

一枚の画像からの時空間周波数を用いた揺らめき動画生成法

山本 薫^{1,a)} 坂本 雄児^{1,b)}

概要: 一枚の画像に揺らめくような動きを与えて動画を自動生成する手法がある。本研究では、画像の空間周波数領域に時間周波数軸を加えた三次元の時空間周波数領域上に成分を配置し、三次元フーリエ逆変換を用いて自然界に見られるような複雑さをもつ揺らめきを生成する手法を提案した。また動画生成実験を行い、提案手法によってより自然な揺らめきを生成可能であることを確認した。

1. はじめに

近年、高性能なデジタルカメラやカメラを搭載した携帯機器が普及したことによって、パーソナルコンピュータやタブレット端末、デジタルフォトフレームなどのディスプレイ表示によって写真を鑑賞する機会が増加した。現時点では、写真を静止画として鑑賞するのが主流であり、スライドショーとして複数の写真を切り替えて表示することはあっても、写真が自動的に動き出すといった効果はない。そのため、写真から自動的にアニメーションが生成される機能が加われば、写真の鑑賞をより楽しめるようになると思われる。

一枚の静止画からの動画生成に関してはこれまでにいくつかの研究がなされており、油絵や写真の中の水面や船、木を動かす手法 [1] や、風に揺れる樹木や歩行する人間のアニメーションを生成する手法 [2]、煙や滝、雲が流れるアニメーションを生成する手法 [3] などが提案されている。これらの手法では、力学モデルに基づいた物理シミュレーションを行った結果をもとに、画像中の各画素の輝度や位置を変更することで動画を生成している。しかし、動かす対象物体によって異なる計算方法を選択する必要があるなど、ユーザにとって作業的な負担が大きく、また計算時間や多くの計算機リソースが必要になるといった問題があった。これに対し、動きの種類を「揺らめき」に限定することでより手軽に動画を自動生成する手法 [4] が提案された。この手法では、水、火、木といった異なる物体に対しても同一のルールを適用することで、対象物体ごとに計算方法を変える必要がなくなり、ユーザの作業的な負担を軽減して

いる。また、手法 [4] で問題とされていた、生成される揺らめきが周期的で単調となることに関して、自然界で多く見られる $1/f^\beta$ ノイズを用いて振動の振幅および周波数を変化させ、より自然な揺らめきを生成する手法 [5] が提案された。

一方で、手法 [5] では、空間的な変化を考慮していないため、画像全体が均一に揺らめき、不自然な動きが生成される問題があった。そこで本研究では、より自然な揺らめきを生成する手法を提案する。提案手法では、画像の空間周波数領域に時間周波数軸を加えた三次元の時空間周波数領域を作り、この三次元周波数領域上に成分を適切に配置し、三次元フーリエ逆変換を適用することによって複雑かつ自然な揺らめきを生成する。

2. 揺らめきの生成手法

2.1 揺らめきの生成手法の概要

揺らめきの生成手法 [4] の流れを図 1 に示す。まず入力画像の座標 (x, y) における画素の輝度値を、高さ H_{xy} の初期値として設定する。この高さは擬似的な量であり、物体の本来の高さを表すわけではないが、画像固有の値として与えられる。次に、この高さ H_{xy} で作られる凹凸に仮想光源の光を当て、陰影画像を作成する。具体的には、高さ H_{xy} を用いて画像平面に対してバンプマッピング [6] を行い、得られた法線ベクトルと仮想平行光源ベクトルから Lambert 反射を計算し、陰影画像とする。最後に、生成された陰影画像をもとに入力画像の輝度値を変更し、新たに画像を生成・出力する。

高さ H_{xy} は、次式で表される単振動により時間変化する。

$$H_{xy}(t) = A \sin(\omega t + \alpha_{xy}) + C \quad (1)$$

ここで、 $H_{xy}(t)$ は座標 (x, y) および時刻 t における高さである。先に述べた通り、高さの初期値 $H_{xy}(0)$ は、入力画像

¹ 北海道大学大学院情報科学研究科
Graduate School of Information Science and Technology,
Hokkaido University

a) k.yamamoto@ist.hokudai.ac.jp

b) yuji@ist.hokudai.ac.jp

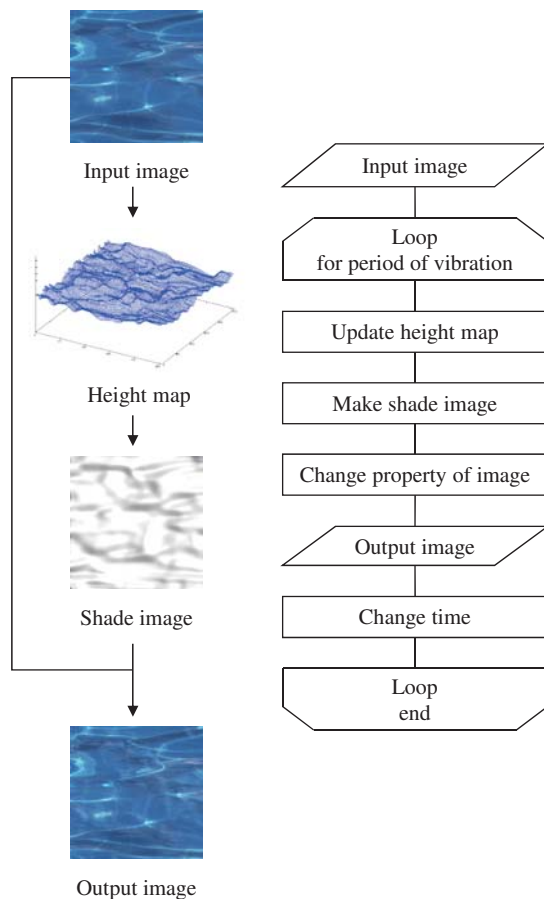


図 1 揺らめきの生成手法の概要
Fig. 1 Overview of generating flicker.

の輝度値で与えられる。高さ $H_{xy}(t)$ は、 $0 \leq H_{xy}(t) \leq 255$ を満たす値を取る。

A は振幅、 C は振動の中心である。これにより、式 (1) の高さ $H_{xy}(t)$ は $C - A \leq H_{xy}(t) \leq C + A$ の範囲で振動する。

ω は角周波数である。振動一回あたりのフレーム数を表す周期 T というパラメータがあり、 ω は T を用いて $\omega = 2\pi/T$ で求められる。

α_{xy} は初期位相である。式 (1) より、時刻 $t = 0$ のときの高さ $H_{xy}(0)$ は次式で表される。

$$H_{xy}(0) = A \sin \alpha_{xy} + C \quad (2)$$

よって、初期位相 α_{xy} は次式で求められる。

$$\alpha_{xy} = \sin^{-1} \left(\frac{H_{xy}(0) - C}{A} \right) \quad (3)$$

α_{xy} が、 $|\alpha_{xy}| \leq \pi/2$ 、または $\pi/2 < |\alpha_{xy}| \leq \pi$ のいずれの範囲の値かによって、高さ $H_{xy}(0)$ の次時刻における増加、または減少が決まる。

高さ $H_{xy}(t)$ の値が時間変化することによって陰画像も変化し、生成される画像も変化するため、生成される画像を連続して出力することで揺らめくようなアニメーションを表現する。

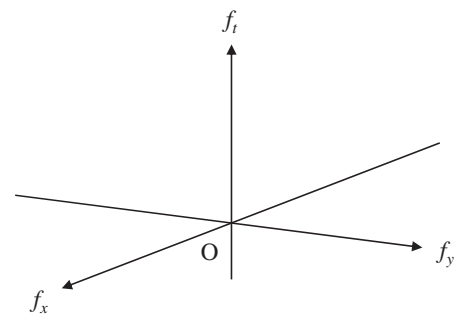


図 2 時空間周波数領域
Fig. 2 Time-space frequency domain.

2.2 自然な揺らめきの生成

動きの生成に用いられる式 (1) において、振幅 A や角周波数 ω は時刻 t によらない定数である。そのため、生成される揺らめきは同一の動きを周期的に繰り返す単調なものとなる。このような単調な動きはユーザに対して不自然な動きという印象を与えることから、より複雑な動きを生成できるように手法の改善が必要であった。

自然界のある特定のノイズは、周波数成分の 2 乗平均であるスペクトル密度について、 f を周波数とすると、 $1/f^\beta$ に比例する特性をもつ。このノイズは風のそよぎや川のせせらぎ、音楽や高速道路上の車の流れなど、自然界の様々な場面で観測される [7]。この $1/f^\beta$ ノイズを利用することによって、揺らめきの時間的な単調さを低減する手法が提案された [5]。

手法 [5] では、生成される揺らめきに空間的な差異がないため、遠近感のある風景や複数の樹木の画像であっても画像全体が等しく動き、不自然な揺らめきが生成される問題があった。そこで、空間的な変化から振動の周期および位相を変化させることによって、より自然な揺らめきを生成する手法を提案する。

3. 提案手法

3.1 時空間周波数法

本手法は、空間的な情報から時間的な情報を得られるという考えに基づいている。例えば樹木について、木全体の動きは遅く葉の動きは速いことが多い。このように、大きな構造は動きが遅く、小さな構造は動きが速い。すなわち、空間的に周波数が低ければ時間的にも周波数が低く、空間的に周波数が高ければ時間的にも周波数が高い場合が多く見られる。これより、時間周波数と空間周波数には関連があると考えられる。

本手法では、画像の空間周波数分布から時間周波数分布を決定する。まず、時空間周波数領域について定義する。時空間周波数領域 (f_x, f_y, f_t) は、二次元の空間周波数領域 (f_x, f_y) に時間周波数軸 (f_t) を加えた、三次元の周波数領域である (図 2)。

図 3 に本手法の流れを示す。はじめに、揺らめきの対

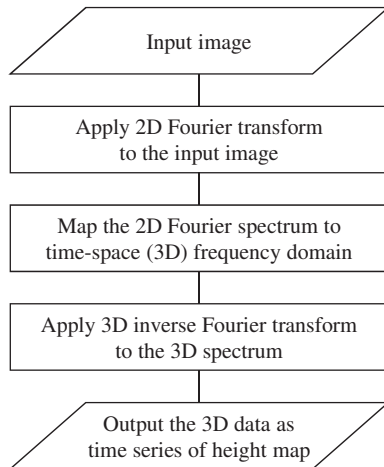


図 3 提案手法の概要

Fig. 3 Overview of proposed study.

象となる静止画像の一部またはすべての領域を取り出し、二次元フーリエ変換を行う。このとき、画像の輝度値ではなく、式 (3) の初期位相に対して二次元フーリエ変換を行う。これにより、揺らめきの生成処理において初期位相が正しく設定される。座標 (x, y) における入力画像の初期位相を $\alpha(x, y)$ とすると、入力画像の二次元フーリエ変換結果 $F(f_x, f_y)$ は次式で表される。

$$F(f_x, f_y) = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} e^{j\alpha(x,y)} e^{-j2\pi(\frac{fx}{X} + \frac{fy}{Y})} \quad (4)$$

ここで、 X, Y はそれぞれ画像の x 方向、 y 方向の画素数、 e はネイピア数、 j は虚数単位である。次に、得られた二次元フーリエスペクトルについて、3.2 節で述べる時空間周波数領域へのマッピングを行い、三次元フーリエスペクトル $F_m(f_x, f_y, f_t)$ を得る。最後に、得られた三次元フーリエスペクトル $F_m(f_x, f_y, f_t)$ に対して次式で表される三次元フーリエ逆変換を行う。

$$I_{out}(x, y, t) = \frac{1}{XYT} \sum_{fx=0}^{X-1} \sum_{fy=0}^{Y-1} \sum_{ft=0}^{T-1} F_m(f_x, f_y, f_t) e^{j2\pi(\frac{fx}{X} + \frac{fy}{Y} + \frac{ft}{T})} \quad (5)$$

ここで、 T は揺らめきの周期であり、生成される動画のフレーム数と等しい。得られた三次元情報 $I_{out}(x, y, t)$ は、時間軸方向から見ると画像の時間変化となっている。 $I_{out}(x, y, t)$ を手法 [4] において高さ $H_{xy}(t)$ として用いることによって、動画を生成する。

3.2 マッピングの種類

二次元のフーリエスペクトルを三次元の時空間周波数領域へマッピングするとき、一般には次の式で表される。

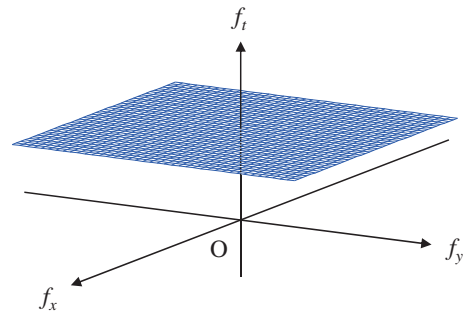


図 4 平面へのマッピング

Fig. 4 Mapping to a plane.

$$\begin{bmatrix} f_x^{3D} \\ f_y^{3D} \\ f_t^{3D} \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x^{2D} \\ f_y^{2D} \\ 0 \\ 1 \end{bmatrix} \quad (6)$$

ここで、 $f_x^{3D}, f_y^{3D}, f_t^{3D}$ は時空間周波数領域上の座標、 f_x^{2D}, f_y^{2D} は画像の空間周波数領域上の座標である。変数 a_{ij} の与え方によって様々なマッピングが考えられるが、現実の揺らめきを考えると意味のない与え方が多いため、本手法では次の二つの条件を満たす場合のみを考える。

一つ目の条件として、 f_t 軸に関して対称であることが挙げられる。 f_t 軸に関して非対称である場合、揺らめきに方向性が生まれるが、本稿では揺らめきの方向性を考慮しない。二つ目の条件として、 f_x, f_y はマッピングの前後で値が等しいことが挙げられる。 f_x, f_y が変化することにより、元の画像の情報が失われてしまうためである。

上記二つの条件を満たすマッピングとして、平面、 f_t 軸の負の方向に凸の曲面、 f_t 軸の正の方向に凸の曲面への 3 種類のマッピングを検討した。

3.2.1 平面へのマッピング

最も単純なマッピングとして、式 (6) について

$$\begin{bmatrix} f_x^{3D} \\ f_y^{3D} \\ f_t^{3D} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & f_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x^{2D} \\ f_y^{2D} \\ 0 \\ 1 \end{bmatrix} \quad (7)$$

としたマッピングが考えられる。ここで、 f_0 は定数である。このマッピングは、図 4 に示すような平面 $f_t = f_0$ へのマッピングであり、画像全体に時間周波数 f_0 が付与されることから、最終的に得られる揺らめきは周波数 f_0 の単振動となる。すなわち、従来手法 [4] のみを用いた場合と同等の揺らめきが生成される。

3.2.2 f_t 軸負方向に凸の曲面へのマッピング

次式で表されるマッピングを考える。

$$\begin{aligned} f_x^{3D} &= f_x^{2D} \\ f_y^{3D} &= f_y^{2D} \\ f_t^{3D} &= (f_x^{2D})^2 + (f_y^{2D})^2 + f_0 \end{aligned} \quad (8)$$

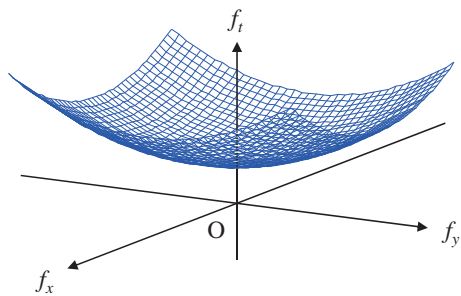


図 5 f_t 軸負方向に凸の曲面へのマッピング
Fig. 5 Mapping to a convex surface.

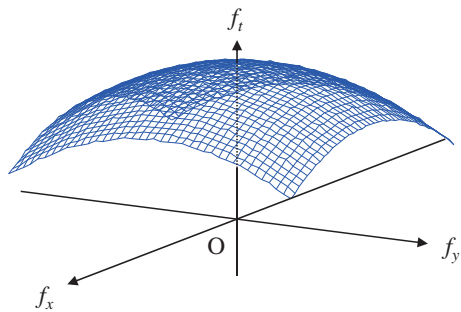


図 6 f_t 軸正方向に凸の曲面へのマッピング
Fig. 6 Mapping to a concave surface.

このマッピングは、図 5 に示すような f_t 軸の負の方向に凸の曲面へのマッピングである。これは空間周波数が高いほど時間周波数が高く、空間周波数が低いほど時間周波数が低くなるようになっている。すなわち、最終的に得られる揺らめきは、構造が大きいほど周期が長く、構造が小さいほど周期が短い振動となる。

3.2.3 f_t 軸正方向に凸の曲面へのマッピング

次式で表されるマッピングを考える。

$$\begin{aligned} f_x^{3D} &= f_x^{2D} \\ f_y^{3D} &= f_y^{2D} \\ f_t^{3D} &= -(f_x^{2D})^2 - (f_y^{2D})^2 + f_0 \end{aligned} \quad (9)$$

このマッピングは、図 6 に示すような f_t 軸の正の方向に凸の曲面へのマッピングである。これは空間周波数が高いほど時間周波数が低く、空間周波数が低いほど時間周波数が高くなるようになっている。すなわち、最終的に得られる揺らめきは、構造が大きいほど周期が長く、構造が小さいほど周期が長い振動となる。ただし、固有振動の観点から、構造が大きいほど周期が長くなる場合が多いため、式 (9) のマッピングによって自然な揺らめきを生成することは困難であると考えられる。

3.3 $1/f^\beta$ ノイズの付加

2.2 節で述べた通り、手法 [5] では、 $1/f^\beta$ ノイズを用いて複雑さを有する揺らめきを生成した。本手法においても、 $1/f^\beta$ ノイズを用いて揺らめきの周期性を抑制することによって、より自然な揺らめきを生成することが可能である



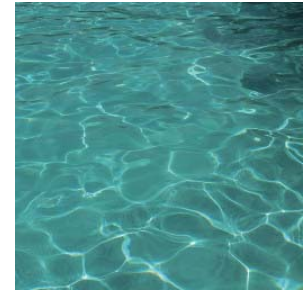
(a) Poplar



(b) Rowan



(c) Fire



(d) Water surface

図 7 動画生成に使用した画像

Fig. 7 Images used for generating animation.

と考えられる。本稿では、手法 [5] において比較的高い評価を得られた $1/f^2$ ノイズを採用する。

本手法においては、マッピングされた面から f_t 軸方向に Δf だけ離れた位置の強度が $1/\Delta f^2$ 倍になるように、時空間周波数領域上のパワースペクトルの分布を決定する。パワースペクトルが $1/\Delta f^2$ に比例するのみではノイズにはならないが、位相をランダムな値とすることによって、 $1/f^2$ ノイズを実現する。このとき、時空間周波数領域上でのスペクトル $U(f_x, f_y, f_t)$ は次式で表される。

$$U(f_x, f_y, f_t) = \frac{1}{|f_t - f_c(f_x, f_y)|} F_m(f_x, f_y, f_t) e^{j\phi(f_t)} \quad (10)$$

ここで、 $f_c(f_x, f_y)$ はマッピングされた面上の (f_x, f_y) における時間周波数であり、式 (7), (8), (9) における f_t^{3D} と同値である。また、 $F_m(f_x, f_y, f_t)$ は入力画像の二次元フーリエスペクトル、 $\phi(f_t)$ は $0 \leq \phi < 2\pi$ の値をとる擬似乱数である。

4. 実験

提案したマッピング手法を比較するため、動画生成実験を行い揺らめきの評価を行った。

4.1 対象とした画像

図 7 に、動画の元となった画像を示す。画像サイズはすべて 512×512 である。また、揺らめかせる領域は、Rowan については文献 [8] の自動推定、Poplar, Fire, Water surface については手作業により設定した。

時空間周波数領域でのマッピングとして、3.2.1 項で述

べた平面へのマッピングを用いた場合, 3.2.2 項で述べた, f_t 軸負方向に凸の曲面へのマッピングを用いた場合, f_t 軸負方向に凸の曲面へのマッピングに $1/f^2$ ノイズを加えた場合の 3 通りを考え, それぞれの場合について動画を生成した. 揺らめきの周期 T は, すべて 64 フレームとし, 式 (7), (8) における f_0 はいずれも 4 とした. 動画のフレームレートはすべて 8fps とした. これにより, いずれの動画も最も強い成分を持つ時間周波数は 0.5Hz であり, 基本的に 0.5Hz で振動しているように見える.

4.2 揺らめき動画生成例

動画生成結果の例として, 図 8-図 11 に, 提案手法を用いて生成した動画のフレームの一部を拡大して示す. この動画において, 時空間周波数領域でのマッピングは, 3.2.2 項で述べた, f_t 軸負方向に凸の曲面へのマッピングを用いた. 各図の (a) はある時刻でのフレームであり, 各図 (b) は (a) の 0.5 秒後, 各図 (c) は (b) の 0.5 秒後のフレームである. 各図 (d) は (a) および (b) の差分画像であり, 各図 (e) は (b) および (c) の差分画像である. 差分画像は視認性を高めるため, ヒストグラム拡張によりコントラストを強調している. これらより, 写真に動きが付加されていることが確認できる.

また, 動画生成の処理時間について, 画像の入力から動画の 1 フレーム目を出力するまでにかかった時間は, すべて 1.5 秒から 2 秒程度であった. なお, 動画生成の際に使用したコンピュータは, CPU が Intel®Core™i7-2600K CPU @ 3.40GHz, メモリが 8.00GB RAM である.

4.3 考察

提案手法を用いて動画を生成した結果について考察を行う.

時空間周波数領域において平面へのマッピングを用いて生成した動画は, いずれも従来手法 [4] を用いた単振動の揺らめき動画と同等であることが確認できた.

また, f_t 軸負方向に凸の曲面へのマッピングを用いて生成した動画では, 一見して単振動の揺らめきと差異がないように感じられるが, 注意深く観察すると, 細かい構造の部分の動きが時間とともに変化していることがわかる. これは, 空間周波数が高いほど時間周波数が高くなるようマッピングを行ったことで, 全体の動きと細かい構造の動きが同期しなくなったためである. これにより, 従来手法の問題点であった画像全体が均一に動くことによる不自然さをある程度緩和できていると考えられる. ただし, Fire に関しては, 単振動の揺らめきと比較してほとんど差異が見られなかった. これは, Fire の画像は高周波成分があまり含まれておらず, 曲面へのマッピングを行っても f_0 以外の時間周波数成分がほとんど現れなかったためであると考えられる.

さらに, f_t 軸負方向に凸の曲面へのマッピングに $1/f^2$ ノイズを加えて生成した動画では, ノイズによって揺らめきが時間変化し, 空間的にも時間的にも単調さを低減できていると考えられる.

5. おわりに

本稿では, 空間的な変化から時間的な変化の情報を得られるのではないかと考えるに基づき, より自然な揺らめきを生成する手法を提案した. 提案手法では, 空間周波数領域に時間周波数軸を加えた三次元周波数領域に成分を適切に配置し, 三次元フーリエ逆変換を行うことによって揺らめき動画の情報を得る. 動画生成実験によって, 従来手法と同等のものを含む, 単調さを低減した揺らめき動画を生成できることを確認した.

今後の課題としては, 主観評価実験を実施して提案手法の有効性を裏付けることが挙げられる. また, 構造の違いによって動き方が変わるようになったが, 画像全体が同時に動く問題は未だ残っており, 領域分割等によって領域ごとに揺らめき方を変えることも求められる.

参考文献

- [1] Chuang, Y.-Y., Goldman, D. B., Zheng, K. C., Curless, B., Salesin, D. H. and Szeliski, R.: Animating Pictures with Stochastic Motion Textures, *ACM Transactions on Graphics*, Vol. 24, No. 3, pp. 853-860 (2005).
- [2] 青木政勝, 新谷幹夫, 筒口 拳: 力学シミュレーションによる 2 次元アニメーション自動生成, 電子情報通信学会論文誌, Vol. J84-D-II, No. 9, pp. 2040-2047 (2001).
- [3] 境野英朋: 流体モデルに基づいた 1 枚の 2 次元画像からの 3 次元的な動画生成法, 映像情報メディア学会誌, Vol. 64, No. 4, pp. 551-562 (2010).
- [4] 森 智美, 坂本雄児: 写真に揺らめき表現を付加した動画生成に関する研究, 映像情報メディア学会技術報告, Vol. 32, No. 46, pp. 39-42 (2008).
- [5] 山本 薫, 坂本雄児: 写真に自然な揺らめき表現を付加した動画生成法, 情報処理学会研究報告, Vol. 2012, No. 7, pp. 1-6 (2012).
- [6] Blinn, J. F.: Simulation of wrinkled surfaces, *Proceedings of the 5th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '78, ACM, pp. 286-292 (1978).
- [7] H.-O. パイトゲン, D. ザウペ (編): フラクタル・イメージ: 理論とプログラミング, シュプリンガー・フェアラーク東京 (1990). 山口 昌哉 (監訳).
- [8] 山本 薫, 坂本雄児: 二枚の写真からの揺らめき動画生成のための対象領域およびパラメータの自動推定法, 第 22 回インテリジェント・システム・シンポジウム講演論文集 (2012).

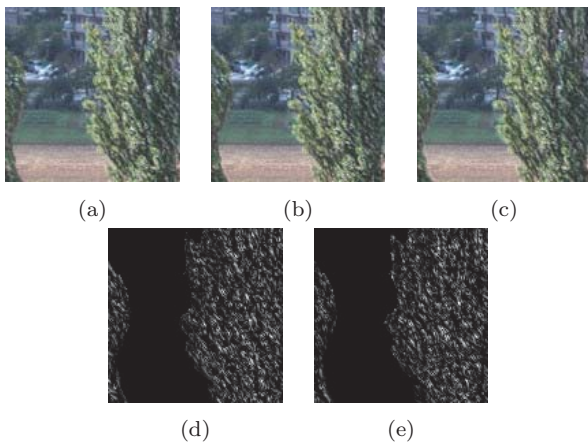


図 8 Poplar に提案手法を適用して生成した動画のフレームおよび差分画像

Fig. 8 Frames in generated animation of image Poplar with proposed method and difference of them.

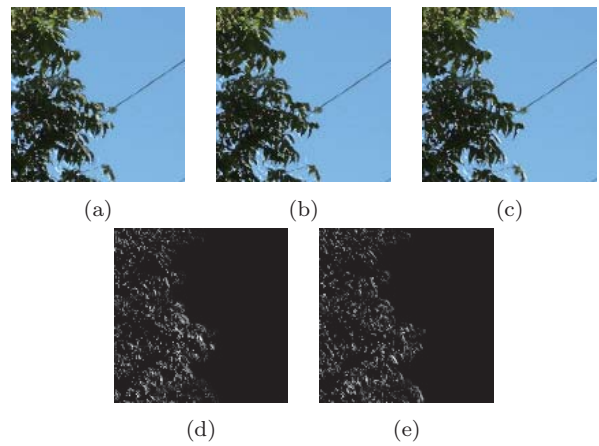


図 9 Rowan に提案手法を適用して生成した動画のフレームおよび差分画像

Fig. 9 Frames in generated animation of image Rowan with proposed method and difference of them.

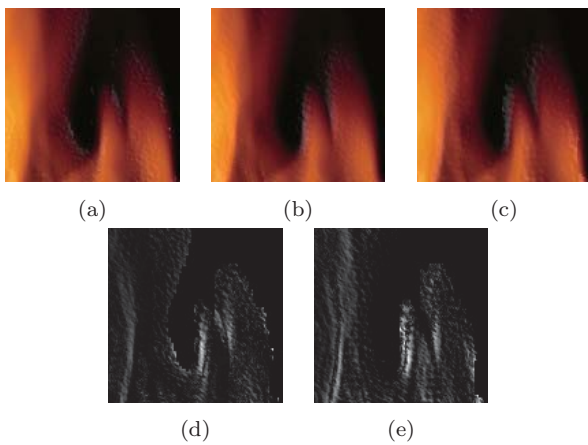


図 10 Fire に提案手法を適用して生成した動画のフレームおよび差分画像

Fig. 10 Frames in generated animation of image Fire with proposed method and difference of them.

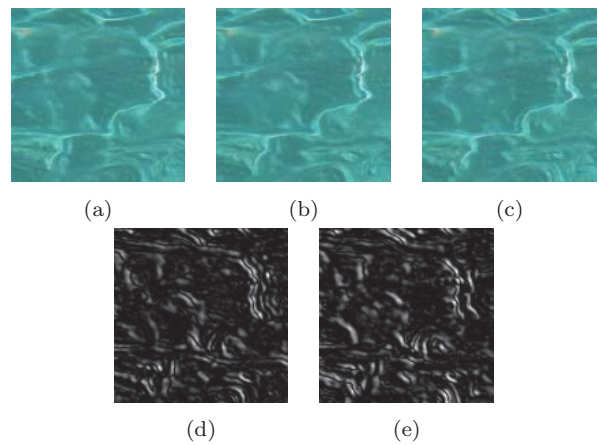


図 11 Water surface に提案手法を適用して生成した動画のフレームおよび差分画像

Fig. 11 Frames in generated animation of image Water surface with proposed method and difference of them.