



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

河口将也¹⁾, Peng Hao-Lun¹⁾, 西田眞也²⁾, 渡辺義浩¹⁾

1) 東京工業大学 工学院 情報通信系 (〒 226-8501 神奈川県横浜市緑区長津田町 4259-G2-31, kawaguchi.s.ae@m.titech.ac.jp)

2) 京都大学 大学院情報学研究所 知能情報学専攻 (〒 606-8501 京都府京都市左京区吉田本町)

概要: 動的な映像投影により, 人間の顔の外観を変化させる応用において, 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. 本稿では, 観測者が知覚する遅延に差が生じる動きの特性を調査する. 具体的には, 表情変化を伴わない動きや, 表情変化を伴う動きに対して, 遅延時間の弁別閾を比較調査する. また, 投影対象である顔の動き分布の特徴量と弁別閾の関係を分析する.

キーワード: 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) は動的な投影対象の位置や形状に合わせて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, わずかな遅延が生じる. 観測者が投影映像の遅延を知覚すると, 投影対象の見た目に違和感を持ち, 没入感が低下することが問題となっている.

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.

また, DFPM とは別に, 顔の知覚に関する研究が行われている. 従来より, 人間の顔とそのほかの物を見る場合では, 異なる脳の認知機構が用いられており, 顔の知覚に特異性があることが知られている [3]. 例えば, 目, 鼻, 口のパーツだけでなく, それらの配置の全体構成として顔を知覚することが示されている [4]. また, 顔全体の中にあるパーツは, 単独で提示された同じパーツよりも認識しやすいことがわかっている [5]. このような特異性により, DFPM での遅延の知覚においても, 投影対象が顔であること特有の知覚性質をもつ可能性がある.

他にも, 視覚科学分野の研究では, 様々な物が動く複雑なシーンを観察する際に, 人間が知覚する動きの方向や速さが, 実際の動きと異なる場合があることが知られている [6]. また, 照明や物の材質によってシーンの光学条件が変化した場合に, 同じ動きをする物であっても知覚される動きの速さが異なることが知られている [7]. そのため, DFPM の遅延の知覚においても, 複雑に動く顔と, 顔の動きに追従する投影映像の間の方向や速度の違いを正確に認識して

いる可能性は低い. 例えば, DFPM において遅延を知覚する手掛かりの 1 つに, 顔の動きと投影映像の動きの間の相対的な速度の違いや, 動きの方向の違いを知覚することが考えられる. しかし, 上記のような錯覚が生じることを考えると, 必ずしも, 顔の動きと投影映像の動きの間の相対的な速度の違いや, 動きの方向の違いが大きいほど, 遅延が知覚され易いわけではないことが予想される.

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示した [8]. 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 [9]. 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. また, Peng らの結果は, 投影対象の動きの速度が速くなり, 顔の動きと追従する投影映像の動きの違いが大きくなるほど, 遅延が知覚され易いわけではないことを示し, 前述した仮説を支持した. さらに, DFPM における遅延時間の弁別閾が, 従来研究 [8] で示されたタッチインタフェース上の遅延時間の弁別閾である平均 6.04 ms を下回り, 平均 3.87 ms であることも確認された.

しかし, 上記で述べた 3 種類の動きの弁別閾の調査において, 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るか調査されていない. さらに, 動きが複雑になるほど遅延を知覚しにくくなる可能性が示唆されたが, 表情変化と頭部の動きを同時に伴うような, 顔で一般的に生じうる複雑な動きの遅延時間の弁別閾について, 調査が行われていない.

そこで, 本稿では, DFPM の投影対象である顔の動きに注目し, 観測者が知覚する遅延に差が生じる動きの特性を調査した. 具体的には, 表情変化と頭部の動きを同時に伴

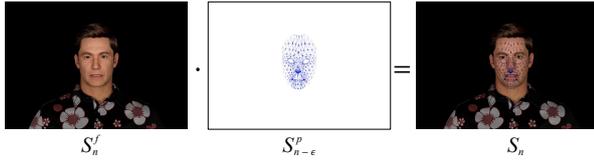


図 1: シミュレーションによる DFPM の投影遅延. 顔映像 S_n^f と投影映像 $S_{n-\epsilon}^p$ を合成し, DFPM の投影遅延映像 S_n を再現.

う複雑な動きに対して, 被験者実験により遅延時間の弁別閾を比較調査し, 特定の刺激間で遅延時間の弁別閾が異なることを確認した. また, 関連研究 [7] を参考に, 投影対象である顔の動き分布の特徴量と弁別閾の関係を分析し, 弁別閾に相関の高い特徴量を確認した.

2. DFPM を模倣したシミュレーション映像

遅延時間の弁別閾を調査するために, 従来研究 [9] と同様に, DFPM を模倣し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を利用することで, 投影対象の顔の動きや追従する投影映像を一定に保ち, 遅延による投影映像のずれを正確に再現して提示することができる. また, 本実験ではシミュレーション映像で作成された刺激を 2 次元のスクリーンで提示するため, 投影対象の顔を立体視することはできない. しかし, 奥行きの変化が少ない顔の表面に投影される映像の遅延知覚において, 立体視できないことが遅延知覚に大きく影響しないと考えた.

DFPM を模倣し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. Meta Human は DFPM の投影対象として用いられた. 本稿では, 従来研究 [9] とは異なり, Live Link Face³ を用いて, 実際に撮影された顔の動きのアニメーションデータから, Unreal Engine 上の Meta Human を動作させることで刺激の動きを作成した.

DFPM を再現するため, 別々にレンダリングされた Meta Human の顔映像と投影映像を, 式 (1) を用いて合成する. また, 図 1 に再現された DFPM の遅延を示す.

$$S_n = S_n^f \cdot S_{n-\epsilon}^p \quad (1)$$

S_n は, 合成後のシミュレーション映像の n フレーム目の画像, S_n^f は顔映像の n フレーム目の画像, $S_{n-\epsilon}^p$ は投影映像の $(n-\epsilon)$ フレーム目の画像を示す. 画像の合成のために, n フレーム目の顔画像 S_n^f と, $(n-\epsilon)$ フレーム目の投影画像 $S_{n-\epsilon}^p$ の対応する位置の RGB の各画素値を乗算した. 変数 ϵ により, 顔画像 S_n^f に投影される投影画像 $S_{n-\epsilon}^p$ の合成フレームを制御し, 顔に対する投影映像の遅延を再現した.

また, 顔画像 S_n^f は 1,000 fps, 投影画像 $S_{n-\epsilon}^p$ は 10,000 fps でレンダリングし, 合成後のシミュレーション画像 S_n を 1,000 fps で投影した. 1,000 fps の顔映像 S^f に対して,

投影映像 S^p を 10 倍の時間分解能でレンダリングすることで, 顔画像 S_n^f に合成される投影画像 $S_{n-\epsilon}^p$ を 0.1 ms 単位で設定できるようにした.

3. 被験者実験

3.1 実験設計

先に述べた従来研究 [9] において, 顔の並進, 左右への回旋, 話している動きで知覚される遅延が, 動きが複雑になるほど気づきにくくなる可能性が示唆された. そこで, 本実験では, 顔が一般的に行う表情変化に注目した. 具体的には, 表情変化と頭部の動きを同時に伴う複雑な動きの刺激を複数用意し, 動きの間で遅延時間の弁別閾に差が生じるかどうか調査することを目的とした.

実験では, 繰り返しペアとして提示される刺激に対して, 遅延が知覚される刺激を選択する Two-Interval-Forced-Choice (2IFC) の方法を利用した [10]. 提示されるペアの映像は, 一方が遅延のない映像, もう一方は図 1 における変数 ϵ によって, 遅延時間が設定された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.

投影の遅延時間は, ステアケース法に基づき決定された [11]. 被験者が, 遅延のない刺激を正しく選択した場合は遅延時間を減少させ, 誤った選択をした場合は遅延時間を増加させた. 遅延時間の増加と減少は, 基本ステップサイズによって制御された. 本実験の場合, 基本ステップサイズの初期値は 1.6 ms に設定し, 遅延時間を増加させる際は, 基本ステップサイズの 3 倍, 減少は常に基本ステップサイズと等しい値だけ減少させた. 被験者の選択で反転 (正しい選択の後に誤った選択, または誤った選択の後に正しい選択) が生じた際, その都度, 基本ステップサイズを半減させた. 実験は, 10 回の反転が起こるまで続けられ, 最終的な遅延時間の弁別閾は, ステップサイズが最小になった後の偶数回目の反転における遅延時間を平均することで求められた.

3.2 実験刺激

図 2 に作成された 5 種類の刺激を示す. 5 種類の刺激は, うなづく動きを 3 回繰り返しながら, 表情が変化する. なお, うなづく動きは 5 種類の刺激で同じアニメーションデータを用いて作成され, 全ての刺激で同じ動きである. Expressionless は, うなづく動きの間, 表情が変化しない刺激である. Blinking は, うなづく動きで顔が正面を向いたタイミングで, 一度瞬きを行う. Smiling は顔が正面を向いたタイミングで目や口が動き, 笑っている表情に変化する. Surprised は, 顔が正面を向いたタイミングで驚いた表情に変化する. Talking は刺激の提示時間中, 常に口が細かく動き, 話している動きを行う.

また, 刺激の速度により変化する投影映像のずれが, 特定の刺激で大きければ, 遅延を知覚するうえでの重要な手掛かりになりうる. 本実験では, 動きの違いによる遅延時間の弁別閾に焦点を当てているため, 顔と投影映像のずれの大きさを統制した. 具体的には, 上記 5 種類の刺激において, 前後 2 フレーム間における画像内での最大移動ベク

¹<https://www.unrealengine.com/ja>

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

³<https://apps.apple.com/jp/app/live-link-face/id1495370836>

表 1: 各刺激の移動ベクトルのノルムの最大値

Expressions	Expressionless	Blinking	Smiling	Surprised	Talking
最大値 (px)	5.22	5.18	5.32	5.25	5.56

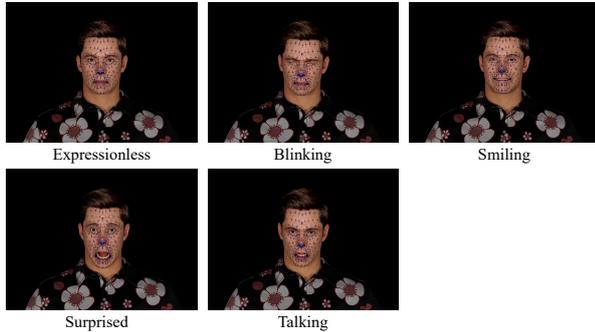


図 2: 表情の異なる 5 種類の実験刺激

トルのノルムを計算し、全フレームを通しての最大値が同程度になるように刺激を作成した。表 1 に各刺激の移動ベクトルのノルムの最大値を示す。

3.3 実験環境

図 3 に被験者実験の環境を示す。実験には、4 名が被験者として参加した⁴。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刺激を、1.5 m 離れた位置から観察した。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、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。なお、刺激の初期遅延時間は 200 ms に設定された。刺激は、24 ビットの RGB 画像を 1,000 fps で投影可能な高速プロジェクタで投影された [12]。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。

3.4 結果

図 4 に、被験者実験により求められた遅延時間の弁別閾の平均値と標準偏差を示す。ANOVA 検定により刺激間の弁別閾の平均値が、少なくともひとつの刺激で、有意に異なることが示された。

表 2 に 5 種類の動きの間で Holm-Bonferroni の多重比較検定を行った結果を示す [13]。表には、算出された p 値が示されており、補正後の p 値が 1.0 を超えた場合は、1.0000 と表記されている。Expressionless と Smiling の刺激間に 1.0 % 水準の有意差、Blinking と Smiling に 5.0 % 水準の有意差が確認された。

表 2: 5 種類の実験刺激における遅延時間の有意差

	Expressionless	Blinking	Smiling	Surprised
Blinking	1.0000	-	-	-
Smiling	0.0035**	0.0167*	-	-
Surprised	0.1733	0.5857	0.3143	-
Talking	0.5857	1.0000	0.0699	1.0000

** $p < 0.01$ * $p < 0.05$

⁴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。

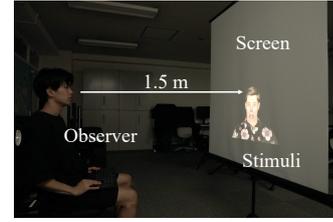


図 3: 被験者実験の環境

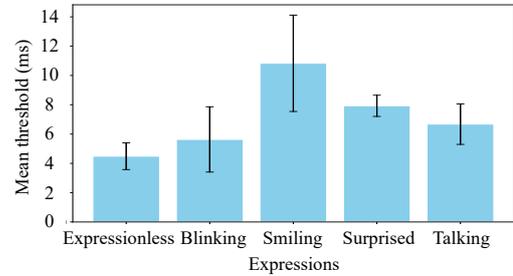


図 4: 5 種類の実験刺激ごとの弁別閾の平均値と標準偏差。青の棒グラフが弁別閾の被験者全体の平均値、エラーバーが標準偏差を示す。

4. 実験刺激の動き分布の特徴量の分析

4.1 概要

次に、被験者実験により求められた弁別閾の差が、どのような動きの特徴に関係しているのかを調査する。そのため、関連研究 [7] を参考に、被験者実験で用いられた刺激の顔の動き分布の特徴量を算出し、遅延時間の弁別閾との関係を分析した。

4.2 オプティカルフロー特徴量

動き分布の分析のため、オプティカルフローによる特徴量を算出した。本実験では、7 つの特徴量を用いた。具体的には、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 x 方向成分の大きさ M_x 、 y 方向成分の大きさ M_y 、ノルム M_{xy} 、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 x 方向成分の大きさの分散 V_x 、 y 方向成分の大きさの分散 V_y 、ノルムの分散 V_{xy} 、方向の分散 V_d である。

5 種類の動きの顔映像に対して、上記の特徴量を算出し、被験者実験により得られた遅延時間の弁別閾と、それぞれの特徴量の間的相关係数を算出した [14]。なお、オプティカルフローは Lucas-Kanade 法を用いて算出した。

4.3 結果

表 3 に、各特徴量と弁別閾との相関係数を示す。オプティカルフローによる動きの特徴量の分析の結果から、弁別閾と特徴量 V_d の間に 0.730 の正の相関があることがわかった。また、表 4 に実際に算出された特徴量 V_d の値を示す。

5. 考察

被験者実験の結果から、異なる表情変化をする Expressionless と Smiling、Blinking と Smiling の間で、遅延知

表 3: 各オプティカルフロー特徴量と弁別閾の相関係数

特徴量	M_x	M_y	M_{xy}	V_x	V_y	V_{xy}	V_d
相関係数	-0.325	-0.576	-0.559	-0.346	-0.593	-0.585	0.730

表 4: 特徴量 V_d の値

Expressions	Expressionless	Blinking	Smiling	Surprised	Talking
V_d	116.746	117.467	119.290	120.464	117.215

覚の弁別閾に違いが生じることを確認した。表 3 の動き分布の特徴量の分析の結果から、 V_d に 0.730 の正の相関があることがわかった。また、表 4 より、有意差が確認された Expressionless と Smiling, Blinking と Smiling における特徴量 V_d の値を比較すると、実際に Smiling の刺激の値が、Expressionless と Blinking の値より高いことが確認された。このことから、遅延時間の弁別閾に関する動きの分布の特徴量として、 V_d が因子の 1 つになりうる可能性があると考えられる。

しかし、表 4 から、Surprised の刺激の特徴量 V_d の値が、2 つの刺激に対して有意差が確認された Smiling の刺激に比べて高く、特徴量 V_d の値の順番が、図 4 から確認される弁別閾の平均値の順番に一致していないことがわかった。このことから、 V_d の特徴量 1 つのみで弁別閾を説明できる可能性は低く、その他の因子が関わっている可能性が高いと考える。

6. まとめ

本稿では、DFPM の投影対象である顔の動きと観測者が知覚する遅延の弁別閾の関係を調査した。本実験では、非剛体的で複雑な顔の動きである表情変化に注目し、遅延時間の弁別閾を比較調査した。被験者実験の結果、表情の違いにより、特定の刺激間で弁別閾の有意差を確認した。次に、弁別閾に差を生じさせる要因を調査するため、オプティカルフロー特徴量と弁別閾の相関係数を分析した。分析の結果、動き方向の分散である特徴量 V_d と遅延時間の弁別閾の相関係数が 0.730 であることを確認した。しかし、特徴量 1 つのみで弁別閾を説明できる可能性は低く、その他の因子が関わっている可能性が高いため、更なる調査が必要であると考えられる。

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] Jennifer Richler, Thomas J. Palmeri, and Isabel Gauthier. Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology*, Vol. 3, pp. 1–6, 2012.
- [4] Kanwisher Nancy, McDermott Josh, and Marvin M. Chun. The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized For Face Perception. *Journal of Neuroscience*, Vol. 17, No. 11, pp. 4302–4311, 1997.
- [5] James W. Tanaka and Martha J. Farah. Parts and Wholes in Face Recognition. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, Vol. 46, No. 2, pp. 225–245, 1993.
- [6] Yang Yung-Hao, Fukiage Taiki, Sun Zitang, and Shin'ya Nishida. Psychophysical measurement of perceived motion flow of naturalistic scenes. *iScience*, Vol. 26, No. 12, pp. 1–19, 2023.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] Gustav Theodor Fechner. *Elemente der Psychophysik*, Vol. 2. Breitkopf & Härtel, Leipzig, 2nd edition, 1889.
- [11] Kaernbach Christian. Simple adaptive testing with the weighted up-down method. *Perception Psychophysics*, Vol. 49, No. 3, pp. 227–229, 1991.
- [12] Watanabe Yoshihiro and Ishikawa Masatoshi. High-speed and high-brightness color single-chip DLP projector using high-power led-based light sources. In *The International Display Workshops*, 2019.
- [13] Hommel Günter. A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, Vol. 75, No. 2, pp. 383–386, 1988.
- [14] Pearson Karl. VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London*, Vol. 187, pp. 253–318, 1896.