

顔への追従投影における遅延知覚と表情動作の関係のモデル化の検討

河口 将也^{*1}Peng Hao-Lun^{*1}西田 眞也^{*2}渡辺 義浩^{*1}

Abstract — 近年、顔への追従投影により、外観を変化させる Dynamic Facial Projection Mapping (DFPM) の研究が行われている。DFPM では、顔のセンシングから投影までの遅延を観測者が知覚する問題がある。そのため、システム設計時に投影遅延を観測者が知覚しない弁別閾以下に収める必要がある。しかし、投影対象の顔の動きと知覚される遅延の関係性の知見は、十分に得られていない。そこで、本稿では DFPM において観測者が知覚する遅延と、投影対象である顔の動きの関係を表すモデルを提案する。具体的には、顔の表情動作に対する動き分布の特徴量を入力として利用することで、遅延時間の弁別閾と投影映像の遅延を知覚する確率を予測するモデルを提案する。

Keywords : 視覚, 遅延知覚, 拡張・複合現実, プロジェクタ

1 はじめに

Dynamic Projection Mapping (DPM) は投影対象の位置や形状に合わせ、映像を投影することで物体の見た目を変化させることができる技術である。DPM では、投影対象のセンシングからプロジェクタによる投影の間に生じる、わずかな遅延を観測者が知覚し、没入感が低下することが問題となっている。そのため、DPM では、システム設計時に投影遅延を観測者が知覚しない範囲に収める必要がある。

本稿では、DPM の中でも、投影対象を人間の顔に限定した Dynamic Facial Projection Mapping (DFPM) において、観測者が知覚する遅延に注目する。DFPM は人間の顔に映像を投影することで、舞台における演出やメイクのシミュレーションなどに応用されている [1, 2]。DFPM においても DPM と同様に、顔の動きと投影映像の間に遅延が生じると、観測者は遅延を知覚する。特に、DFPM では、目、鼻、口といったパーツへの投影ずれを手掛かりに遅延が知覚される可能性が考えられる。

人間が知覚する遅延に関する研究として、Ng らは、平面のタッチインタフェース上で指に追従して動く正方形の遅延が、平均 6.04 ms 以下である場合にユーザから知覚されないことを報告した [3]。しかし、同研究は投影対象の形に重ね合わせるように映像を投影する DFPM とは、実験の条件が異なるものである。

そこで、Peng らは DFPM における遅延の知覚について、観測者が遅延に気づかない遅延時間の弁別閾を異なる顔の動きの条件下で調査した [4]。具体的には、投影対象である顔の並進、左右への回旋、話して

いる動きの 3 種類の違いや、動きの速度の違いによって、遅延時間の弁別閾が異なることを示した。さらに、DFPM における遅延時間の弁別閾が、先行研究で示されたタッチインタフェース上の弁別閾である平均 6.04 ms を下回り、最も厳しい条件で平均 3.87 ms であることも報告した。しかし、同研究では、具体的にどのような動きの特徴が遅延時間の弁別閾に差を生じさせているか調査されていない。また、DFPM において観測者が知覚する遅延と投影対象である顔の動きの関係に関する知見は未だ十分に得られていない。

そこで、本稿では、DFPM において観測者の知覚する遅延と、投影対象である顔の動きの関係を表すモデルを提案する。具体的には、遅延時間の弁別閾に相関の高い顔の動き分布の特徴量を入力として、動きに応じた弁別閾を予測するモデルを構築した。さらに、弁別閾の予測モデルを利用し、観測者が遅延を知覚する確率を予測するモデルを構築した。

2 関連研究

2.1 視覚に関するモデル

視覚刺激に対する人間の知覚反応をモデル化した事例は多数報告されている。中でも、視覚刺激の特徴量を入力として人間の知覚を予測するモデルの構築を行った事例として、Tursun らは、映像変化の検出についてのモデルを報告した [5]。同研究では、映像変化の検出に影響を与える主要因である時間周波数、空間周波数、偏心の 3 つの特徴量を変更した実験刺激を複数作成し、被験者実験により各実験刺激を検出する輝度コントラストの弁別閾を調査した。さらに、被験者実験で得られた弁別閾をもとに、観測者が映像変化を検出する確率を予測するモデルも報告された。

*1東京科学大学

*2京都大学

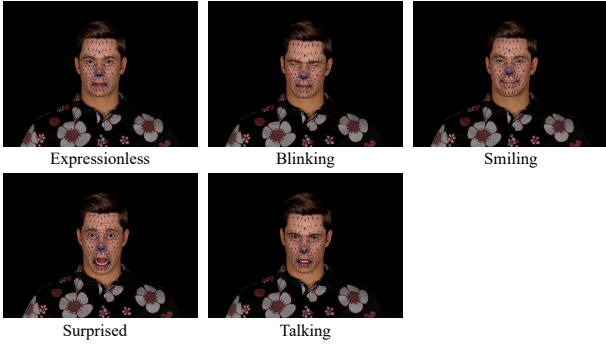


図1 先行研究で利用された5種類の実験刺激 [7]. 投影対象の顔に重なる青色のワイヤーフレームが DFPM での投影映像の効果を再現している。

Fig.1 Five types of experimental stimuli used in previous study [7]. The blue wire-frame overlapping the face represents the effect of the projected image in DFPM.

また, Assen らは, 様々な照明環境や物の材質の光学条件が変化した場合に, 人間が錯覚的に知覚する動きの速度について, 動きベクトルの分布であるオプティカルフローを分析した [6]. 同研究では, 速度の知覚に強いオプティカルフローの特徴量があることを確認し, 確認された特徴量を入力として, 観測者が知覚する速度を予測するモデルを構築した。

2.2 DFPM の遅延知覚と動きの関係の分析

先行研究では, DFPM の遅延の知覚と動きの関係性を分析した事例が報告されている。事例の1つとして, 河口らは被験者実験により複数の表情動作に対する遅延時間の弁別閾と, 弁別閾に差を生じさせる動きの特徴の調査を行った [7]. 同研究では, 最大速度の統制された複数の表情動作に対する遅延時間の弁別閾と, オプティカルフローの特徴量の相関係数を分析した。分析の結果, 動き方向の分散を表すオプティカルフローの特徴量と弁別閾の相関係数が 0.730 であることを確認した。しかし, 分析された特徴量をもとに観測者の知覚する遅延を予測する試みはされておらず, DFPM で知覚される遅延について, モデルの構築が行われた事例は未だ報告されていない。

3 遅延時間の弁別閾予測モデル

3.1 概要

本節では, 動きベクトルの分布であるオプティカルフローの特徴量を入力として, DFPM における投影遅延時間の弁別閾を予測するモデルを提案する。本モデルは, DFPM の投影対象である顔の動きから, 遅延時間の弁別閾を予測し, DFPM の遅延時間の設計要件を定めることができる。なお, 本モデルの設計は, 時間周波数, 空間周波数, 偏心をもとに人間の輝度コ

表1 15種類の実験刺激の最大速度 v_{\max} .

Table 1 The maximum speed v_{\max} of the 15 experimental stimuli.

| v_{\max} (m/s) | Slow | Normal | Fast |
|------------------|-------|--------|-------|
| Expressionless | 0.190 | 0.279 | 0.844 |
| Blinking | 0.173 | 0.277 | 0.873 |
| Smiling | 0.168 | 0.285 | 0.850 |
| Surprised | 0.160 | 0.280 | 0.872 |
| Talking | 0.176 | 0.297 | 0.844 |

ントラストの知覚のモデル化を提案した関連研究 [5] を参考にした。

3.2 被験者実験

弁別閾予測モデルの設計のため, 複数の表情動作に対する遅延時間の弁別閾を被験者実験により調査する。弁別閾の調査のため, 繰り返しペアとして提示される実験刺激に対して, 遅延が知覚される実験刺激を選択する Two-Interval-Forced-Choice (2IFC) の方法を利用した [8]. 提示されるペアの映像は, 一方が遅延のない映像で, もう一方が遅延のある映像である。被験者は遅延がないと感じる映像を選択するように求められた。なお, 投影の遅延時間は, ステアケース法に基づき決定された [9].

図1に本実験で利用した実験刺激を示す。実験刺激は, 先行研究 [7] で利用された実験刺激と同様の表情動作の実験刺激を利用した。5種類の表情動作に対して, 最大速度が異なる3つの実験刺激を作成し, 合計で15種類の実験刺激を用意した。各実験刺激の最大速度は v_{\max} として定義され, 全フレームを通してのオプティカルフローのノルムの最大値を速度換算することで計算された。各実験刺激の最大速度 v_{\max} を表1に示す。実験刺激の最大速度 v_{\max} は, Slow, Normal, Fast として表情動作の異なる刺激間で, それぞれ同程度の速度帯に設定された。

実験刺激は, 先行研究 [7] と同様の方法で作成されたシミュレーション映像を用いた。シミュレーション映像は, ゲームエンジンである Unreal Engine 5.3¹と, リアルな人間の外観をもつデジタルヒューマンである Meta Human²により作成された。

図2に被験者実験の環境を示す。実験には3名の被験者が参加した³。実験実施中, 被験者は環境光のない暗室で白色のスクリーンに投影される実験刺激を, 1.5 m 離れた位置から観察した。実験刺激内の顔の大きさは, 視野角換算で縦 7.628 deg, 横 5.724 deg であり, 実験刺激の明るさは, 104.4 lx であった。2IFC の方法での選択タスクを行うため, 被験者はキーボードを持った状態で実験を行った。実験刺激として, 24 ビツ

¹<https://www.unrealengine.com/ja>

²<https://www.unrealengine.com/ja/metahuman>

³本実験は, 東京工業大学研究推進部による人を対象とする研究倫理審査の承認 (承認番号:2024086) を得たうえで実施された。

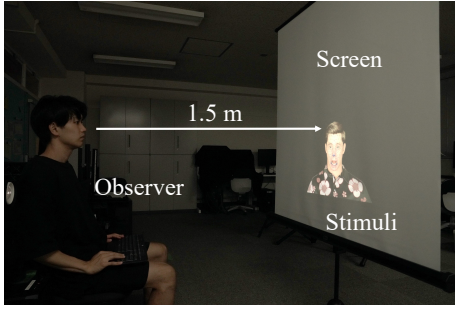


図2 被験者実験の環境.

Fig. 2 Experimental setup for user study.

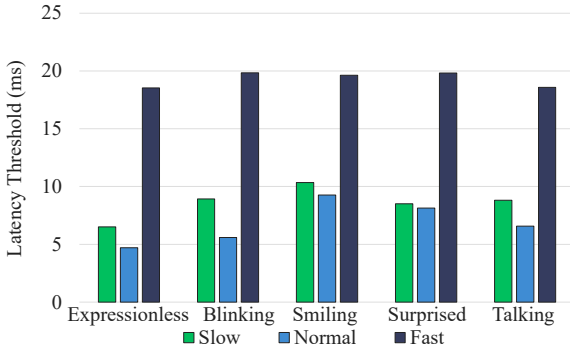


図3 被験者実験の結果. 各速度帯における実験刺激の弁別閾の平均値を表情動作ごとに示す.

Fig. 3 The results of the user study. The average latency thresholds for each speed range are shown for each facial expression.

トのRGB画像を1,000 fpsで投影可能な高速プロジェクタで投影した[10]. 画像の解像度は 1024×672 であった. 実験では, Core i9-12900K, RTX 3090, 64 GBのメモリを搭載した計算機を利用した.

図3に被験者実験によって得られた各実験刺激の遅延時間の弁別閾の平均値を示す. グラフの縦軸は遅延時間の弁別閾を示し, 各速度帯の実験刺激の弁別閾が表情動作ごとに示されている.

3.3 モデルの設計方法

遅延時間の弁別閾予測モデルの設計のために, 被験者実験により得られた各実験刺激の弁別閾と, オプティカルフローの特徴量を利用する. オプティカルフローの特徴量は, 前述した動作の最大速度を表す v_{\max} と, 先行研究[7]で弁別閾に対して高い相関を示すことが報告された特徴量 d_v を入力とする. 特徴量 d_v は, 前後2フレーム間で計算し, 全フレームを通して平均したオプティカルフローの方向成分の分散である. なお, 特徴量 d_v は0から1の範囲で正規化された.

遅延時間の弁別閾予測モデルの設計では, まず, 動作の最大速度 v_{\max} と遅延時間の弁別閾に3次の多項式回帰を適用する. この処理は, 顔の動作速度が遅延時間の弁別閾に大きく影響を与えることが, 先行研究[4]により報告されているため, 最大速度 v_{\max} に対す

る基本となる弁別閾の変化を表現するために施される. 3次の多項式回帰式 S_p は式(1)で示され, パラメータ $\beta_0, \beta_1, \beta_2, \beta_3$ が最小2乗法によって推定される.

$$S_p(v_{\max}) = \beta_0 + \beta_1 v_{\max} + \beta_2 v_{\max}^2 + \beta_3 v_{\max}^3 \quad (1)$$

次に, 3次多項式により得られた曲線 S_p に対して, 式(2)に示すソフトプラス関数を適用した曲線 S_{SP} を得る. この処理は, 3次多項式により得られた曲線 S_p が, 負の値をとらないようにするため, 施される処理である.

$$S_{SP}(v_{\max}) = \ln [1 + \exp (S_p(v_{\max}))] \quad (2)$$

次に, もう一方の特徴量 d_v が弁別閾に与える影響を, 曲線 S_{SP} の形状の調整により反映させるため, 2つの関数を設定する. 具体的には, 式(3)に示すスケーリング関数 T と, 式(4)に示すシフト関数 U を利用する. スケーリング関数 T は式(2)で得られた曲線 S_{SP} のスケールを調整する役割を持つ関数であり, シフト関数 U は, 式(2)で得られた曲線 S_{SP} の横軸方向の位置の調整を行うために設定される関数である. b_1, b_2, b_3, b_4 はスケーリング関数, シフト関数でそれぞれ調整されるパラメータである.

$$T(d_v) = b_1 + b_2 d_v \quad (3)$$

$$U(v_{\max}, d_v) = v_{\max} + b_3 + b_4 d_v \quad (4)$$

最終的な遅延時間の弁別閾予測モデル S を式(5)に示す. スケーリング関数 T は, ソフトプラス関数を適用した多項式回帰式である S_{SP} に乗算されることで曲線をスケーリングする. シフト関数 U は S_{SP} の入力として与えられる.

$$S(v_{\max}, d_v) = T(d_v) S_{SP}(U(v_{\max}, d_v)) \quad (5)$$

以上のような手順で, 最終的な遅延時間の弁別閾の予測モデルを設計する. なお, 本モデルの構築にあたり, 前述した被験者実験により得られた3名の遅延時間の弁別閾データを利用した.

3.4 結果

図4にオプティカルフローの特徴量 v_{\max} と遅延時間の弁別閾に3次の多項式回帰によるフィッティングを行った結果, 及びソフトプラス関数を適用した結果を示す. また, 表2にモデルの設計で推定されたパラメータを示す.

図5に最終的な弁別閾予測モデル S の可視化を示す. 緑色のプロットは, モデルの設計に利用された被

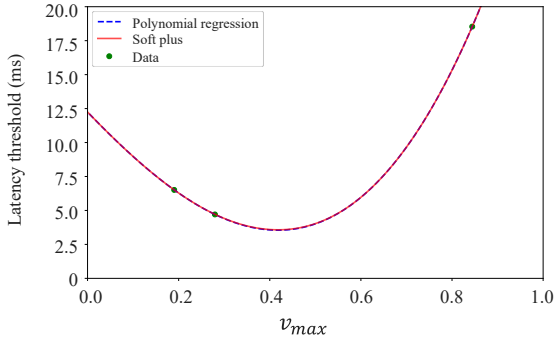


図4 多項式回帰でのフィッティングとソフトプラス関数適用後の結果.

Fig. 4 Results of polynomial regression fitting and after applying the softplus function.

表2 遅延時間の弁別閾予測モデルで使用した推定パラメータの値.

Table 2 Values of estimated parameters used in the latency threshold prediction model.

| Term | Coefficient |
|-----------|-------------|
| β_0 | 12.2450 |
| β_1 | -34.8702 |
| β_2 | 17.4783 |
| β_3 | 38.6558 |
| b_1 | 1.0937 |
| b_2 | 0.1369 |
| b_3 | -0.0241 |
| b_4 | -0.0316 |

験者のデータである。また、赤色のプロットはモデルにより予測された各実験刺激ごとの弁別閾を示す。

モデルの評価のために、モデルがどれだけ目的変数を説明しているかを表す決定係数 R^2 を用いた。分析の結果、決定係数 R^2 の値は 0.6903 となった。予測された弁別閾と実際の弁別閾の平均絶対誤差は、2.802 ms となった。

4 投影遅延を知覚する確率予測モデル

4.1 概要

本節では、DFPM システムの投影遅延時間に対して、観測者が遅延を知覚する確率を予測するモデルを提案する。本モデルは、前節の遅延時間の弁別閾予測モデルから得られた弁別閾を利用し、どれほどの確率で観測者が遅延を知覚するか予測するモデルである。前節での遅延時間の弁別閾予測モデルを拡張し、遅延を知覚する確率を予測することで、DFPM システムの遅延時間を、遅延を知覚する確率をもとに評価することが可能になる。本モデルも前節と同様に関連研究 [5] を参考にモデルの設計を行った。

4.2 被験者実験

確率予測モデルの設計にあたり、投影遅延時間が異なる複数の実験刺激を用いて、遅延のある映像を判別

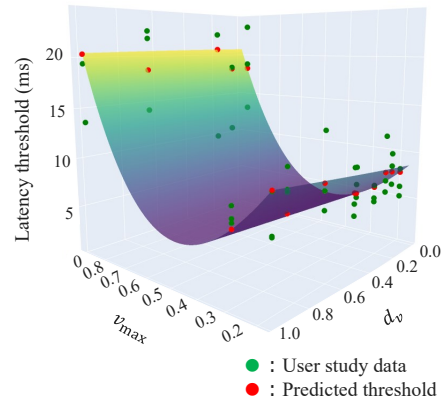


図5 遅延時間の弁別閾予測モデル S の可視化。緑色のプロットは、モデルの設計に利用された被験者のデータ。赤色のプロットはモデルにより予測された各実験刺激ごとの弁別閾。

Fig. 5 Visualization of the latency threshold prediction model S . The green plots represent the user study data used for model construction. The red plots show the predicted latency thresholds for each experimental stimulus.

できる確率を被験者実験により調査する。被験者実験に用いる実験刺激の作成のため、遅延レベル L を式 (6) のように定義する。遅延レベル L は、投影映像の遅延時間を予測された弁別閾で割った比で定義され、弁別閾に対する相対的な遅延時間のレベルを表す。

$$L = \frac{\text{Latency time (ms)}}{\text{Threshold (ms)}} \quad (6)$$

被験者実験では、表 1 に示す Expressionless, Blinking, Smiling, Surprised, Talking の 5 種類の表情動作のなかで Normal の速度帯の実験刺激を利用する。5 種類の表情動作それぞれに対して、遅延レベル L を 0.25, 0.5, 1.0, 2.0, 4.0 で設定した異なる遅延時間の実験刺激を用意した。実験刺激は、合計で 25 種類である。

被験者実験では、前節と同様に 2IFC の方法を利用した [8]。本実験では、表情動作ごとに 5 つの遅延レベルで設定された実験刺激が合計で 10 回提示されるように設定した。被験者は遅延のない映像と遅延のある映像を比較し、遅延のない映像を選択するように求められた。以上の作業により、遅延レベルごとに投影映像の遅延を判別できる確率を 2IFC タスクの正答率として求めた。なお、本被験者実験には 3 名が被験者として参加した。

4.3 モデルの設計方法

被験者実験により得られた各遅延レベルごとの 2IFC タスクの正答率を、心理測定関数にフィッティングする。フィッティングのため、2IFC タスクの正答率であ

る条件付確率 $P(T | x)$ を式 (7) のように定義する.

$$P(T | x) = 0.5 + 0.5 \cdot P(\text{detect} | x) \quad (7)$$

ここで, T は 2IFC タスクにおける正答を示し, x は, 遅延レベルを示す. $P(T | x)$ 内の $P(\text{detect} | x)$ を最終的な遅延の知覚確率を表す関数として, 式 (8) に示すロジスティック関数を用いて定義する [11]. なお, α , β は推定されるパラメータであり, 最小 2 乗法により推定される.

$$P(\text{detect} | x) = 1 + e^{-\frac{x-\alpha}{\beta}} \quad (8)$$

4.4 結果

図 6 に各遅延レベルごとの 2IFC タスクの正答率とフィッティングにより得られた心理測定関数の結果を示す. 図 6 の (a) は被験者実験により得られた遅延時間レベルごとの 2IFC タスクの正答率のプロットと, 式 (7) の $P(T | x)$ に示す 2IFC タスクの正答率のフィッティング結果を示す. 図 6 の (b) は遅延の知覚確率として得られた式 (8) のロジスティック関数 $P(\text{detect} | x)$ のフィッティング結果を示す. なお, (a), (b) のグラフの横軸は $\log_{10} L$ で表記された遅延レベルを示す. また, 本モデルで推定された式 (8) のパラメータは, $\alpha = 0.01$, $\beta = 10.877$ であった.

確率予測モデルの設計に用いられた被験者データとの決定係数 R^2 と相関係数 R の値は, $R^2=0.903$, $R=0.950$ であった. 本確率予測モデルは, 3 名の被験者データに対して高いフィッティングを示した.

5 考察

遅延時間の弁別閾予測モデルでは, 決定係数 R^2 が 0.6903 であり, 遅延時間の予測値と実際の弁別閾の平均絶対誤差が 2.8020 ms であった. 遅延の知覚に関する弁別閾のモデルの先行事例は未だ報告されていないが, 輝度コントラストに関して弁別閾を予測する関連研究 [5] のモデルの決定係数は 0.7 以上の高い水準を報告している. このモデルと比較すると, 本稿で述べた遅延時間の弁別閾予測モデルの決定係数は, 高い水準ではないと考える. そのため, 本モデルには改善の余地があると考えられる.

弁別閾予測モデルの精度の向上に関して, シフト関数, スケーリング関数の式の係数やパラメータを最適化することや, 空間周波数等の新たな特徴を入力として検討することが改善案として挙げられる.

投影遅延を判別する確率の予測モデルでは, 3 名での被験者実験の結果に対して, 決定係数 $R^2 = 0.903$, 相関係数 $R = 0.950$ となり, 高い精度を示した. このことから, 遅延時間の弁別閾に対するシステムの遅延

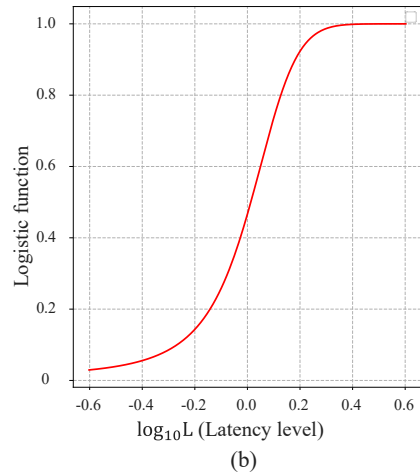
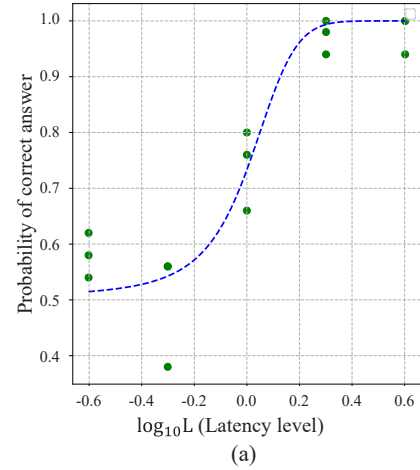


図 6 遅延レベルに対する遅延の知覚確率のプロットとフィッティング結果. (a) は遅延時間レベルごとの 2IFC タスクの正答率のプロットと式 (7) によるフィッティング結果. (b) は遅延知覚確率として得られた式 (8) のロジスティック関数 $P(\text{detect} | x)$.

Fig. 6 Plots and fitting results of the latency perception probability for different latency levels. (a) shows the plot of the probability of correct answer in 2IFC task and the fitting result using equation (7). (b) presents the logistic function $P(\text{detect} | x)$ shown by equation (8).

レベルが正確に予測されることで, 遅延の知覚確率を高い精度で予測可能であると考えられる.

6 まとめ

本稿では, DFPM における遅延時間の弁別閾を予測するモデルを構築した. さらに, 遅延時間の弁別閾を予測するモデルを拡張し, 観測者が投影遅延を知覚する確率を予測モデルを構築した. 提案した遅延時間の弁別閾予測モデルは, DFPM の投影対象である顔の動きから, 遅延時間の弁別閾を予測することができるため, DFPM の設計時の基準を定める際に利用さ

れることが期待される。また、遅延を知覚する確率予測モデルを利用することで、システムの遅延時間を、遅延を知覚する確率をもとに評価することや、遅延を知覚する確率を最小化するように投影対象の動きの最適化を検討することが可能となる。

参考文献

- [1] Amit H. Bermano, Markus Billeter, Daisuke Iwai, and Anselm Grundhöfer. Makeup Lamps: Live Augmentation of Human Faces via Projection. *Computer Graphics Forum*, Vol. 36, No. 2, pp. 311–323, 2017.
- [2] Nao Tsurumi, Kaoru Ohishi, Ryo Kakimoto, Fumihiko Tsukiyama, Hao-Lun Peng, Yoshihiro Watanabe, and Yuji Masubuchi. Rediscovering your own beauty through a highly realistic 3D digital makeup system based on projection mapping technology. In *International Federation of Societies of Cosmetic Chemists*, 2023.
- [3] Albert Ng, Julian Lepinski, Daniel Wigdor, Steven Sanders, and Paul Dietz. Designing for low-latency direct-touch input. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, pp. 453–464, 2012.
- [4] Hao-Lun Peng, Shin'ya Nishida, and Yoshihiro Watanabe. Studying User Perceptible Misalignment in Simulated Dynamic Facial Projection Mapping. In *2023 IEEE International Symposium on Mixed and Augmented Reality*, pp. 493–502, 2023.
- [5] Cara Tursun and Piotr Didyk. Perceptual visibility model for temporal contrast changes in periphery. *ACM Trans. Graph.*, Vol. 42, No. 2, 2022.
- [6] Jan Jaap R. van Assen, Mitchell J.P. van Zuijlen, and Shin'ya Nishida. Perceived object motion variance across optical contexts. *Computational and Mathematical Models in Vision*, 2022.
- [7] 河口将也, 彭浩倫, 西田真也, 渡辺義浩. 顔への追従投影における遅延知覚と表情変化の関係の分析. 第29回バーチャルリアリティ学会大会, pp. 2E2–09, 2024.
- [8] Rolf Ulrich and Jeff Miller. Threshold estimation in two-alternative forced-choice (2AFC) tasks: The Spearman-Kärber method. *Perception & Psychophysics*, Vol. 66, pp. 517–533, 2004.
- [9] Kaernbach Christian. Simple adaptive testing with the weighted up-down method. *Perception Psychophysics*, Vol. 49, No. 3, pp. 227–229, 1991.
- [10] Yoshihiro Watanabe and Masatoshi Ishikawa. High-speed and high-brightness color single-chip DLP projector using high-power led-based light sources. In *The International Display Workshops*, 2019.
- [11] Maurice Vogels, Rita Zoeckler, Donald M. Stasiw, and Lawrence C. Cerny. Notice sur la loi que la population suit dans son accroissement. *Journal of Biological Physics*, Vol. 3, pp. 183–192, 1975.