

注視点の変化に追従するゲームキャラクタの 頭部および眼球運動の自動合成

鍵山 裕貴¹ 川井 正英¹ 桑原 大樹¹ 森島 繁生²

概要: 眼球運動とそれに伴う頭部運動は、ゲームキャラクタの動きをよりリアルにみせるために重要な要素である。しかし、現状では眼球運動はアーティストの手作業によって作成されており、その表現には多大な労力と時間がかかっている。そこで本研究では、注視点の位置と表示時間を考慮したキャラクタアニメーションを自動合成する手法を提案する。具体的には、注視点が移る際の頭部及び眼球運動を実測し、その実測データから運動を決定づける複数のパラメータを抽出する。このパラメータをモデル化し、注視点の位置と表示時間に応じて決定することで、キャラクタの詳細な頭部及び眼球運動の合成が可能となる。

1. はじめに

近年、コンピュータ・ゲーム機の演算・描画機能向上に伴い、CGシーンをリアルに再現することが可能となってきた。CGをリアルに再現する要素としては、3つの要素が挙げられる。第1にキャラクタやオブジェクトの形状を再現するモデリング、第2にキャラクタやオブジェクトの動作を制御するアニメーション、第3に光の反射や屈折を計算し、物体の質感を写実的に表現するレンダリングである。3つの要素はどれもCGにおいて欠かせない要素であり、これまでに多くの研究が行われてきた。本研究では、上に挙げた3つの要素のうち、アニメーションについて着目した。

キャラクタアニメーションを作成するための研究は数多く存在する。例えば、Kovarら[1]はモーションキャプチャシステムを用い人間の様々な動作を取得し動作合成を行うことで、キャラクタの身体動作を作成した。しかしながら、この研究は身体のおおよその動きを再現することは可能だが、顔のような詳細な表現が必要とされる動きには対応できない。

次に、顔の動き（フェイシャルアニメーション）をリアルに表現する研究も多く行われている。顔の動きは非常に微細な変化を含むため、身体の動きと区別して動作を制御することが多い。フェイシャルアニメーションの代表的な手法としては、「モーションキャプチャシステムを使って実

際の人間の動きを取得し、それをCGモデルに適用する手法」[2], [3], 「アーティストが作成したキーシェイプの線形結合によって、キャラクタの表情を再現するブレンドシェイプと呼ばれる手法」[4], [5]などが挙げられる。これらの手法を用いることで顔の動きを正確に表現することができるようになってきているが、顔の一つ一つのパーツに着目すると、再現性には未だに問題が残っている。特に目に関しては、モデリングすらされていない場合が多い。

顔のパーツの中で、人に与える印象に強く影響するものの一つとして「目」が挙げられる。眼球運動の研究は生理学等の分野[6], [7], [8]で広く行われており、近年では、CG分野[9], [10]においても少しずつ研究が進んできている。しかしながら、眼球運動は周りの環境などによって多様なバリエーションがあるため、その再現は非常に難しい。したがって、実際のCGシーンにおいては、眼球運動はアーティストの手作業によってキーフレーム毎に定義されており、その表現には多大な労力と時間がかかっている。

そこで本研究ではモーションキャプチャと視線測定装置を同時に用いて、注視点が変化した際の、頭部及び眼球運動の関連性を分析し、注視点の位置と表示時間を入力としてキャラクタに合成する手法を提案する。注視点が移り変わるシーンはCGコンテンツに多く登場するシーンであり、このシーンを分析し、頭部及び眼球の動きのモデルを構築することで、CGコンテンツ内の様々なシーンに適用可能である。

2. 関連研究

2.1 眼球運動をCG再現する研究

眼球運動を再現する代表的な手法としては、実測結果に

¹ 早稲田大学

Waseda University

² 早稲田大学理工学術院理工学総合研究所

Waseda Research Institute for Science and Engineering

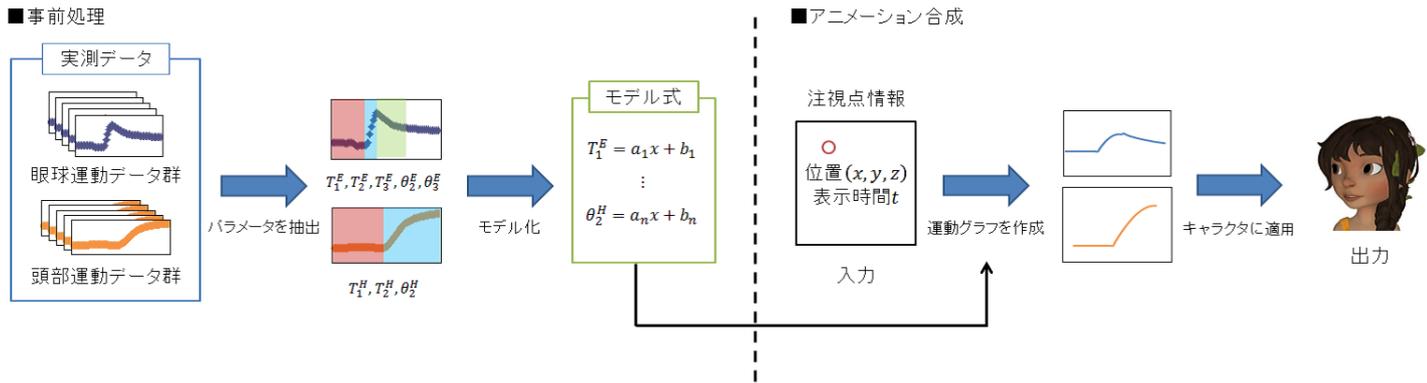


図 1 本研究の概要

に基づき、眼球運動のモデルを構築する手法 [11], [12], 生理学, 心理学等の知見を利用し、眼球運動のモデルを構築する手法 [13], [14] が挙げられる。しかし、いずれの手法にも問題がある。まず、実測結果に基づき、モデルを構築する手法においては、計測に視線測定装置を用いることが多いが、視線測定のみでは眼球運動と頭部の動きとの関連性を考慮できないことに問題がある。眼球は頭部と相互に作用しながら運動をしているため、視線動作のみを計測し、モデルを構築すると、眼球と頭部の動作の同調がうまくいかない。また、身体と眼球の動きを測定し、関連性を分析した研究 [15], [16] もあるが、分析結果を適用できるシーンが測定した環境のみであり、合成シーンが限定的であることが大きな問題である。このように頭部と眼球運動の関連性を分析し、合成する研究はこれまでにほとんど行われていない。次に、従来の生理学, 心理学の知見を利用し、モデルを構築する手法であるが、この手法では、定性的に説明された眼球運動の性質のみを使って表現するため、定量的なモデル構築ができないことが問題である。

2.2 頭部及び眼球運動の関連性を考慮した研究

頭部及び眼球運動を CG 再現する研究として、Iwao ら [15] はモーションキャプチャと視線測定装置を同時に用いて、足下環境が変化した際の、歩行時の身体動作と頭部及び眼球運動の関連性を分析し、キャラクターに合成した。

また、Yeo ら [16] も同様に、モーションキャプチャと視線測定装置を用いて、ボールを捕る際の人間の眼球運動及び身体動作について測定し、その結果を分析し、合成する手法を提案した。彼らの手法では、頭部の動きだけではなく、手の動きとの関連性にも着目して、リアルな身体動作及び眼球運動を合成することができていた。しかしながら、これらの研究は歩行や捕球といった限定的なシーンにしか適用できないといった問題がある。

汎用的な頭部及び眼球運動を CG 再現する研究として、Itti ら [17] は、カメラ 3 台と視線測定装置を用いて、テレビ画面に映った動画を見る際の頭部及び眼球運動を測定し、

この動きをモデル化した。注視点を表示された動画内の顕著性が最も高い部分とすることで、様々な状況下における頭部及び眼球動作を再現することが可能であった。しかしながら、実際の人間の動きは注視点が同じであったとしてもその時その時で動作が変化するはずであるが、この手法では一意に動き方が決定されるといった問題点がある。また、前庭動眼反射と呼ばれる、頭が動いたときにこれと反対方向に眼球を動かして、視界内のぶれを防ぐための反射を再現できていない。

2.3 本手法の新規性

関連研究 (2.1, 2.2 章) の問題点を踏まえて、本手法の新規性について述べる。まず、Iwao らと Yeo らの手法では、歩行動作や捕球動作など限定的なシーンでしか適用ができていなかった。そこで本手法では、眼球及び頭部運動の根幹である注視点が動く動作について、実測を行うことによりこれを改善する。次に Itti らの手法では成しえなかった、動きのパリエーションに関しては、実験で注視点の表示時間を変更することにより、また前庭動眼反射に関しては、データ解析の際に前庭動眼反射まで解析範囲を広げることにより、改善する。新規性をまとめると次の 3 つのようになる。

- (1) 適用可能なシーンが豊富
- (2) 注視点の表示時間を考慮することで、多様な動作が再現可能
- (3) 前庭動眼反射を含めたモデル式を作成することで、動きがリアルになる

この 3 点を達成することで、汎用的かつ詳細にキャラクターアニメーションを作成することが可能となる。

3. 頭部及び眼球運動の実測

3.1 実験環境と測定装置

本研究では被験者である 20 代の男女 9 名を対象に、注視点に移る際の頭部及び眼球運動を測定する (図 2)。頭部運動の測定装置として Vicon 社製の光学式モーションキャ

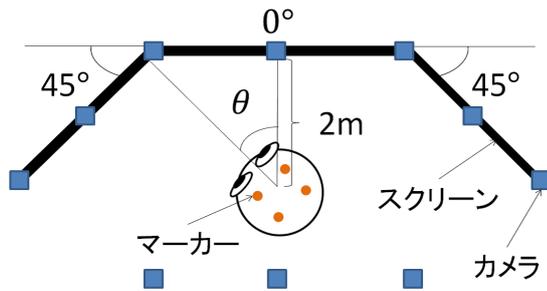


図 2 実験の模式図

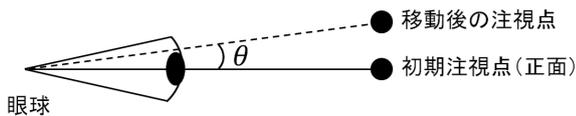


図 3 眼球の変異角の定義

プチャシステムを用いた。また、測定の際には 10 台のカメラと 4 個のマーカーを用いた。マーカーは頭部の右前、左前、右後、左後に張り付けた。モーションキャプチャ装置のサンプリングレートは 120Hz である。

視線測定については頭部動作と同時に視線を測定可能なヘッドマウント型の視線測定装置である Nac 社の EMR-9 を用いた。EMR-9 は、額と眼球付近にそれぞれ小型カメラを装着し、眼球が見ている景色と眼球のみの画像をそれぞれ撮影している。なお、本研究の額付近の小型カメラの仰角は 121° に設定した。また、小型カメラから測定される画像サイズは (640,480) である。視線測定装置のサンプリングレートは 60Hz である。

3.2 注視点に移る際の頭部及び眼球運動の実測

測定は図 2 に示すように、まず、被験者が正面を向いた状態で、スクリーンの X° の方向に T 秒間、注視点となる黄色い円を表示させる。次に被験者にこの黄色い円を注視してもらうことで、表示時間 T 秒のときの正面からの X° までの頭部の座標変化と、注視点の遷移を測定できる。この試行を $X=90^\circ$ から 90° まで 10° 刻みで、表示時間 $T=0.5$ 秒から 2 秒まで 0.5 秒刻みで測定し、これを 1 セットとする。これを 1 被験者あたり 10 セット測定する。表示させる時間を変更させることで、表示時間が異なるときに物体を確認する際の頭部及び眼球運動を測定することができる。

4. 頭部及び眼球運動の分析

4.1 眼球の変異角

本研究では、注視点変更の際の頭部と眼球の運動を分析する。そのため、注視点変更の際の回転する眼球と頭部の

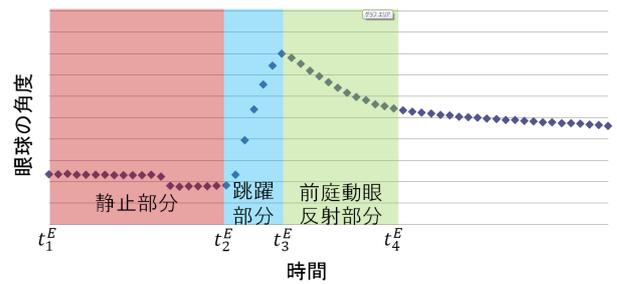


図 4 実際の眼球運動の様子

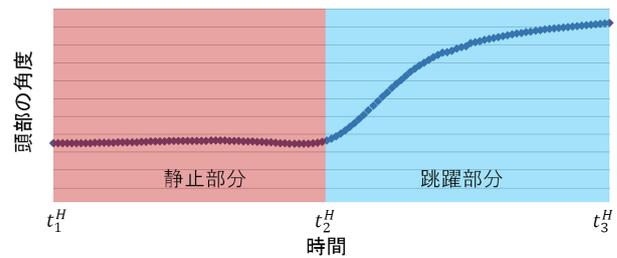


図 5 実際の頭部運動の様子

変位の大きさ (以下、変位角) をフレームごとに決定する必要がある。以下、眼球運動の場合について説明する。本研究では、眼球運動の変位角の大きさ θ を図 3 のように定義する。視線測定装置の小型カメラから得られる画像サイズ (w, h) が (640,480) であること、測定時の小型カメラの仰角 ϕ が 121° であることを考慮すると、初期注視点からのピクセルのずれ p を変位角 θ に変換する式は、式 (1) のように表せる。

$$\theta = \arctan\left(\frac{p \tan \phi / 2}{w / 2}\right) \quad (1)$$

ただし、瞬きしたフレームに関しては視線が測定できないため、前後のフレームから線形補完をするものとした。頭部の変位角の大きさについても、モーションキャプチャのマーカー座標を取得することにより、上記と同様に定義することができる。図 4 に実際の眼球運動の様子を、図 5 に実際の頭部運動の様子を示す。

4.2 頭部及び眼球運動のパラメータの抽出

次に、頭部及び眼球運動を決定づけるパラメータを実測データから抽出する手法について述べる。眼球運動に関しては図 4 に示すように注視点が表示されてから動き始めるまでの「静止部分」、注視点をとらえるまでの「跳躍部分」、頭の回転と逆方向に回転し視線方向を維持し続けるための「前庭動眼反射部分」の 3 つに分割する。注視点が表示された時刻を t_1^E とし、上記の 3 つに分割する時刻をそれぞれ t_2^E , t_3^E , t_4^E とする。具体的には前フレームとの差分が 0.3° 以上であるフレームが 3 フレーム以上続いた時の最初のフレームを t_2^E 、跳躍方向と逆方向に動いたフレームが 5 フレーム以上続いた時の最初のフレームを t_3^E 、前フレームとの差分が 0.1° 以下であるフレームが 2 フレーム以上

続いた時の最初のフレームを t_4^E と定義する。また、時刻 t_i^E のときの角度を $\theta(t_i^E)$ ($i = 1 \sim 3$) とする。これにより眼球運動の各区間に要する時間 T_i^E ($i = 1 \sim 3$) 及び角度変化 θ_i^E ($i = 2, 3$) が式 (2)(3) により得られる。静止部分の角度変化は 0 であると仮定した。

$$T_i^E = t_{i+1}^E - t_i^E \quad (2)$$

$$\theta_i^E = \theta(t_{i+1}^E) - \theta(t_i^E) \quad (3)$$

さらに、跳躍部分に関しては角度方向と時間方向を 0 から 1 に正規化し、Itti ら [17] が用いた以下の式でフィッティングを行う。

$$\theta = e^{at} \left(\sin \frac{\pi}{2} t \right)^b \quad (4)$$

ただし前庭動眼反射部分に関しては、(4) 式の正負を逆にしたものをフィッティングに用いる。 θ は角度、 t は時間、 a, b はフィッティングパラメータである。跳躍部分、前庭動眼反射部分のフィッティングパラメータをそれぞれ $a_2^E, b_2^E, a_3^E, b_3^E$ とする。

頭部運動に関しては図 5 に示すように眼球運動とは違い前庭動眼反射のような戻る動きが顕著に観測されないため、静止部分と跳躍部分の 2 つに分割する。注視点が表示された時刻を t_1^H 、前フレームとの差分が 0.1 以上であるフレームが 27 フレーム以上続いた時の最初のフレームを t_2^H 、跳躍方向と逆方向に動いたフレームを t_3^H と定義する。頭部運動の各区間に要する時間を T_i^H ($i = 1, 2$)、角度変化を θ_i^H ($i = 2$)、フィッティングパラメータを a_2^H, b_2^H とし、眼球運動と同様の解析を (2)~(4) 式を用いて行う。

4.3 頭部及び眼球運動のモデル構築

本章では、前章で分析した身体動作と頭部及び眼球運動の関係を CG シーンへ適用できるように、前章で分析された結果を、数式を用いてモデル化する。前章で抽出されたパラメータ $T_1^E, T_2^E, T_3^E, \theta_2^E, \theta_3^E, T_1^H, T_2^H, \theta_2^H$ を注視点角度ごとに、全被験者・全試行について平均をとった値をプロットしたグラフを図 6 に示す。眼球の静止時間 [a] に関しては、0.5 秒、1.0 秒の場合は注視角度による変化があまり見られないことがわかる。1.5 秒、2.0 秒の場合は注視角度が大きくなるにつれて静止時間も長くなっている。眼球の跳躍時間 [b] も眼球の静止時間 [a] と同様のことが言える。眼球の前庭動眼反射時間 [c] に関しては、前庭動眼反射時間は表示時間だけに依存しており、注視角度の影響はほとんどないことがわかる。眼球の跳躍角度 [d] については、表示時間別で見ると、0.5 秒の時が最も大きく 1.5 秒、2.0 秒の時が最も小さい。また 30° 付近までは単調増加しているが、その後 50° まで減少、70° まで増加、90° まで減少している。眼球の前庭動眼反射角度 [e] は、データにばらつきがあるがほとんどのデータが 15° ~ 25° の範囲に収まっている。次に、頭部の静止時間 [f] に関してであるが、

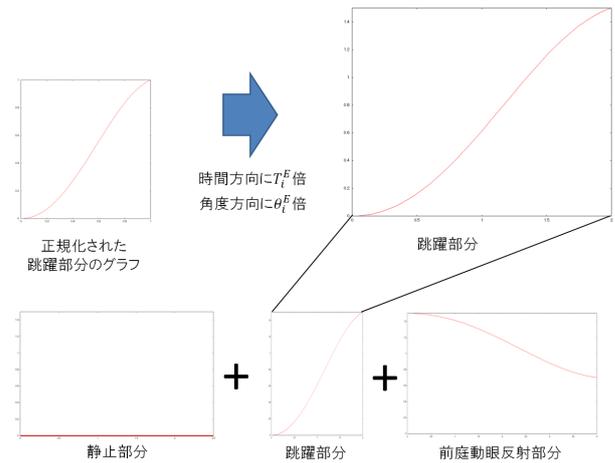
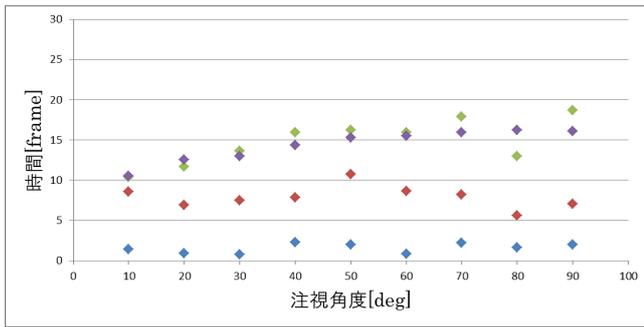


図 7 眼球運動の生成過程

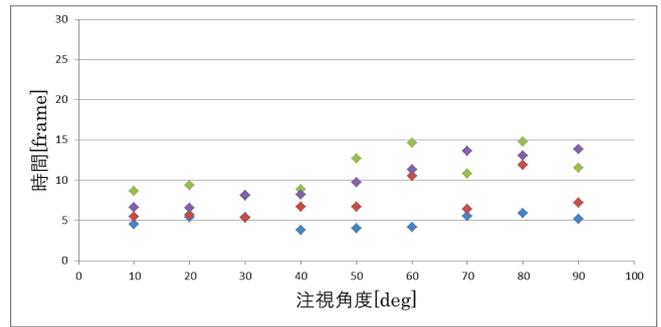
これは眼球の静止時間 [a] とほぼ同様のことがいえる。頭部の跳躍時間 [g] に関しては、注視点角度が小さいほど表示時間による差がなく、逆に注視点角度が大きいと、表示時間が長いほど跳躍時間も長いとわかる。頭部の跳躍角度 [h] に関しても同様のことが言える。つまり跳躍時間が増加すると跳躍角度も増加しているの、注視点の表示時間によらず頭の回転速度はほぼ一定であるといえる。各表示時間のデータを線形で近似する。その傾きと切片を表 1, 2 に示す。 $a_2^E, b_2^E, a_3^E, b_3^E, a_2^H, b_2^H$ は全データを平均したものを使用する。これを表 3 に示す。

5. 頭部及び眼球運動の合成

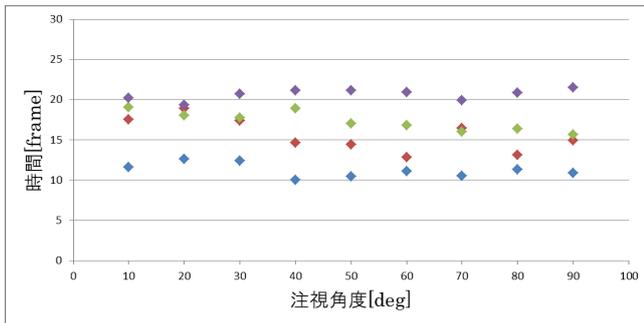
次に、注視点角度を入力として決定された各パラメータを使った頭部及び眼球運動の合成法を述べる。眼球運動に関して、まず初めに、 $a_2^E, b_2^E, a_3^E, b_3^E$ を用いて眼球の跳躍部分、前庭動眼反射部分の正規化されたグラフ (式 (4)) をそれぞれ作成する。次に、 T_1^E, T_2^E, T_3^E を用いて全ての区間に要する時間を決定し、図 7 のようにその時間にあわせて作成したグラフを伸縮させる。同様に、 θ_2^E, θ_3^E を用いて跳躍部分と前庭動眼反射部分で変化する角度を決定し、その角度にあわせて作成したグラフを伸縮させる。最後に、全てのグラフを結合することで、図 4 のような一連の眼球運動のグラフを決定する。頭部運動についても、推定されたパラメータ a_2^H, b_2^H を使って、跳躍部分のグラフ (式 (4)) を決定し、時間 T_1^H, T_2^H と角度 θ_2^H のパラメータを使って、図 5 のような一連の頭部運動のグラフを決定する。注視点を 55°、表示時間を 2.0 秒間としたときの従来手法で合成したアニメーションを図 8 に示す。また、提案手法を用いて頭部と眼球を合成したキャラクターアニメーションを図 9 に示す。提案手法では、まず最初に注視点に視線を合わせ、そのあとに頭部の動きを打ち消すような眼球の回転があるため、注視点が固定されていることがわかる。



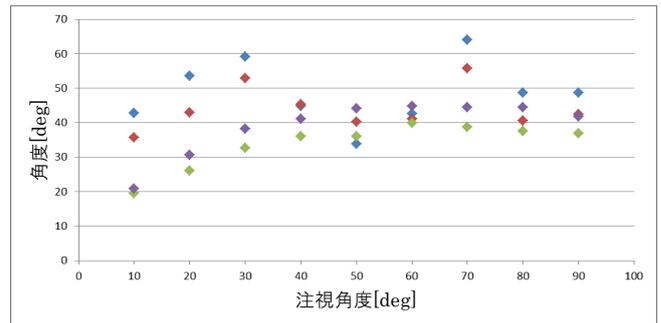
[a] 眼球の静止時間 (T_1^E)



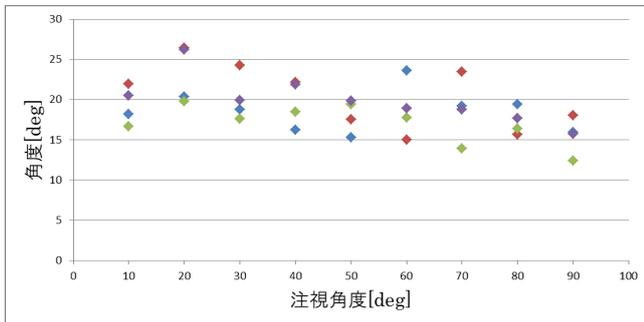
[b] 眼球の跳躍時間 (T_2^E)



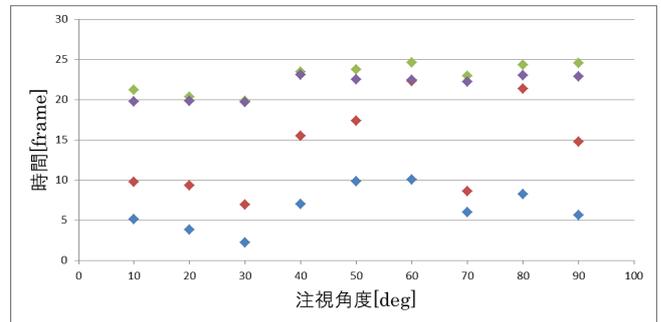
[c] 眼球の前庭動眼反射時間 (T_3^E)



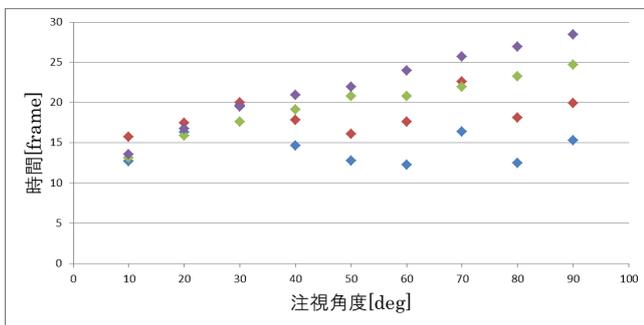
[d] 眼球の跳躍角度 (θ_2^E)



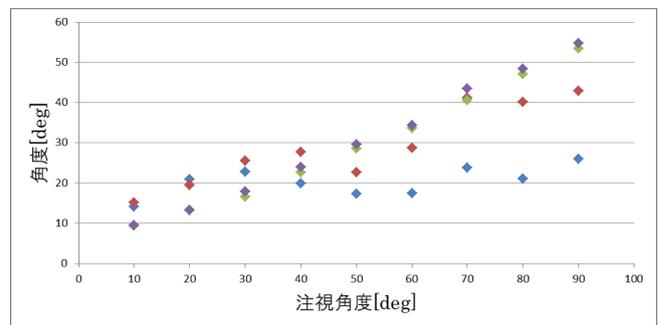
[e] 眼球の前庭動眼反射角度 (θ_3^E)



[f] 頭部の静止時間 (T_1^H)



[g] 頭部の跳躍時間 (T_2^H)



[h] 頭部の跳躍角度 (θ_2^H)

図 6 各パラメータと注視点角度の関係
 表示時間 0.5 秒：青，表示時間 1.0 秒：赤，表示時間 1.5 秒：緑，表示時間 2.0 秒：紫
 フレームレートは 30fps である。

表 1 眼球パラメータの線形近似の傾きと切片

表示時間 (s)	T_1^E		T_2^E		T_3^E		θ_2^E		θ_3^E	
	傾き	切片	傾き	切片	傾き	切片	傾き	切片	傾き	切片
0.5	0.01	1.04	0.01	4.48	-0.02	11.98	0.03	47.32	-0.01	18.84
1.0	-0.01	8.54	0.05	4.67	-0.05	18.20	0.03	42.36	-0.09	25.21
1.5	0.07	11.00	0.06	7.81	-0.04	19.31	0.20	23.73	-0.06	19.87
2.0	0.06	11.02	0.10	4.88	0.01	20.00	0.24	27.18	-0.08	24.12

表 2 頭部パラメータの線形近似の傾きと切片

表示時間 (s)	T_1^H		T_2^H		θ_2^H	
	傾き	切片	傾き	切片	傾き	切片
0.5	0.04	4.30	-0.02	15.54	0.08	16.40
1.0	0.11	8.49	0.04	16.43	0.34	12.18
1.5	0.05	20.06	0.13	13.13	0.56	1.42
2.0	0.04	19.50	0.18	13.18	0.58	1.56

表 3 フィッティングパラメータの平均値

a_2^E	b_2^E	a_3^E	b_3^E	a_2^H	b_2^H
1.21	0.07	0.01	0.39	-0.19	1.08



© 2009–2015 Mixamo, Inc.

図 8 従来手法



© 2009–2015 Mixamo, Inc.

図 9 提案手法

6. まとめ

本研究ではまず、注視点に移る際の、頭部動作と眼球運動を表示時間を変えて測定した。得られた実測データから頭部及び眼球運動を決定づける複数のパラメータを抽出し、注視点の位置と表示時間に応じたモデルを構築した。注視点の位置と表示時間を入力とすることで、リアルな頭部及び眼球の運動の自動合成が可能となる。本手法によって構築されたモデルは、従来手法では表現できていなかった前庭動眼反射の再現もでき、モデル式のパラメータ値を変更するだけで、実験で測定した環境以外の状況にも適用可能な、汎用的なモデルである。

7. 今後の課題

本研究で測定した動作は目が瞬間的に動く跳躍運動のみ

であり、物体を追いかける動きである追従運動に関しては考慮されていない。今後は、本研究で測定していない様々なバリエーションの動作を測定し、分析することにより、本手法の汎用性をより厳密に確かめることが必要である。また、本手法ではある動作中の注視対象が一つであるとして、動作合成を行っているが、実環境や CG シーンでは注視対象が複数ある場合も多いと考えられる。そのようなシーンの合成のために、今後は顕著性マップを算出する手法と本手法を組み合わせることにより、注視点を分散させつつ、リアルな頭部及び眼球運動を合成する手法について検討していきたい。

謝辞 本研究に 3D キャラクタ「Battery Girl」を提供していただいた Mixamo, Inc に深く感謝申し上げます。

参考文献

- [1] Baerlocher P., and Boulic R., "An inverse kinematics architecture enforcing an arbitrary number of strict priority levels", *The Visual Computer*, Springer Verlag, 20(6), pp.402-417, 2004
- [2] Li H., Yu J., Ye Y., and Bregler C., "Realtime facial animation with On-the-fly correctives", *ACM Transactions on Graphics (TOG) SIGGRAPH 2013 Conference*.
- [3] Cao C., Weng Y., Lin S., Zhou K., "3D shape regression for real-time facial animation", *ACM Transactions on Graphics (TOG) SIGGRAPH 2013 Conference Proceedings*, Volume 32, Issue 4, July, Article No. 41 (2013)
- [4] Bouaziz S., Wang Y., and Pauly M., "Online modeling for realtime facial animation". *ACM Transactions on Graphics (TOG) SIGGRAPH 2013 Conference Proceedings*, Volume 32 Issue 4, July, Article No.40 (2013)
- [5] Tena J. R., Torre F. D. la, Mathews I., "Interactive Region-Based Linear 3D Face Models", *ACM SIGGRAPH2011*, Vol.30, Article 76. (2011)
- [6] Becker, W., "The Neurobiology of Saccadic Eye Movements. Metrics", In: RH Wurtz and ME Goldberg(eds), Elsevier Science Publishers BV (Biomedical Division), New York, NY ch.2, pp13-67, 1989
- [7] Evinger, C., Manning, K. A., and Sibony P. A., "Eyelid Movements Mechanisms and Normal Data", *Investigate Ophthalmology and Visual Science*, Vol32, No.2, February 1991
- [8] Becker, W., "The Neurobiology of Saccadic Eye Movements. Metrics", In: RH Wurtz and ME Goldberg(eds), Elsevier Science Publishers BV (Biomedical Division), New York, NY ch.2, pp 13-67, 1989
- [9] Lance, B. J. and Marsella, S. C., "The Expressive Gaze Model: Using Gaze to Express Emotion", *Computer Graphics and Applications*, IEEE, Volume30 Issue4, pp62-73, 2010
- [10] Colburn, A., Cohen, M. F., and Drucker, S. M., "The role of eye gaze in avatar mediated conversational interfaces." Microsoft Tech Report 2000-81
- [11] Lee, S. P., Badler, J. B., and Badler, N. I., "Eyes Alive", *ACM Transactions on Graphics (TOG)- Proceedings of ACM SIGGRAPH*, Volume 21 Issue 3, pp637-644, July 2002
- [12] Gu, E., Lee, S. P., Badler, J. B., and Badler, N. I., "Eye Movements, Saccades, and Multiparty conversations", *Data-Driven 3D Facial animation*, pp79-97, December 11, 2007
- [13] Masuko, S., and Hoshino, J., "Head-eye animation corresponding to a conversation for CG characters", *Computer Graphics Forum, Journal of the European Association for Computer Graphics*, vol.26, no.3, pp.303-311, 2007.
- [14] ITTI, L., "Realistic Avatar Eye and Head Animation Using a Neurobiological Model of Visual Attention", *Tech.rep.*, DTI Document, 2003
- [15] Iwao, T., Kubo, H., Maejima, A., and Morishima, S., "A Visuomotor Coordination Model for OBstacle Recognition", *Journal of WSCG*, vol.22, , pp49-56, July 2014
- [16] Yeo, S. H., Lesmana, M., Neog, D. R., and Pai, D. K., "Eyecatch: Simulating Visuomotor Coordination for Object Interception", *ACM Transactions on Graphics*, vol.31 issue 4, no.42, July 2012
- [17] Itti, L., Dhaval, N., and Pighin, F., "Photorealistic Attention-Based Gaze Animation", *Multimedia and Expo, 2006 IEEE International Conference on*, pp521-524, July 2006