** 頭部運動が誘発されるのは, 注視位置が顔正面からおおよそ  $\pm 30^\circ$  以上の位置にあった場合である [14]. この時, 頭部運動の開始は眼球の運動開始から 20ms 後であるが, 注視位置が予測可能な場合や能動的探索に伴う視線移動等の場合は, 頭部運動が先に起こる事もある [14].

以上より, 頭部運動の開始条件は, 次の二つを同時に満たす事とした.

- 注視位置が顔正面から  $\pm 30^\circ$  以上
- 眼球運動開始から 20ms が経過している, または, トップダウン性注意に基づく注視である

■ **目標方向決定** 注視位置に対し, 首がどの程度まで運動するかは状況に依存する. 最も動く場合は顔面が注視位置に対して正面を向き, 最も動かない場合は注視位置の方向の  $30^\circ$  手前までで止まる [14]. 文献 [14] では, 被験者が「正確な視線

移動を行う」ように求められた場合に首が最大の運動を行うなど, 首が運動する程度は被験者の意志に依存するとしている.

本システムでは, 3.3.4 節で述べた「注意を引き付ける量」が大きいほど, VH にとって「重要な視線移動」であると言え, この場合に「正確な視線移動を行うよう求められた人間の被験者」同様, 首が最大の運動を行うと考える.

よって, 頭部運動方向は, 注視方向  $-30^\circ$  から注視方向正面まで, 注意を引き付ける量に従って滑らかに変化するよう, 次のような式によって決定した.

$$\theta_{head} = \theta_{attention} - \frac{30}{1 + e^{-7\left(\frac{2A - 2A_1}{A_2 - A_1} - 1\right)}} [\text{度}] \quad (7)$$

ただし

$\theta_{attention}$ : 注視位置

$A$ : 注意を引き付ける量

$A_1$ : 頭部運動が最小となる下限値

$A_2$ : 頭部運動が最大となる上限値

■ **運動制御** この目標方向  $\theta_{head}$  に対し, 実際の頭部運動は

$$\theta'_{head} = \theta_{head}(6s^5 - 15s^4 + 10s^3) \quad (8)$$

(ただし頭部運動開始からの経過時間を  $t$ ,

目標到達時刻を  $t_f$  としたとき  $s = t/t_f$ )

によって与えられる副目標位置  $\theta'_{head}$  に頭部を向けるよう PD 制御する事によって行う. 式 (8) は, バーチャルボクシングにおいて躍度最小な到達運動を実現するために用いられた式であり, “滑らかな” 運動を実現する. 今回は, 滑らかな頭部運動の実現のために使用した.

## 4 結果・考察

実際に本システムを用いて次のようなインタラクションを行った.

- ユーザが VH を攻撃し, VH は防御する
- VH がユーザを攻撃する
- VH の前でユーザが手を振る



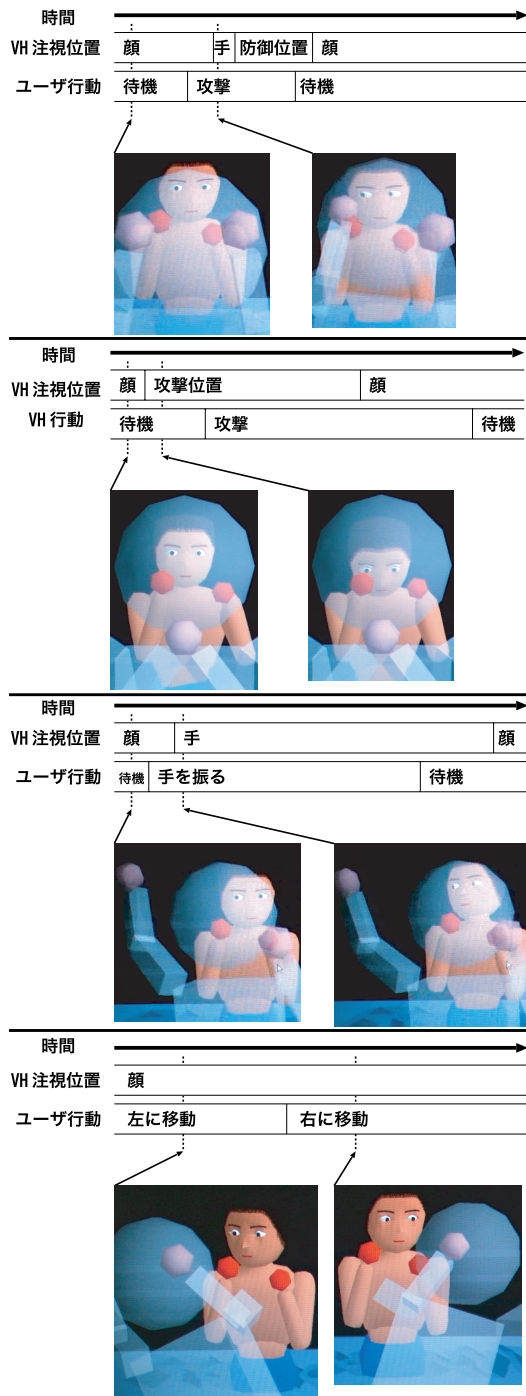


図5 実験結果

- VHの前をユーザが左右に移動する

その結果、VHの視線移動は図5のようになった。この事から、1節で述べた

- パンチする手を自然と見る

- 攻撃前に攻撃予定位置を向く
- 目の前で振られる手を見る
- 極力相手の顔を見ようとする

といった視線は実現していると言える。

一方、ボクシングの試合を記録したビデオにおいて、次のような視線を観察することができた。

- 極力相手の顔を見ようとする。殴られた等の理由で視線が逸れた場合も、すぐに視線を相手の顔に戻そうとする。
- 相手が攻撃してくる際に、攻撃を行う相手の手を見る事がある。ただし一瞬である。
- 攻撃前に攻撃予定位置を向く事がある。

このように、実際のボクサーは相手の顔を注視する傾向が強く、それ以外の場所を注視する場合があっても一瞬だけであった。これは、「相手の顔を注視」というトップダウン性注意の優先度が非常に大きいためではないかと考えられる。

本システムにおいては、3.3.4節で述べた『トップダウン性注視の重要度』の数値を変化させることで、この傾向を変化させる事ができる。数値の調整により、「強そうなボクサー」「弱そうなボクサー」など、さまざまな性質のVHを実現することができる可能性があり、今後検証していきたいと考えている。

## 5 課題

今後の課題として、提案システムがVHのリアリティ向上のために有効か評価を行う事を考えている。

また、ボクシング以外のアプリケーションにおける視線の構築も課題である。特にトップダウン性注意についてはボクシングの行動計画に基づいており、ボクシング以外のVHのアプリケーションでは、そのVHにおける行動計画に基づいたものを作り直す必要がある。トップダウン性注意を生成するためのより汎用な方法の検討を含め、今後の課題としたい。

## 参考文献

- [1] S.Hasegawa, T.Ishikawa, N.Hashimoto, M.Salvati, Y.Koike, M.Sato:  
**“Human Scale Haptic Interaction with a Reactive Virtual Human in a Realtime Physics Simulator”**,  
ACM SIGCHI Intl. Conf. on Advances in Computer Entertainment Technology ACE 2005(2005).
- [2] S.Hasegawa, M.Sato:  
**“Real-time rigid body simulation for haptic interactions based on contact volume of polygonal objects”**,  
Computer Graphics Forum, September 2004.
- [3] S.Jeong, N.Hashimoto, M.Sato:  
**“A novel interaction system with force feedback between real - and virtual human: an entertainment system: “virtual catch ball””**,  
Proc. of the 2004 ACM SIGCHI Intl. Conf. Advances in computer entertainment technology,p.61-66(2004).
- [4] C.Peters, C. O’ Sullivan:  
**“Bottom-Up Visual Attention for Virtual Human Animation”**,  
Computer Animation for Social Agents(2003).
- [5] M.Garau, M.Slater, S.Bee, M.A.Sasse:  
**“The Impact of Eye Gaze on Communication using Humanoid Avatars”**,  
Conference on Human Factors in Computing Systems, Proc. of the SIGCHI Conf. on Human factors in computing systems, p.309-316(2001).
- [6] V.Vinayagamorthy, M.Garau, A.Steed, M.Slater:  
**“An Eye Gaze Model for Dyadic Interaction in an Immersive Virtual Environment: Practice and Experience”**,  
Computer Graphics Forum 23(1),p.1-12(2004).
- [7] S.P.Lee, J.B.Badler, N.I.Badler:  
**“Eyes alive”**,  
ACM Transactions on Graphics 21(3),p.637-644(2002).
- [8] 広田 健一, 益子 宗, 星野 准一:  
**“インタラクティブ CG キャラクタのための反射的注意動作の生成”**,  
FIT2005, LJ-005.
- [9] 中村 正宏, 星野 准一:  
**“協調動作モデルを利用した CG キャラクタのための表情表現手法”**,  
FIT2005, LJ-004.
- [10] 苧阪 良二, 中溝 幸夫, 古賀 一男:  
**“眼球運動の実験心理学”**,  
p.20-21, 名古屋大学出版会 (1993).
- [11] 独立行政法人 産業技術総合研究所 デジタルヒューマン研究センター:  
**“人体寸法データベース 1991-92”**,  
<http://www.dh.aist.go.jp/AIST91DB/91-92/>
- [12] C.M.Harris, D.M.Wolpert:  
**“Signal-dependent noise determines motor planning”**,  
Nature 394,p.780-784(1998).
- [13] 張 曉林, 若松 秀俊:  
**“両眼眼球運動制御メカニズムの数学モデルと視軸制御システムの構築”**,  
日本ロボット学会誌,20,1,pp.89-97(2002).
- [14] R.H.S.Carpenter:  
**“VISION AND VISUAL DYSFUNCTION VOLUME 8 - Eye Movements”**,  
p.107-110, CRC Press(1991).