

写真に自然な揺らめき表現を付加した動画生成法

山本 薫^{1,a)} 坂本 雄児^{1,b)}

概要：一枚の静止画から動画を自動生成する手法が多く提案されている。その手法の一つとして、一枚の写真に揺らめくような動きを与えて動画を生成する手法がある。先行研究における揺らめきは、パラメータが時間変化しない単振動によるものであり、動きが単調となる問題があった。そこで本研究では、自然界で多く見られる $1/f^\beta$ ノイズを用いて、複雑かつ自然な揺らめきを生成する手法を提案する。また主観評価実験を行い、提案手法の有効性について従来手法との比較により確認した。

Generation Method of Animation from Pictures with Natural Flicker

Abstract: Some methods to create animation automatically from one picture have been proposed. There is a method that gives a flicker to a picture to make animation. However, the flicker generated by the simple harmonic motion causes monotonous animation. In this study, we propose the method of creating complex and natural flicker by using $1/f^\beta$ noise seen often in nature. We carried out the experiment that compares the proposed method to the conventional method by the subjective appraisal, and confirmed the adequacy of the proposed method.

1. はじめに

高性能なデジタルカメラやカメラを搭載した携帯機器の普及により、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えた。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枚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ことはない。そのため、写真から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と考えられる。

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、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、ユーザにとって手間がかかる部分が多く、また計算時間や多くの計算機リソースが必要になることから、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、ユーザの手間を軽減している。また、この手法でも、広い対象に対して、比較的良好な揺らめきを与えられることが示された。さらに、必要とされる計算機リソースも少なく、フォトフレームなどへの移植も期待される。

一方で、先行研究 [4] では、動きの計算に単振動の式を用いているため、生成される揺らめきが周期的で単調となることが問題であった。そこで本研究では、より自然な揺らめきを生成することを目的とする。提案手法では、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、揺らめきにランダム性を与える。また、従来手法と比較する主観評価実験を行い、結果について考察した。

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

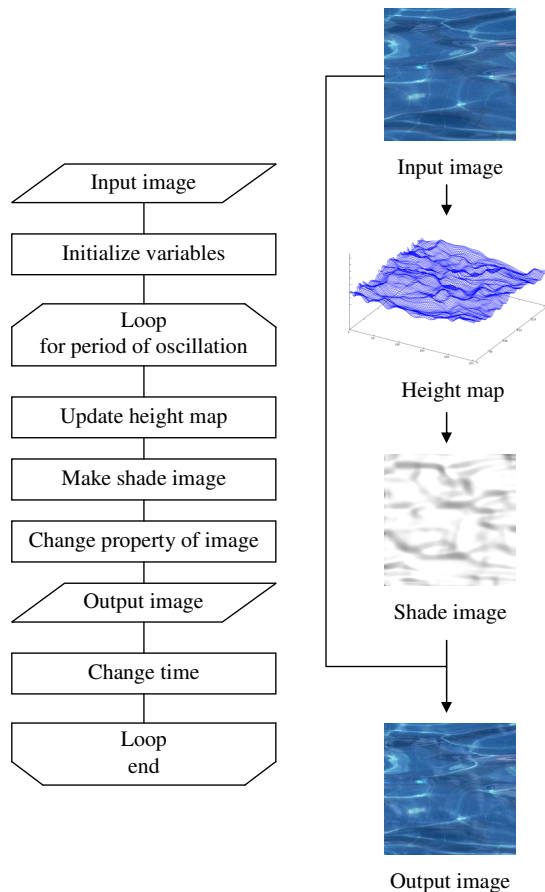


図 1 先行研究の概要

Fig. 1 Overview of preceding study.

2. 揺らめきの原理

2.1 先行研究の概要

先行研究 [4] の手法の流れを図 1 に示す。まず入力画像の座標 (i, j) における画素の輝度値を、高さ H_{ij} 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_{ij} 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_{ij} から画像平面に対してバンプマッピング [5] 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、作成された陰影画像をもとに入力画像の輝度値を変更し、新しく画像を生成・出力する。

高さ H_{ij} は、次式で表される単振動により時間変化する。

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

ここで、 $H_{ij}(t)$ は座標 (i, j) 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 $H_{ij}(0)$ は、入力画像の輝度値で与えられる。高さ $H_{ij}(t)$ は、 $0 \leq H_{ij}(t) \leq 255$ 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 $H_{ij}(t)$ は $C - A \leq H_{ij}(t) \leq C + A$ の範囲で振動する。 A および C の与え方は 2 種類あり、一方は $A = 127.5, C = 127.5$ で固定する「全域振動」、もう一方は近傍の画素の輝度値の最大値と最小値から決定する「近傍考慮」である。この A および C の決定法はユーザが決定する。

ω は角周波数である。一回振動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_{ij} は初期位相である。式 (1) より、時刻 $t = 0$ のときの高さ $H_{ij}(0)$ は次式で表される。

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

よって、初期位相 α_{ij} は次式で求められる。

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

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

高さ $H_{ij}(t)$ の値の時間変化によって陰影画像も変化し、生成される画像も変化するため、生成される画像を次々に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。

2.2 先行研究の問題点

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る。

3. 提案手法

3.1 $1/f^\beta$ ノイズ

自然界の複雑な形状を表す概念として、フラクタルがある [6]。フラクタルは一部を拡大しても全体と複雑さが変わらないという自己相似性を持ち、コンピュータグラフィックスの分野で海岸線や山などの形状を記述する方式として知られている。

自然界では空間的な変化に限らず、時間的に変化するノイズに関してもフラクタル性が観測されることが多い。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、 $1/f^\beta$ の特性をもつ。このノイズは、 β が小さいほどランダム性が強く、 β が大きいほど相関が強い。 $\beta = 0$ のノイズはスペクトル密度が全周波数について一定であり、ホワイトノイズと呼ばれる。 $\beta = 1$ のノイズは $1/f$ ノイズ、 $\beta = 2$ のノイズは $1/f^2$ ノイズ、またはブラウンノイズと呼ばれる。これらのノイズのうち、 $1/f$ ノイズとブラウンノイズによって作られる

グラフはフラクタル曲線となる。

上述のノイズは自然界の様々な場面で観測され、特に $1/f$ ノイズについては、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、自然界で最も多く見られるタイプのノイズであるとされる [6]。

3.2 $1/f^\beta$ ノイズを用いた複雑な動き

本手法では、3.1 節で示したホワイトノイズ、 $1/f$ ノイズおよびブラウンノイズを式 (1) へ適用することによって、複雑な揺らめきを生成する。式 (1) において、振動の様子を決めるパラメータは振幅 A 、角周波数 ω 、初期位相 α であり、これらをノイズによって時間変化させることで、複雑な揺らめきを生成することが可能である。ただし、位相の時間微分が周波数であることから、初期位相 α の時間変化は角周波数 ω の変動の一部と考えることができるため、本手法では、ノイズによって時間変化させるパラメータを振幅 A と角周波数 ω の二つとする。

3.2.1 振幅へのノイズの適用

$1/f$ ノイズを用いることで作成される振幅は以下の式で表される。

$$A(t) = A_0 \cdot X_A(t) \quad (4)$$

ここで、 A_0 は振幅の基本値、 $X_A(t)$ はノイズである。 $X_A(t)$ については区間 $[0, 1]$ に値が集中するようなノイズを用いるが、 $X_A(t)$ が 0 を下回った場合は 0 に、1 を上回った場合は 1 に、 $X_A(t)$ を修正する。 ω を一定として、式 (4) を用いると、式 (1) で表される高さは、角周波数 ω で $A(t)$ の包絡線をもつ振動として時間変化する。振幅へノイズを適用することによって、揺れの大きさが不規則になり、従来手法での揺らめきの単調さが軽減されることが期待できる。

なお、本研究で揺らめきの対象としている木や水などには慣性質量があり、外力が入力された場合、一定時間、振動や移動を続けることになる。よって、式 (4) の $X_A(t)$ は厳密にはノイズそのものではなく、次式で表される畳み込みの結果である。

$$X_A(t) = \sum_{i=0}^n N(t-i)P(i) \quad (5)$$

ただし、 $X_A(t)$ が区間 $[0, 1]$ を大きく外れないよう、式 (5) の右辺には適当な係数をかける。ここで、 n は応答の有効時間である。 $N(t)$ は時刻 t でのノイズ出力であり、ホワイトノイズ、 $1/f$ ノイズ、ブラウンノイズのいずれかを用いる。 $P(t)$ は減衰関数であり、時定数 τ を用いて次式で表される。

$$P(t) = \exp\left(-\frac{t}{\tau}\right) \quad (6)$$

式 (5) より、 $P(t)$ は有限インパルス応答 (FIR: Finite impulse response) フィルタのインパルス応答であり、 n は

FIR フィルタのフィルタ次数であることがわかる。インパルスが入力された場合、式 (4)、(5)、(6) より、高さ $H(t)$ は減衰振動を行う。

3.2.2 周波数へのノイズの適用

振幅の場合と同様、角周波数へのノイズの適用は、次式で表される。

$$\omega(t) = \omega_0 \cdot X_f(t) \quad (7)$$

ただし、 ω_0 は角周波数の基本値、 $X_f(t)$ はノイズであり、 $0 \leq X_f(t) \leq 2$ を満たすものとする。これは上限が +1 オクターブであることを意味する。角周波数へノイズを適用することによって、振動の周期が変動するため、振幅の場合と同様に、従来手法での揺らめきの単調さが軽減されることが期待できる。

3.2.3 振幅および周波数へのノイズの適用

振幅 A および角周波数 ω のいずれにもノイズを適用した場合、揺れの大きさおよび周期がともに変化し、より自然な揺らめきを生成できる可能性がある。そのため、本稿では振幅 A および角周波数 ω へ同時にノイズを適用する場合も考える。振幅 A および角周波数 ω のいずれにもノイズを適用した場合、式 (4)、(7) より、式 (1) は次式に示す形となる。

$$H_{ij}(t) = A_0 X_A(t) \cdot \sin(\omega_0 X_f(t)t + \alpha_{ij}) + C \quad (8)$$

3.3 $1/f^\beta$ ノイズの生成方法

3.1 節で述べたように、 $1/f^\beta$ ノイズにはホワイトノイズやブラウンノイズ、 $1/f$ ノイズがある。本稿では、どのノイズ生成方法が適しているのかについても考察するため、幾つかのノイズ生成方法を用いた。以下に、計算機によって $1/f^\beta$ の特性をもつノイズ $N(t)$ を生成する方法を示す。

3.3.1 ホワイトノイズ

中心極限定理より、区間 $[0, A]$ 上に一様分布する乱数 n 個の和を Z_n とおくと、 n が大きいほど、 Z_n の確率分布は平均 $n/2 \cdot A$ 、分散 $n/12 \cdot A^2$ のガウス分布に近づく。この Z_n から平均 $n/2 \cdot A$ を減ずれば、ホワイトノイズであり、同時にガウス性も満たす、ホワイトガウスノイズ $N(t)$ が得られる。

3.3.2 間欠カオスによる $1/f$ ノイズ

$1/f$ ノイズの生成方法の一つとして、間欠カオス [7] がある。間欠カオスは以下の漸化式によって得られる。

$$N(t+1) = \begin{cases} N(t) + uN(t)^z & N(t) < 0.5 \\ 2N(t) - 1 & 0.5 < N(t) < 1 \\ 0 & N(t) > 1 \end{cases} \quad (9)$$

ここで、 $u > 0$ 、 $1 < z < 2$ である。

3.3.3 デジタルフィルタによる $1/f$ ノイズ

$1/f$ の特性の伝達関数をもつデジタルフィルタにホワイトノイズを入力することによっても、 $1/f$ ノイズ $N(t)$



図 2 動画生成に使用した画像

Fig. 2 Images used for making animation.

を生成することができる [8].

3.3.4 ブラウンノイズ

ブラウンノイズは、ホワイトガウスノイズの積分によって得られる。正規分布 $N(0, 1)$ に従うホワイトガウスノイズを $N_w(t)$ とすると、ブラウンノイズ $N(t)$ は次式で表される。

$$N(t) = \sum_{s=0}^t N_w(s) \quad (10)$$

すなわち、現時刻のノイズの値にホワイトガウスノイズを加えることで、次時刻のブラウンノイズの値が得られる。

4. 実験

本手法の有効性を確認するため、実験を行った。適用するノイズの手法や、振幅や周波数のどちらに適用すべきかなどは、対象とする物体画像によって異なると予想されるため、様々な対象に関して適した手法が何であるかについても考察する。実験では、提案手法によって生成された揺らめき動画について、従来手法との比較による主観評価を行った。

4.1 対象とした画像

図 2 に、動画の元となった 4 種類の画像を示す。画像サイズは Poplar が 512×512 , Rowan, River, Waterfall が 640×480 である。4 種類の画像の物体はそれぞれ異なる物理現象によって動くため、適用するノイズによる影響も異なると予想される。Poplar および Rowan はともに木の画像であるが、Poplar が遠距離からの撮影画像であるのに対し、Rowan は近距離からの撮影画像であり、本来の動きの様子は異なる。生成される揺らめきの振動の周期の基本値は、すべて 1 回振動あたり 20 フレームとした。すなわち 1 フレームあたり 0.05 回振動であり、後述するようにフ

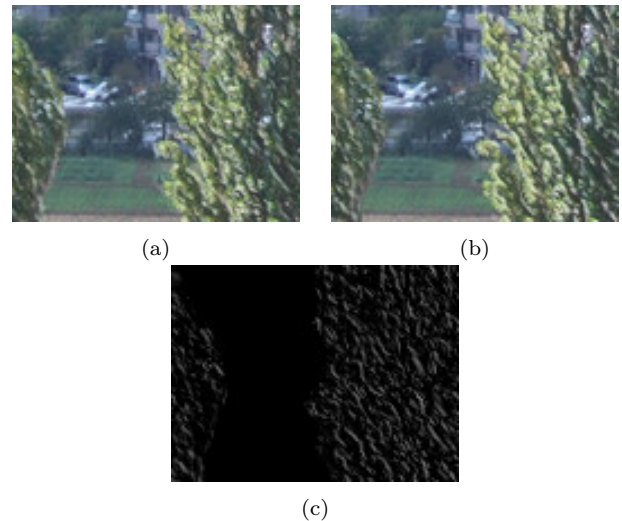


図 3 $1/f$ ノイズを適用して生成した動画のフレームおよび差分画像
Fig. 3 Frames in generated animation of image Poplar with $1/f$ noise and difference of them.

レームレートが約 12fps であることから、生成する動画の角周波数の基本値 ω_0 は、すべて $1.2\pi[\text{rad/s}]$ である。振幅の基本値 A_0 については、Poplar, Rowan が 127.5, River, Waterfall が近傍考慮であり座標によって異なる値を持つ。また、揺らめかせる領域は、Rowan, River については文献 [9] の自動推定、Poplar, Waterfall については手作業により設定した。

本手法では、揺らめき動画は、ファイルへ出力せず、プログラム上で逐次計算して画像を画面に出力し続けることによって実現している。動画生成プログラムはプログラミング言語 C# で実装された GUI アプリケーションである。生成された揺らめき動画のフレームレートはすべて約 12fps であった。なお、揺らめき動画生成の際に使用した PC は、CPU が Intel®Core™i7-2600K CPU @ 3.40GHz, メモリが 8.00GB RAM である。

4.2 揺らめきの例

動画生成結果の例として、図 3 に、Poplar に $1/f$ ノイズを適用して生成した動画の、フレームの一部を拡大して示す。図 3(a) はある時刻でのフレームであり、図 3(b) は (a) の約 0.5 秒後のフレームである。図 3(c) は (a) および (b) の差分画像であり、写真に動きが付加されていることが確認できる。

実験で生成した揺らめき動画における、振幅の基本値 $A_0 = 127.5$ の場合の振幅 $A(t)$ の時間変化の例を図 4 に示す。図 4(a) がホワイトノイズを適用した場合、(b) が $1/f$ ノイズを適用した場合、(c) が $1/f^2$ ノイズを適用した場合、(d) がノイズを適用しない場合である。(a) から (d) への順に時間方向の相関が強くなり、生成される揺らめきが単調となる。

また、図 5 に、振幅を一定にし、角周波数 ω に $1/f$ ノ

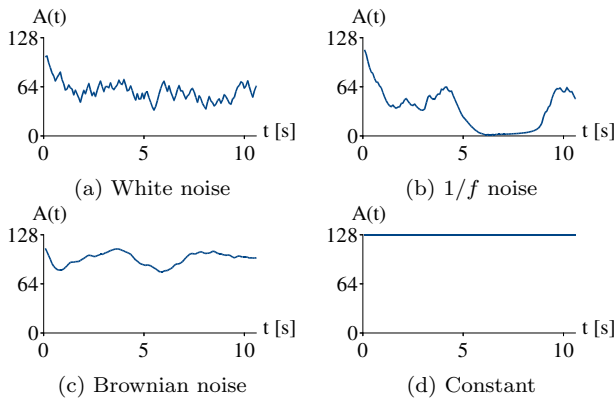


図 4 振幅 $A(t)$ の時間変化 ($A_0 = 127.5$ の場合)

Fig. 4 Amplitude $A(t)$ in time domain (in case $A_0 = 127.5$).

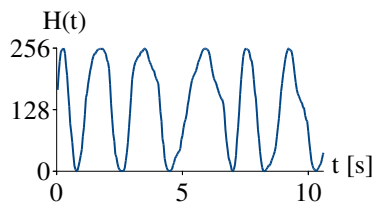


図 5 周波数に $1/f$ ノイズを適用した場合の高さ $H(t)$ の時間変化

Fig. 5 Height $H(t)$ using frequency with $1/f$ noise.

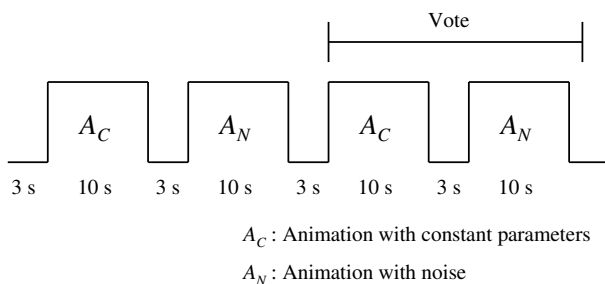


図 6 動画の提示方法

Fig. 6 Presentation of animation.

イズを適用した例を示す。図より、時間経過とともに周期が変化していることがわかる。なお、角周波数 $\omega(t)$ の時間変化のグラフは振幅 $A(t)$ の時間変化と形状が等しくなるため、ここでは省略する。

4.3 主観評価実験

生成された揺らめき動画の適切さについて、8 人の被験者による主観評価実験を行った。二重刺激劣化尺度法 [10] を参考にして、図 6 に示すように、従来手法と提案手法の動画の比較再生を行った。ノイズを適用せずに生成した動画、ノイズを適用して生成した動画をそれぞれ 10 秒間、交互に 2 回ずつ再生し、それぞれの動画の間には 3 秒間のインターバルを設けた。各提案手法による動画の揺らめき方が、従来手法による動画と比較して自然かどうかを 5 点満点で点数をつけてもらい、平均を取ったものを評価結果とした (平均オピニオン評点)。表 1 に、主観評価における評点を示す。この評価方法では、従来手法との比較によ

表 1 主観評価実験の評点

Table 1 Score of subjective evaluations.

5	提案法の方が極めて自然である
4	どちらかといえば提案法の方が自然である
3	どちらともいえない (変わらない)
2	どちらかといえば提案法の方が不自然である
1	提案法の方が極めて不自然である

て、従来手法と同等とみなせる場合は評価が 3 となる。

4.3.1 ノイズの手法による違い

同一の対象物体画像を入力し、適用するノイズを変えた場合の評価結果について考察を行い、ノイズの違いによる効果を検証する。

入力画像は Poplar とした。ノイズの適用に関しては、振幅、周波数それぞれに対して、ホワイトノイズ、間欠カオスによる $1/f$ ノイズ、デジタルフィルタによる $1/f$ ノイズ [8]、ブラウンノイズのいずれかを適用した場合、またはノイズを適用しなかった場合の、 $5 \times 5 = 25$ 通りについて考える。ただし振幅、周波数ともにノイズを適用しなかった場合については、従来手法と同一となるため、除外する。よって、24 通りの場合について評価を行った。

主観評価の結果を表 2 に示す。表 2 では、行が振幅へ適用するノイズの指定、列が周波数へ適用するノイズの指定に対応する。評価結果より、振幅、周波数にブラウンノイズを適用した場合が最も評価が高く、従来手法よりもやや自然な揺らめきであると評価されている。一方、振幅にデジタルフィルタによる $1/f$ ノイズ、周波数に間欠カオスによる $1/f$ ノイズを適用した場合が最も評価が低く、従来手法よりも不自然な揺らめきであると評価されている。

全体的に、周波数にノイズを適用すると評価が低くなる傾向にある。これは、揺らめきの周期が急に変化すること、なめらかに揺らめく従来手法と比較して不自然であるという印象を与えたことが原因であると推測される。これに対して、ブラウンノイズは比較的穏やかに変化するノイズであるため、周波数にブラウンノイズを適用した場合はより自然であると評価されたと推測される。

4.3.2 対象物体画像による違い

同じ「揺らめく」物体であっても、木の揺れと水面の揺らめきは異なる物理現象であることから、対象物体ごとにノイズの適用による効果は異なると推定される。そのため、図 2 のそれぞれの画像にノイズを適用した場合の評価結果について考察し、対象物体画像の違いによる効果を検証する。

主観評価の結果として、図 2 の各画像について最も適したノイズおよびその評価点を表 3 に示す。評価結果より、対象物体によってノイズの効果が異なることが確認できる。River についてはノイズを適用したことで評価が低くなっているが、Waterfall については従来手法と同程度の評

表 2 主観評価結果（ノイズによる違い）

Table 2 Result of subjective evaluation (for noise).

Freq. Amp.	Constant	White noise	1/f by chaos	1/f by filter	Brownian noise
Constant	—	3.0	2.8	2.9	2.9
White noise	3.2	3.0	3.0	2.1	3.0
1/f by chaos	3.1	2.8	2.7	2.2	2.5
1/f by filter	3.2	2.6	2.0	2.4	2.8
Brownian noise	3.3	3.3	3.0	3.2	3.5

表 3 主観評価結果（対象画像による違い）

Table 3 Result of subjective evaluation (for input image).

Image	Noise for amp.	Noise for freq.	Score
Poplar	$1/f^2$	$1/f^2$	3.5
Rowan	1/f by filter	Constant	3.2
River	1/f by filter	Constant	2.6
Waterfall	1/f by filter	Constant	3.2

価となっている。また、Poplar と Rowan については、同じ「風に揺れる木」であっても最適なノイズが異なる結果となった。このことから、自然な揺らめきの生成には、対象物体や撮影距離によって適切なノイズを選択する必要があることが確認できる。

5. おわりに

本稿では、ノイズを用いて複雑かつ自然な揺らめき動画を生成する手法を提案した。提案手法では、揺らめきを与える高さの時間変化式の振幅 A および角周波数 ω に対して、ホワイトノイズ、 $1/f$ ノイズ、およびブラウンノイズを適用した。主観評価実験により、全体としてはノイズの適用によって、従来手法で問題となっていた単調さによる不自然さが改善された。また、振幅 A にブラウンノイズを適用することで最も高い評価が得られることを確認した。さらに、対象物体画像によって最適なノイズは異なり、対象物体や撮影距離によって適切なノイズを選択する必要があることを確認した。

今後は、より多くの対象物体画像に対して本手法を適用することで、様々な対象物体とノイズの関係性を明らかにすること、また、主観評価実験の被験者の数を増やすことによって、評価の精度を上げるについて検討する必要がある。さらに、対象物体によって適切なノイズを選択する必要があることは、先行研究 [4] における同一のルールの適用という利点を損なうため、ノイズを統一することについても今後検討する必要がある。

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, 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 次元動的な動画生成法, 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] Blinn, J. F.: *Simulation of wrinkled surfaces*, ACM SIGGRAPH 78 Proceedings, pp. 286-292 (1978).

[6] Peitgen, H. O. and Saupe, D.: *The Science of Fractal Images*, Springer-Verlag New York (1988). 山口昌哉 (監訳): 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).

[7] 趙華安, 内田大輔: カオス現象における $1/f$ ゆらぎの観察, 電子情報通信学会ソサイエティ大会講演論文集, 基礎・境界, p. 24 (2002).

[8] 細田 隆之: $1/f$ 雑音生成アルゴリズム, 入手先 (<http://www.finetune.jp/~lyuka/technote/pinknoise/>) (2012.11.01).

[9] 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, 1B2-1 (2012).

[10] Subjective video quality assessment methods for multimedia applications, ITU-T Recommendation P. 910 (1999).