Proceedings of the World Congress on Engineering 2014 Vol I,
WCE 2014, July 2 - 4, 2014, London, U.K.

Fluctuation Analysis for Photographs of Tourist Spots and Music Extraction from Photographs

Hidefumi Kawakatsu*

***Abstract*—This study proposes the following two methods applying one-dimensional FFT (fast Fourier transform) algorithm: (1) a method for finding pleasant photographs of local tourist spots, and (2) a method for creating music from these photographs. We define "pleasant photograph" as the photograph containing $1/f$ noise components since it has been suggested that the $1/f$-noise structure in visual art as well as in music can stimulate the perception of pleasant. We analyze 198 photographs published in the book edited by Onomichi city to find the pleasant photographs. The method for the music extraction from the picture is developed taking into account that the sequence of the brightness in the horizontal direction on each row can be decomposed into periodic waves, and then, by regarding each wave obtained as sound waves, successive chords can be created from the image. A method to arrange the music by reflecting features of the image is also provided.**

***Index Terms*—fast Fourier transform, $1/f$ noise, pleasant photograph, music extraction**

I. INTRODUCTION

Local cities in which there are many resources for tourism attempt to increase tourist arrivals. This study provides a way helping to recommend tourist attractions. Mathematics and art works such as music have some vague sort of affinity, but several researchers have shown that the music displays regularities in scaling properties and long-range correlations[1], [2], [3], [4]. They indicated that pleasant music for humans displays a behavior similar to $1/f$ noise. For visual arts, E. Rodriguez et al. [5] have clarified the existence of the factuality and scaling properties in Pollock's drip paintings. They suggested that paintings which contain $1/f$-noise structures can also stimulate the perception of pleasant. They applied two-dimensional DFA method to analyze the gray-scaled images obtained from paintings. The DFA algorithm for two dimensions has simpler structure than the two-dimensional FFT (fast Fourier transform), but its computation time becomes greater than that of the one-dimensional case. Musha[6] has described the method which evaluates paintings by using one-dimensional Fourier transform as follows: (1) computing the power spectral densities of brightness in the horizontal direction in each row, and then (2) obtaining the average values for the spectral densities in each column. In this study, we propose the method for finding pleasant photographs of tourist spots applying the one-dimensional FFT[7] for saving the computational time.

We also develop the method for extracting music considering the fact that the sequence of the brightness in the horizontal direction in each row can be resolved into periodic waves, and then, by regarding each wave obtained as sound waves, successive chords can be created from the image.

Our methods may be available to recommend a sightseeing route which connects the pleasant spots for tourists, and to create the music from photographs of these spots.

Manuscript received March 11, 2014; revised March 31, 2014.
H. Kawakatsu is with Department of Economics & Information Science, Onomichi City University, 1600-2 Hisayamadacho, Onomichi 722-8506 Japan (corresponding author to provide phone: +81-848-22-8312 (ex.617); fax: +81-848-22-5460; e-mail: kawakatsu@onomichi-u.ac.jp).

II. CLASSIFICATION OF DIGITAL PHOTOS

The $1/f^\alpha$ noise is defined in terms of the shape of its power spectral density $P(f)$, where f is the frequency. When the signal follows a scaling law, a power-law behavior for the power spectral density $P(f)$ is observed:

$$P(f) \sim 1/f^\alpha, \tag{1}$$

where α is called the scaling (spectral) exponent. The scaling exponent α is computed as the slope of the plot $\{\log(f)$ versus $\log(P(f))\}$. The sequence, denoted by $u(i)$, can be classified as follows: (a) If $\alpha = 0$, the sequence $u(i)$ is non-correlated (i.e. white noise); (b) If $\alpha = 1$, the correlation of the sequence is the same of $1/f$ noise; and (c) If $\alpha = 2$, the sequence behaves like Brownian motion (random walk)[1], [8].

We define "pleasant photograph" as the photograph containing $1/f$ components and propose a procedure for finding the pleasant photographs using the (one-dimensional) fast Fourier transform (FFT) algorithm. The color images are transformed into gray scale images by using a simple linear combination of the original RGB channels. The size of these images is normalized to M (horizontal) $\times N$ (vertical) pixels, where $M = 2^k$ $(k = 2, 3, \cdots)$ and N is determined so that the aspect ratio of the image can be maintained. Let $G_n(m)$ $(m = 1, 2, \cdots, M, n = 1, 2, \cdots, N, G_n(m) = 0, 1, \cdots, 255)$ be the brightness (gray scale level) at the coordinates (m, n) of the gray scale image. The procedure consists of the following steps:

1) $n = 1$.
2) Set $u(m) = G_n(m)$ $(m = 1, 2, \cdots, M)$, and pass $u(m)$ through a bandpass filter in the range from a Hz to b Hz, where $a = 0.015M$ and $b = 0.35M$. The values of a and b are determined considering the Nyquist frequency (folding frequency) and the distortion of the spectrum.
3) Compute the spectral density $s_n(f)$ of the sequence $u(m)$ using the FFT algorithm. The power spectral density $P_n(f)$ can therefore be obtained by $P_n(f) = s_n^2(f)$.
4) If $n < N$, increment n by one and go to Step 2), otherwise go to Step 5).
5) Compute $\bar{P}(m)$ which is given by $\bar{P}(m) = (1/N)\sum_{n=1}^{N} P_n(m)$ $(m = 1, 2, \cdots, M)$ and the scaling exponent α as the slope of the plot $\{\log(f)$ versus $\log(P(f))\}$ $(f = a, a+1, \cdots, b)$[6].

ISBN: 978-988-19252-7-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

6) Classify the image into one of the following three cases: (a) If $\alpha \le 0.75$, the image is classified into "Random" (or "White"): (b) If $0.75 < \alpha \le 1.25$, the image is classified into "$1/f$ noise" (or "Pink"): (c) If $\alpha > 1.25$, the image is classified into "Brownian" (or "Red").

III. Music Extraction from Images

This section proposes a simple method for extracting music from the image using the FFT algorithm. The image is transformed into the gray scale image $G_n(m)$, and the size of the image is normalized to M (horizontal) $\times N$ (vertical) pixels as mentioned in Section II.

The horizontal axis of the image can be regarded as the time axis by using the following equation.

$$t = m/S_f, \ m = 1, 2, \cdots, M \tag{2}$$

where t represents the time corresponding to m and S_f expresses a sampling frequency. We assign an integer to each musical note of the scale such that middle C (C4) is assigned 37, the note just above (C#4) is 38 and the musical rest is 0. In this case, the frequency F_j (Hz) of each sound can be given by

$$F_j = \begin{cases} 440 \times 2^{\frac{j-46}{12}}, & \text{if } j = 1, 2, \cdots, 96, \\ 0, & \text{if } j = 0. \end{cases} \tag{3}$$

Let H $(\le b - a)$ be the maximum value of the number of sounds (notes) per chord, then the following procedure can be used to extract the music from the gray scale image:

1) Set a suitable value for S_f.
2) $l = 1$.
3) Divide the image into L $(\le N)$ regions from the first row of the image (The size of each region becomes $\lfloor N/L \rfloor$ (vertical) $\times M$ (horizontal)).
4) Compute the average of the power spectral density $\bar{P}_l(m)$ in the lth $(l = 1, 2, \cdots, L)$ region for $m = 1, 2, \cdots, M$, which is given by

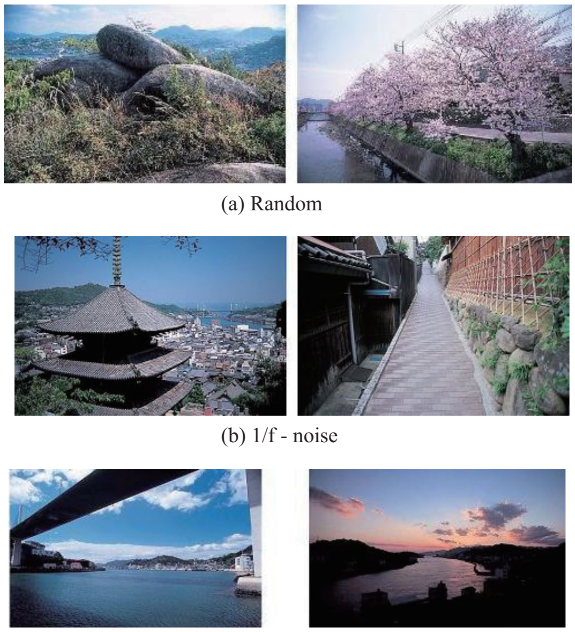
$$\bar{P}_l(m) = \frac{1}{w} \sum_{k=(l-1)w+1}^{l \cdot w} s_k^2(m), \tag{4}$$

where $w = \lfloor N/L \rfloor$ and $s_k(m)$ can be obtained by the procedure of Step 3) in Section II.

5) Sort m in descending order of $\bar{P}_l(m)$, and let m'_i be the value corresponding to the sorted m such that $\bar{P}_l(m'_1) \ge \bar{P}_l(m'_2) \ge \cdots \ge \bar{P}_l(m'_H)$.
6) $i = 1$.
7) Compute the frequency f_i corresponding to m'_i and the absolute value d_j of the difference between f_i and F_j, where $f_i = m'_i \cdot (S_f/M)$ and $d_j = |f_i - F_j|$ $(j = 1, 2, \cdots, 96)$.
8) Find F_k such that $d_k = \min d_j$ and let $q_{i,l} = F_k$ be a note in the chord $\boldsymbol{Q}(l)$ which is generated from lth region, where $\boldsymbol{Q}(l)$ is expressed by Eq. (5) in the next step.
9) If $i < H$, increment i by one and go to Step 7), otherwise set

$$\boldsymbol{Q}(l) = (q_{1,l}, q_{2,l}, \cdots, q_{H,l})^T \tag{5}$$

and go to Step 10).



(a) Random

(b) 1/f - noise

(c) Brownian

Fig. 1. An example of classification

10) If $l < L$, increment l by one and go to Step 4), otherwise go to Step 11).
11) Arrange $\boldsymbol{Q}(l)$ in order from $\boldsymbol{Q}(1)$ to $\boldsymbol{Q}(L)$ and transform them into the musical score.

We refer to the music generated from above procedure as an "original music". A method for arrangement the original music by reflecting features of the image will be discussed in Section V.

IV. Results of Classifying Images

This section analyzes the 198 photographs published in the book edited by Onomichi city[9]. These photographs are classified into (a) "Random", (b) "$1/f$ noise" and (c) "Brownian", using the method proposed in Section II. Figure 1 shows an example of the result of classification, and Table I summarizes the number of images classified into each category. Figure 2 depicts the behavior of the brightness of the part indicated by the broken white lines.

A. Feature of the images in category (a)

Figure 1(a) shows pictures classified into the category of "Random". Many plants with small leafs occupy the surface of the photographs in Fig. 1(a). We have confirmed that other pictures with the similar feature are classified into this category. We can also observe in Fig. 2(a) that the brightness in the horizontal direction oscillates randomly in the part of the image that contains overgrown plants with the small leafs.

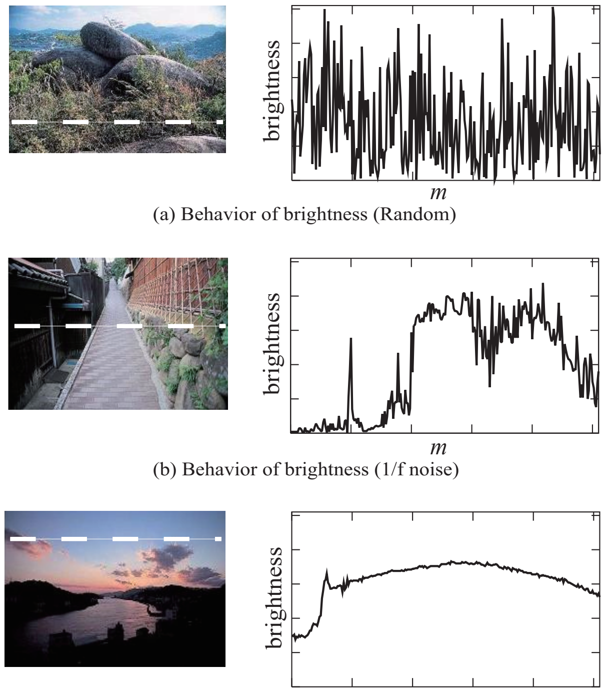
B. Feature of the images in category (b)

Figure 1(b) shows pictures classified into the category of "$1/f$" noise. We can notice, in Fig. 1(b), that the photographs are composed of scenery which artificial and natural objects harmonize well. Figure 2(b) displays the behavior of the brightness in the horizontal direction, which resembles the behavior of $1/f$ noise. There are many old style objects such

ISBN: 978-988-19252-7-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE I
THE NUMBER OF IMAGES CLASSIFIED INTO EACH CATEGORY

Category	Number of images	%
(a) Random	7	3.5
(b) $1/f$ noise	122	61.6
(c) Brownian	69	34.8



(a) Behavior of brightness (Random)

(b) Behavior of brightness (1/f noise)

(c) Behavior of brightness (Brownian)

Fig. 2. Behavior of brightness

as shrines, temples, and slopes with houses in Onomichi. People in this region have well preserved these cultural heritages as well as nature. This may be one of the reasons why most of the photographs (61.6%) are classified into this category as shown in Table I.

C. Feature of the images in category (c)

Figure 1(c) shows the pictures classified into the category of "Brownian". We can observe in Fig. 1(c) that contrast of these pictures is relatively low and flat, and they have large dark area. This signifies that there is little, if any, change in hue. Brownian noise is correlated showing only slow changes, which behaves as (one-dimensional) random walk. E. Rodriguez et al. [5] have indicated that pictures containing Brownian noise create monotonous or boring impression. We can also confirm, in Fig. 2(c), that the brightness in the horizontal direction exhibits Brownian-like motion.

V. ARRANGEMENT OF "ORIGINAL MUSIC"

This section develops a method of arrangement of the "original music" which is created by the procedure proposed in Section III. We consider the case where, as the average value of the brightness in the region l ($l = 1, 2, \cdots, L$) increases, the number of sounds (notes) in the chord $\boldsymbol{Q}(l)$ decreases, tempo increases and the ratio of containing dissonance per chord decreases. By this way, a part of the music extracted from the bright area in the picture can be played lightly, while it is played heavily corresponding to the brightness of the dark area.

We also consider the case where some sounds are interpolated into two successive chords $\boldsymbol{Q}(l)$ and $\boldsymbol{Q}(l+1)$ ($l = 1, 2, \cdots, L-1$) reflecting the scaling (spectral) exponent in the region l.

We here introduce some notation with respect to the region l ($l = 1, 2, \cdots, L$) in the picture for simplicity as listed below:

$\bar{g}_l$**:** the average value of the brightness in the region l.
α_l**:** the scaling (spectral) exponent.
$\tau(l)$**:** the tempo of the part corresponding to the region l.
$\theta(l)$**:** the number of tones sounded simultaneously.
$\lambda(l)$**:** the ratio of containing dissonance.

Let $q_{H,l}$ correspond to a melody part and the other sounds in $\boldsymbol{Q}(l)$ without $q_{H,l}$ be an accompaniment part.

In this case, the procedure for the arrangement of the original music can be proposed as follows:

1) $l = 1$.
2) Compute the average value $\bar{g}_l$ of the brightness and the scaling exponent α_l in the region l.
3) Set the values for $\tau(l)$, $\theta(l)$ and $\lambda(l)$ such that these values are proportional to $\bar{g}_l$, where $\tau(l)$, $\theta(l)$ and $\lambda(l)$ take the values on the interval $[40, 255]$, $[0, 0.1]$ and $[1, H]$, respectively.
4) Adjust the values for the components of $\boldsymbol{Q}(l)$ without $q_{H,l}$ such that the ratio of containing dissonance in $\boldsymbol{Q}(l)$ is less than or equal to $\lambda(l)$.
5) Set $\boldsymbol{S}\left((l-1)(B+1)+1\right) = \boldsymbol{Q}(l)$, where B denotes the number of interpolating sounds.
6) If $l = 1$, go to Step 9), otherwise go to to Step 7).
7) For the melody part, set the value of $s_{H,(l-2)(B+1)+j}$ ($j = 2, \cdots, B-1$) corresponding to the value for α_l as follows:
   a) $\alpha_l \le 0.75$.
      Set
      $$s_{H,(l-2)(B+1)+j} = r^a_{j-1}, \tag{6}$$
      where r^a_k ($k = 1, 2, \cdots, B-2$) is a positive integer, which is given as the uniform random number on the interval $[q_{H,l-1}, q_{Hl}]$.
   b) $0.75 < \alpha_{\le} 1.25$.
      Set
      $$s_{H,(l-2)(B+1)+j} = r^b_{j-1}, \tag{7}$$
      where r^b_k is a positive integer, which is given as the $1/f$ random number with mean $(q_{H,l-1} + q_{Hl})/s$.
   c) $\alpha_l > 1.25$.
      Set
      $$s_{H,(l-2)(B+1)+j} = \begin{cases} s_{H,(l-2)(B+1)+j-1} - 1, & p_{j-1} < 0.5 \\ s_{H,(l-2)(B+1)+j+1} + 1, & p_{j-1} \ge 0.5 \end{cases} \tag{8}$$
      where p_k expresses the uniform random variable on the interval $(0, 1)$.
8) For the accompaniment part, set $s_{ij} = qil$ ($i = 1, 2, \cdots, H-1$, $j = (l-1)(B+1)+2, (l-1)(B+1)+3, \cdots, (B+1)l$).

ISBN: 978-988-19252-7-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

9) If $l < L$, increment l by one and go to Step 2), otherwise go to Step 10).
10) Arrange $\boldsymbol{S}(l)$ in order from $\boldsymbol{S}(1)$ to $\boldsymbol{S}(L)$ and transform them into the musical score.

VI. Conclusion

This study has developed the following two methods applying one-dimensional FFT (fast Fourier transform) algorithm: (1) a method for finding pleasant photographs of local tourist spots, and (2) a method for creating music from these photographs. We have defined "pleasant photograph" as the photograph containing $1/f$ noise components since the recent studies (such as E. Rodriguez et al. [5]) had suggested that paintings which contain $1/f$-noise structures stimulate the perception of pleasant. We have analyzed 198 photographs published in the book edited by Onomichi city[9]. These photographs are classified into (a) Random, (b) $1/f$ noise and (c) Brownian, using the method proposed in Section II. When we had shown the results obtained from the analysis using the 198 pictures at a seminar for the general public held in Onomichi, and asked for the participants' opinions, they mostly agreed with our observation. This indicates that our method can efficiently be applied to classify the pictures of tourist areas where scenery with old houses, shrines and temples. In contrast, the comment on music created from these pictures was that music should be created enough to evoke the scene in the pictures. The pictures used in this study are transformed into gray scale images. In many cases, the pictures of the tourist spots where countryside scenery are preserved consist of mountains, plants, coasts, temples and so on. The color of mountains, coasts and temples are usually green, blue and brown, respectively. Various melodies can be created by changing of musical key depending on the color of the objects in the picture. Taking account of such factors is an interesting extension.

References

[1] R. F. Voss and J. Clarke, ""$1/f$ noise" in music: Music from $1/f$ noise," *J. Acoust. Soc. Am.*, vol. 63, no. 1, pp. 258–263, 1978.
[2] G. R. Jafari, P. Pedram, and L. Hedayatifar, "Long-range correlation and multifractality in bach's inventions pitches," *Journal of Statistical Mechanics: Theory and Experiment*, vol. P04012, 2007.
[3] L. Dagduga *et al.*, "Correlations in a mozart's music score (k-73x) with palindromic and upside-down structure," *Physica A*, vol. 383, no. 2, pp. 570–584, 2007.
[4] W. Ro and Y. Kwon, "$1/f$ noise analysis of songs in various genre of music," *Chaos Solitons & Fractals*, vol. 42, pp. 2305–2311, 2009.
[5] E. Rodriguez *et al.*, "$1/f$ noise structures in pollocks' drip paintings," *Physica A*, vol. 387, pp. 281–295, 2008.
[6] T. Musha, *Yuragi no Sekai [The World of Fluctuations]*. Kodansha, 1980.
[7] P. Bloomfield, *Fourier Analysis of Time Series: An Introduction*. John Wiley & Sons, 2000.
[8] E. Rodriguez, J. C. Echeverria, and J. Alvarez-Ramirez, "$1/f^\alpha$ fractal noise generation from Grünwald-Letnikov formula," *Chaos, Solitons & Fractals*, vol. 39, no. 2, pp. 882–888, 2009.
[9] Onomichi city, *Onomichi Sansaku 204 Sen [Selected 204 exquisite spots in Onomichi]*. Toppan Printing, 2004.