

ビデオマグニフィケーションが体感品質に与える影響の分析

武田 翔一郎^{1,a)} 亀田 明男¹ 磯貝 愛¹ 木全 英明¹

概要: ビデオマグニフィケーションは、ヒトの視覚では捉えることが困難な映像中の微細な変化を強調・可視化する技術だが、被写体に素早く大きい変化があると、強調後の映像にノイズが発生するという課題を抱えている。これまでに、時系列データの滑らかさを評価する躍度という指標を用いることで、素早く大きい運動を無視しながら、微細な変化のみを強調する新たなビデオマグニフィケーション手法が提案されている [9]。この手法は、定性的および定量的評価を通じて、従来発生していたノイズを著しく抑制できることが確認されているが、強調された映像がユーザに与える体感品質については明らかにされていない。本稿では、微細な変化の強調率に応じた、変化に対するユーザの気づきやすさと自然さの関係性について主観評価を行い、その結果を報告する。

A Study of Quality of Experience Assessment for Video Magnification

Abstract: Video magnification reveals subtle changes invisible to the naked eye, but artifacts appear in magnified video in the case of scene including non-linear large motion. Jerk-aware video magnification can only pass subtle change under non-linear large motions by using the jerk-based smoothness and is expected to obtain impressive magnification results without artifacts. However, it is not clear what quality of user experience is affected. In this paper, we evaluate quality of user experience for video magnified according to the rate of amplification factor.

1. はじめに

映像中に存在する微細な変化を強調して表現する映像合成技術であるビデオマグニフィケーションは、今までヒトの視覚では捉えることが困難であった、スポーツ中の肌や筋の微細な変形や、エンジンなどで生じる異常振動、または、音を奏でる弦楽器の振動など、様々な現象を容易に捉えることができるようになる映像合成技術として、近年注目を浴びている [7][5][6][1][3]。しかし、微細な変化だけでなく、被写体やカメラが横に移動したり、飛び跳ねたりするよ提案うな、ゆっくり又は素早く大きい運動が映像中に存在する場合、微細な変化だけを検出することが困難になり、ノイズが発生することが知られている。

従来研究 [6][1][3] は、ゆっくりで大きい運動が存在する映像から、微細な変化だけを強調することに成功しているが、素早く大きい運動が存在する映像に適用すると、強調後の映像にノイズが発生するという課題は残っている。

このような課題に対して、近年、映像中の素早く大きい

運動に対して微細な変化は時系列的に滑らかに変化しているという観察結果を基に、滑らかさを評価する指標として神経科学や機械工学の分野で用いられてきた躍度を用いたビデオマグニフィケーションが提案されている [9]。

この手法が、素早く大きい運動が存在する映像から微細な変化だけを検出し、ノイズを抑えながら微細な変化のみ強調できることは定性的および定量的評価により確認されているが、強調された映像がユーザに与える体感品質は明らかとなっていない。

素早く大きい運動が存在する映像中の微細な変化を強調することで、ユーザはこれまで捉えられていなかった変化に気づきやすくなり、副次的効果として、ユーザに与える迫力感の増加が期待される。一方で、微細な変化を強調しすぎると、物体の歪みや本来伝えたい現象以外の微細な変化の強調など、違和感のある変化が顕著となり、自然さが損なわれると考えられる。

そこで、本稿では、躍度に基づくビデオマグニフィケーションを用いて、微細変化の強調率を増加させたとき、ユーザが感じる変化の気づきやすさと自然さについて主観評価を行い、その関係性を報告する。

¹ NTTメディアインテリジェンス研究所
画像メディアプロジェクト

^{a)} shoichiro.takeda.us@hco.ntt.co.jp

2章で関連研究を紹介した後、3章で本評価で用いた躍度に基づいたビデオマグニフィケーションを説明する。その後、4章で評価実験方法を示し、5章で結果と6章で考察を述べる。

2. 関連研究

ビデオマグニフィケーションは、映像中のある固定されたブロック毎に時間方向に変動する輝度または位相情報を捉えることで、そのブロック毎に被写体の色（輝度）または運動（位相）を検出・強調する手法である [7][5]。しかし、ブロックの大きさによって捉えることのできる変化の大きさが異なるため、映像を複数解像度で表現された映像ピラミッドに分解し、様々な大きさの輝度または位相情報を映像中から取得する。その後、各ピラミッド階層のブロック毎の輝度または位相の時系列変化に対して、バンドパスフィルタを適用し、任意周波数の微細な色または運動の変化を得る。そして、取得した微細変化を元の変化に α 倍して加算し、微細変化を強調する。最後に映像ピラミッドを再構成することで、映像中の微細変化が強調された映像を得る。しかし、従来手法では、映像中の微細変化と大きな運動をブロック内で切り分けることができないという問題を抱えていた。

大きな運動が存在する中でも、微細変化のみを検出・強調するためにレイヤーに基づいた手法が提案されている [1][3]。これらの手法は、ユーザーによる強調領域をの指定 [1]、または、デプスカメラを用いて強調領域と背景を切り分けることで [3]、大きい運動による影響を軽減する手法である。しかしながら、これらの手法は、人手による操作や [1]、デプスカメラなどの特殊な撮影環境が要求され [3]、手法の適用範囲が限定的である。

一方で Zhang らは、ゆっくりで大きい運動は時間スケールで見ると線形近似できることに着目し、微細な加速度変化のみを検出・強調することを提案した [8]。この手法は、ゆっくりで大きい運動が映像中に存在していても、追加の手間や情報なしで微細変化のみを強調することを可能にした。しかし、映像中に素早く大きい運動が存在すると、それを誤認識し、大きなノイズを生じさせてしまうという課題を持っている。

3. 躍度に基づいたビデオマグニフィケーション [9]

本稿では、素早く大きい運動にも頑健性の高い躍度に基づいたビデオマグニフィケーション [9] を評価に用いる。躍度に基づいたビデオマグニフィケーションは、Zhang らの手法 [8] に対して、滑らかさを評価する指標である躍度に基づく時空間フィルタを適用した手法 [9] である。最初に、Zhang らの手法 [8] について説明し、次に、躍度に基づ

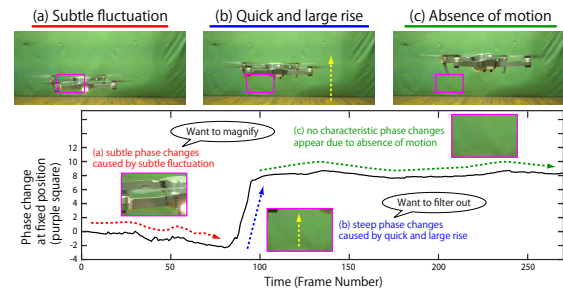


図 1 映像中の運動（位相）変化。ドローンの微細な揺れによる微細な位相変化と、素早い浮上による急峻な位相変化が生じている。この急峻な変化と比較して、微細な変化はその振幅の小ささゆえに、滑らかに変化していると観察できる。

づいた時空間フィルタを Zhang らの手法に適用する方法を述べる。

3.1 Video Acceleration Magnification [8]

本稿ではスペースの都合上、微細な運動変化の強調手順について説明する（色の強調に関しては、[8]を参照）。入力映像を $I(\mathbf{x}, t)$ とする時（ \mathbf{x} は位置 (x, y) 、 t は時間）、ある解像度を持つピラミッドの階層を l 、空間周波数を ω 、及び空間方向選択性を θ とした時、複数解像度において運動情報を表す位相変化 $\phi_{\omega, \theta}^l(\mathbf{x}, t)$ を得るため、complex-steerable filter $\psi_{\omega, \theta}^l$ を入力映像に対して以下のように適用する：

$$(I(\mathbf{x}, t) \otimes \psi_{\omega, \theta}^l)(\mathbf{x}, t) = A_{\omega, \theta}^l(\mathbf{x}, t) e^{i\phi_{\omega, \theta}^l(\mathbf{x}, t)}. \quad (1)$$

その後、任意周波数 f に応答性を持たせたガウシアンフィルタ $G_f(t)$ を時系列方向に畳み込んだ上で、2回微分 $\frac{\partial^2}{\partial t^2}$ を適用し、微細運動を検出する：

$$C_{f, \omega, \theta}^l(\mathbf{x}, t) = \frac{\partial^2 G_f(t)}{\partial t^2} \otimes \phi_{\omega, \theta}^l(\mathbf{x}, t). \quad (2)$$

最後に、強調率 α を乗算し、元の位相変化に加算することで、微細運動成分のみが強調された位相変化を得る：

$$\hat{\phi}_{f, \omega, \theta}^l(\mathbf{x}, t) = \phi_{\omega, \theta}^l(\mathbf{x}, t) + \alpha C_{f, \omega, \theta}^l(\mathbf{x}, t). \quad (3)$$

この Zhang らの手法 [8] は、位相変化の加速度成分のみを検出することで、線形近似が可能なくゆっくりで大きい運動を無視できるが、素早く大きい運動が存在すると、それを誤検出し、強調してしまう問題がある。

3.2 躍度に基づいた時空間フィルタ

微細変化は素早く大きい運動と比較して時系列的に滑らかに変化するという観察結果（図 1）から、時系列データの滑らかさを評価する指標として神経科学や機械工学の分野で用いられてきた躍度と呼ばれる指標を利用し、微細変化のみを透過させ、素早く大きい運動の影響を除去するフィルタを設計する。

式 (1) で得られた位相変化 $\phi_{\omega, \theta}^l(\mathbf{x}, t)$ に対する躍度を計

算するために、任意周波数 f に応答性を持たせたガウシアンフィルタ $G_f(t)$ を時系列方向に畳み込んだ上で、躍度を導出するための三回微分 $\frac{\partial^3}{\partial t^3}$ を適用する：

$$Jerk_{f,\omega,\theta}^l(\mathbf{x}, t) = \frac{\partial^3 G_f(t)}{\partial t^3} \otimes \phi_{\omega,\theta}^l(\mathbf{x}, t). \quad (4)$$

次に、この位相変化から算出された躍度を基に、位相変化の滑らかさを求める。滑らかな位相変化の時には1、素早く大きい位相変化が生じている時には0になるように、以下のように躍度の値を正規化し、反転させる。

$$nJerk_{f,\omega,\theta}^l(\mathbf{x}, t) = \frac{|Jerk_{f,\omega,\theta}^l(\mathbf{x}, t)| - \min_{\mathbf{x}, t} |Jerk_{f,\omega,\theta}^l(\mathbf{x}, t)|}{\max_{\mathbf{x}, t} |Jerk_{f,\omega,\theta}^l(\mathbf{x}, t)| - \min_{\mathbf{x}, t} |Jerk_{f,\omega,\theta}^l(\mathbf{x}, t)|}, \quad (5)$$

$$smoothness_{f,\omega,\theta}^l(\mathbf{x}, t) = 1 - nJerk_{f,\omega,\theta}^l(\mathbf{x}, t). \quad (6)$$

最後に、フィルタの抑制効果を調整可能にするために、ハイパーパラメータ β を導入し、躍度に基づいた時空間フィルタ $JAF_{f,\omega,\theta}^l(\mathbf{x}, t)$ を得る：

$$JAF_{f,\omega,\theta}^l(\mathbf{x}, t) = smoothness_{f,\omega,\theta}^l(\mathbf{x}, t)^\beta. \quad (7)$$

3.3 階層情報を利用したフィルタ補正

上記のフィルタは、映像ピラミッドの各階層 l 毎に求められるが、階層毎の解像度の違いにより、算出される躍度の大きさが変わってくる [4][2]。そこで、ダウンサンプルファクタ ($0 < \lambda < 1$) を用いて、階層毎に求められたフィルタを以下の式で補正する：

$$JAF_{\lambda,f,\omega,\theta}^l(\mathbf{x}, t) = JAF_{f,\omega,\theta}^l(\mathbf{x}, t)^{1/\lambda}. \quad (8)$$

また、躍度に基づく時空間フィルタは、オプティカルフロー等で用いられてきた疎密戦略と同じように、階層を跨いだ情報のやり取りによって補正する必要がある [4][2]。したがって、階層 l のフィルタをそれよりも上の階層 $l+N$ のフィルタで補正する、伝搬補正というものを以下の式で定義する。

$$pJAF_{\lambda,f,\omega,\theta}^l(\mathbf{x}, t) = \prod_{i=l}^{l+N} res(JAF_{\lambda,f,\omega,\theta}^i(\mathbf{x}, t), l). \quad (9)$$

$res(JAF_{\lambda,f,\omega,\theta}^i(\mathbf{x}, t), l)$ は、バイキュービック補間を示している。これらの補正を通じて、躍度に基づいた時系列フィルタ $pJAF_{\lambda,f,\omega,\theta}^l(\mathbf{x}, t)$ を得ることができる。

このフィルタを、Zhang らの手法 (式 (2, 3)) に以下のように適用することで、ゆっくり及び素早く大きい運動を無視し、微細運動のみを映像中から検出し、強調することができる。

$$\hat{C}_{\lambda,f,\omega,\theta}^l(\mathbf{x}, t) = pJAF_{\lambda,f,\omega,\theta}^l(\mathbf{x}, t) \times C_{f,\omega,\theta}^l(\mathbf{x}, t). \quad (10)$$

$$\hat{\phi}_{\lambda,f,\omega,\theta}^l(\mathbf{x}, t) = \phi_{\omega,\theta}^l(\mathbf{x}, t) + \alpha \hat{C}_{\lambda,f,\omega,\theta}^l(\mathbf{x}, t). \quad (11)$$

4. 主観評価実験

本章では、主観評価で用いた評価コンテンツおよび評価観点を説明した後、実験方法について説明する。

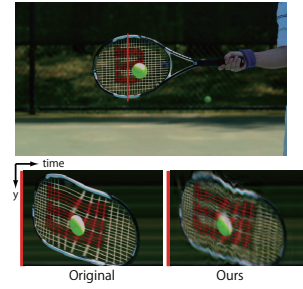


図 2 Tennis：素早く大きい運動と微細な変化。下段 (左) は、元映像において赤線のスライスを時間方向にスタックした画像。下段 (右) は、強調後の映像において赤線のスライスを時間方向にスタックしたものであり、ボールが当たった瞬間にガットに広がる微細な変化が強調されている。

表 1 評価コンテンツ

Video	Resolution	time(s)
Gun	720 × 576	11.5
Tennis	640 × 360	12.07
Wood	640 × 360	12.13
Slam dunk	582 × 328	15.0

4.1 評価コンテンツ

素早く大きい運動に微細な変化が含まれる 4 種類のコンテンツを用意した (表 1)。例えば、評価コンテンツ Tennis には、スイングという素早く大きい運動中に、ボールがラケットに当たった瞬間にガットに伝わる微細な変化が含まれている (図 2)。

各コンテンツに対して、強調率 α は $\alpha = [6, 12, 18, 24, 30]$ と変化させ、微細な変化を強調した映像を作成した。以降、微細な変化が強調されていない元映像は、強調率 $\alpha = 0$ の映像として表現する。

4.2 評価観点

微細な変化の強調率 α の増加に伴い、ユーザが変化の違いに気づきやすくなり、迫力を感じるようになる一方で、不自然さにどのような影響を与えるのか評価するため、下記のような評価観点を設定した。

- 変化の気づきやすさ：微細な変化に気づきやすくなったか、映像中の「変化の度合い」を 5 点：強い～1 点：弱い の 5 段階で評価する。
- 迫力感の増大効果：素早く大きい運動中の微細な変化を強調することで、映像に「迫力感」があったかを 5 点：迫力がある～1 点：物足りない の 5 段階で評価する。
- 自然さの維持効果：微細な変化を強調したとき、違和感のある変化であると感じ、「不自然さ」が気になるかを 5 点：気にならない～1 点：非常に気になる の 5 段階で評価する。

4.3 実験方法

微細変化の強調率の大きさによって与えるユーザ体感品質を評価するため、基準映像(元映像)と強調率 α を変化させた評価映像を対にして2回提示し、2回目の提示時に両映像に対して評価を行った。このとき、DCR (Degradation Category Rating) 法 [10] を参考に、基準映像の後に評価映像を提示することで順序効果の影響を抑制している。

被験者は20代から50代の男女12名で、事前に注視してほしい領域を伝えて評価する公開グループと、事前情報なく評価する非公開グループの2つに分けて実施した。被験者には、評価尺度が連続的に表現されているスケール上に、評価を棒線で記入してもらい、小数第一位までを各映像の評価値として換算する。その後、基準映像に比べてどのくらいの差があるか評価するため、DSCQS(Double Stimulus Continuous Quality Scale) 法 [11] を参考に、微細な変化を強調した評価映像の評価値と基準映像の評価値との差分を最終的なスコアとする。

5. 評価結果

図3-図6に、スコアの平均を縦軸、強調率 α を横軸として、変化の度合い、迫力感、不自然さの体感品質に関する主観評価結果を示す。

公開・非公開グループともに、強調率が大きくなるにしたがって、変化の度合いが大きいと感じるという結果が得られた。特に、事前に注視してほしい領域を伝えている公開グループでは、コンテンツによって強調率は異なるものの、基準映像と強調された評価映像との間に有意差 ($p < 0.05$) が認められた。その一方で、強調率が大きくなるにしたがって、強調された映像に不自然さを感じる傾向が見られる。また、副次的な効果として想定していた迫力感は、強調率の違いによって大きな差がなく、有意差は認められなかった。

6. 考察

全体的な傾向として、微細変化の強調率を大きくすることで、変化に気づきやすくなることから、これまでヒトの視覚では捉えることが困難であった事象を伝えるのに有効であることが示唆された。その一方で、微細変化の強調率にしたがって、基準映像に比べて評価映像の不自然さを強く感じるという結果が得られた。今回の評価実験では、基準映像と評価映像のそれぞれに評価するものの、必ず基準映像と評価映像を対で視聴するため、基準映像にはない変化そのものを不自然であると感じ評価した可能性が考えられる。今後は、「不自然さ」のように映像全体を評価する形容詞ではなく、物体形状の歪みや、目的とした現象以外の微細な変化の強調が与える影響を表現する形容詞の検討、単一の映像で評価を行うACR(Absolute Category Rating Method) 法での評価が課題である。

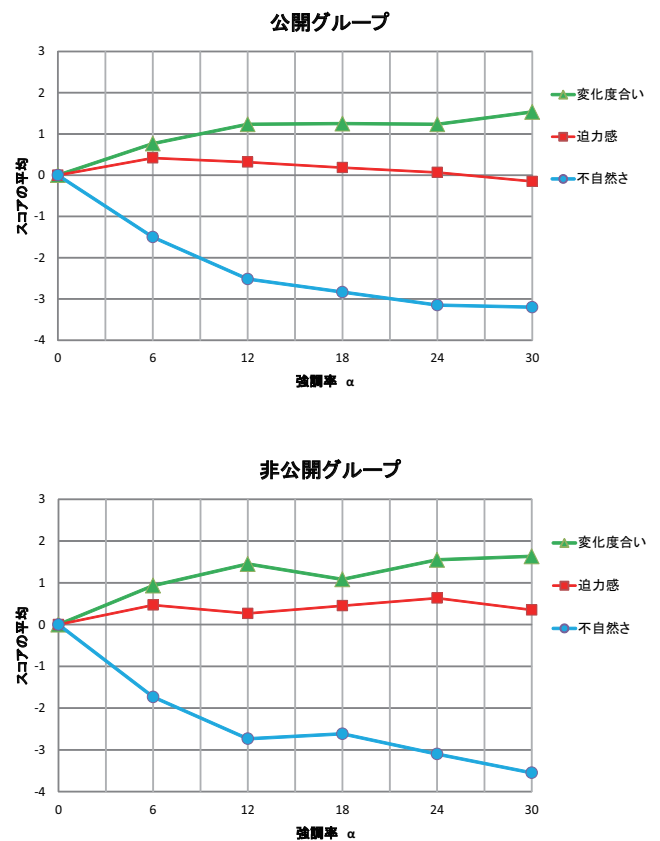


図3 コンテンツ Gun の評価結果。

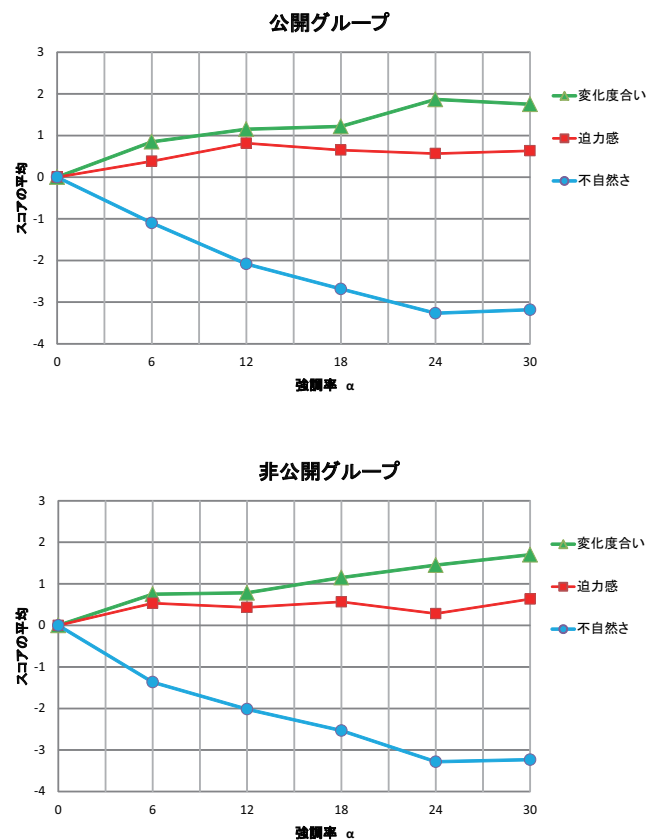


図4 コンテンツ Tennis の評価結果。

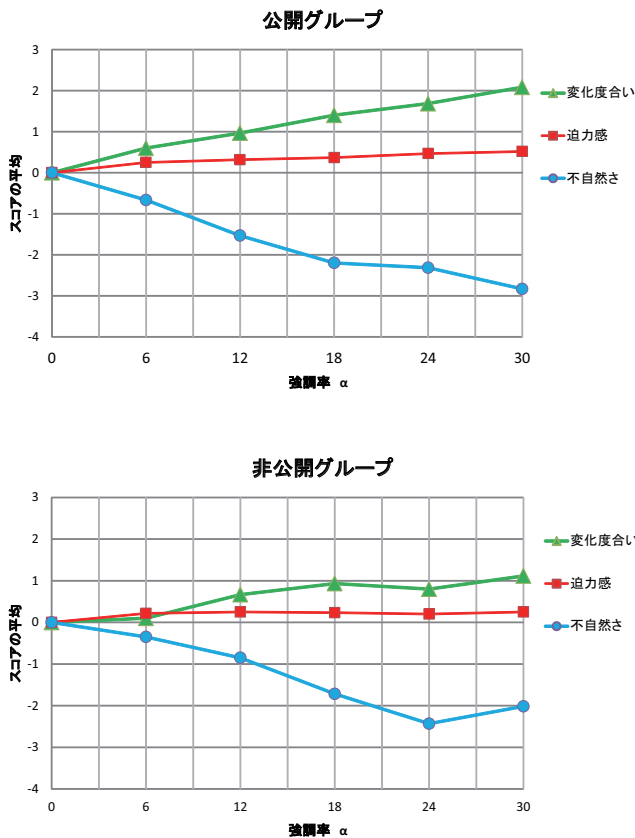


図 5 コンテンツ Wood の評価結果.

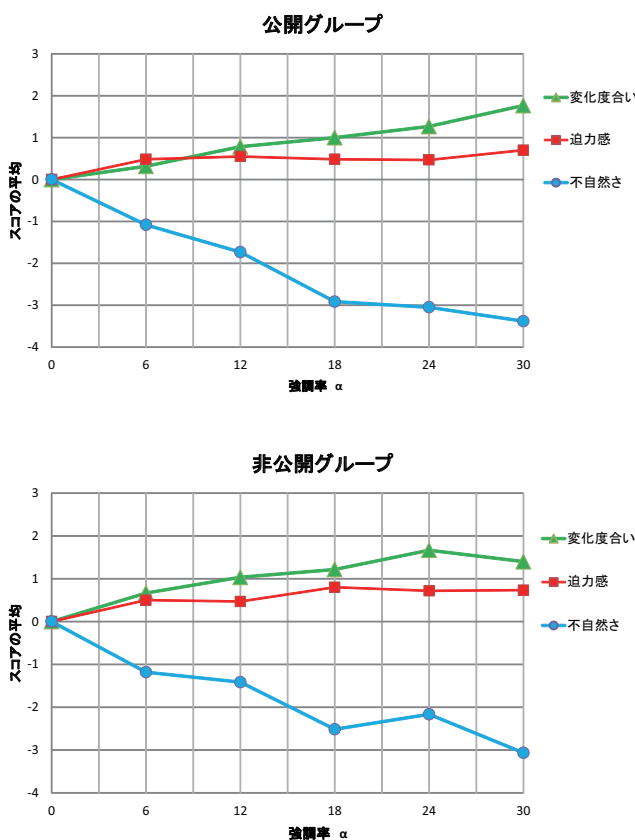


図 6 コンテンツ Slam dunk の評価結果.

最後に、迫力感については、基準映像と評価映像の評価値の差分であるスコアには有意差は見られないものの、基準映像に対して迫力感が上昇する傾向はある。微細変化の強調が迫力感に寄与する可能性を確認するため、基準映像と評価映像の評価値 (MOS 値) を用いて、基準映像と評価映像群で t 検定を実施したところ、公開グループで評価した Gun 以外は、基準映像と評価映像群の間に有意差が認められた。ただし、微細変化を強調したときに感じた迫力とコンテンツ内容との依存関係については明らかとなっておらず、微細変化の強調が与える副次的な体感品質の評価は課題である。

7. まとめ

本稿では、素早く大きい運動が存在する映像中の微細な変化を強調することで、ユーザに与える体験品質について評価を行った。主観評価実験の結果、強調率を大きくすることで変化に気づきやすくなり、これまでヒトが捉えられていなかった事象を伝えるのに有効であることが確認された。今後は、微細変化の強調によって生じる不自然さの評価手法の再検討および評価、主観評価に基づく強調率の最適化など課題として取り組んでいく。

参考文献

- [1] Elgharib, M. A., Hefeeda, M., Durand, F. and Freeman, W. T.. Video magnification in presence of large motions, CVPR (2015).
- [2] Hu, Y., Song, R. and Li, Y.. Efficient coarse-to-fine patch-match for large displacement optical flow, CVPR (2016).
- [3] Kooij, J. F. and van Gemert, J. C.. Depth-aware motion magnification, ECCV (2016).
- [4] Meyer, S., Wang, O., Zimmer, H., Grosse, M. and Sorkine-Hornung, A.. Phase-based frame interpolation for video, CVPR (2015).
- [5] Wadhwa, N., Rubinstein, M., Durand, F. and Freeman, W. T.. Phase-based video motion processing, SIGGRAPH, Vol. 32, No. 4, p. 80 (2013).
- [6] Wadhwa, N., Rubinstein, M., Durand, F. and Freeman, W. T.. Riesz pyramids for fast phase-based video magnification, ICCP (2014).
- [7] Wu, H.-Y., Rubinstein, M., Shih, E., Gutttag, J., Durand, F. and Freeman, W.. Eulerian Video Magnification for Revealing Subtle Changes in the World, SIGGRAPH, Vol. 31, No. 4 (2012).
- [8] Zhang, Y., Pinteá, S. L. and van Gemert, J. C.: Video Acceleration Magnification, CVPR (2017).
- [9] Takeda, S., Okami, K., Mikami, D., Isogai, M. and Kimata, H.. Jerk-Aware Video Acceleration Magnification, CVPR (2018).
- [10] ITU-T Recommendation P.910: "Subjective video quality assessment methods for multimedia applications" (1999)
- [11] ITU-R Recommendation BT-500-11: "Methodology for the subjective assessment of the quality of television pictures", (2002)