

Identify a Specified Fish species by the Co-occurrence and Confusion Matrix

Ziyue Yan^{1,2}, Lifeng Zhang¹, Xuelong Hu², Seiichi Serikawa¹

¹Department of Electrical Engineering and Electronics, Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu 804-8550, Japan

²School of Information Engineering, Yangzhou University, 198 Huayangxi Road, Hanjiang, Yangzhou 225127, China

*Corresponding Author: yanziyue07@gmail.com

Abstract

Nowadays, invasive species threaten native species has become a global problem. Invasive species might be carrying pathogenic microorganisms, reduce biological species and even threat to human health. Therefore, in this study, we proposed a method of co-occurrence matrix to texture analysis of three species of fish. We catch the body pattern, and make a judgment based on confusion matrix. Simulation results show that three species of fish can be classified from each other reasonable.

Keywords: invasive species, co-occurrence matrix, confusion matrix.

1. Introduction

In recent years, many foreign creatures destroy the ecosystem of native species, and this becomes a serious worldwide problem now. Foreign creature was called specialized invasive alien species, and Japanese government issued the "invasive alien species act" in 2004. In Japan, there are 14 kinds of alien species about fish. Invasive species can destroy the ecological environment. When a stranger creature allowed into ecosystem, native species would loss of living space and food, stranger creatures may be release chemicals and attack native species. Ultimately, native species will become extinct if they are not be protected timely. In recent years, the main way of catching stranger creature is manual control. But it will cost much money and manpower. Therefore, it is necessary to get rid to them at the beginning of breeding.

Natural texture has sophisticated structure. Now there is no unified standard to describe its characteristics. How to

characterize the natural texture and classification is an important direction in the field of computer vision research. The way of texture characteristics can be divided into statistics method and structural method. Statistical method is more suitable for processing natural texture⁽¹⁾. In statistical method, the co-occurrence matrix is classical way to describe the texture. The generated parameters can describe the statistical characteristics of the texture in many aspects.

In this study, we use co-occurrence matrix for texture analysis to three kinds of fishes, crap, sunfish and black bass. We get a part of fish body surface image, make a co-occurrence matrix and then calculate four kinds of characteristics form co-occurrence matrix. Finally we use confusion matrix to evaluate the simulation result of the proposed method. From the evaluation value, it is proved that the proposed method can identify the fish species effectively.

In section 2, co-occurrence matrix is introduced. In section 3, confusion matrix method is selected for this study. Simulations for fish texture analysis with two methods are revealed in section 3.

2. Co-occurrence matrix

2.1 Co-occurrence matrix definition

In 1973, Haralick proposed co-occurrence matrix to describe the texture feature⁽²⁾. Co-occurrence matrix is defined over an image to be the distribution of co-occurring values at a given offset.

$$G_d(i, j)(i, j = 0, 1, 2, \dots, k-1) \quad (1)$$

Where i and j are the image intensity values of the image, d is the relationship between pixels of location. When two pixel locations d is selected, we can get matrix with $d^{(3)}$.

$$G_d = \begin{bmatrix} G_d(0,0) & G_d(0,1) & \dots & G_d(0,k-1) \\ G_d(1,0) & G_d(1,1) & \dots & G_d(1,k-1) \\ \dots & \dots & \dots & \dots \\ G_d(k-1,0) & G_d(k-1,1) & \dots & G_d(k-1,k-1) \end{bmatrix} \quad (2)$$

2.2 Co-occurrence matrix characteristic parameters

Haralick defines 14 kinds of feature parameters for texture analysis of co-occurrence matrix⁽⁴⁾. Ulaby's study found that based on 14 kinds texture feature, just 4 characteristic parameters are irrelevant⁽⁵⁾. These values are easy to calculate and they deserve a more accurate classification.

Entropy: the amount of information which is the image measurement.

$$ENT = -\sum_{i=1}^k \sum_{j=1}^k G(i, j) \log G(i, j) \quad (3)$$

Energy: the sum of the squares of the gray level co-occurrence matrix elements, so also known as energy, reflects the image grey distribution uniformity degree of thickness and texture.

$$ASM = \sum_{i=1}^k \sum_{j=1}^k (G(i, j))^2 \quad (4)$$

Contrast: reflects the image of the degree of clarity and texture grooving depth. If the value of contrast is small, the picture is fuzzy.

$$CON = \sum_{n=0}^{k-1} n^2 \left\{ \sum_{|i-j|=n} G(i, j) \right\} \quad (5)$$

Relevancy: it is used to measure the similarity degree of the elements of gray level co-occurrence matrix on the row or column, therefore, the relative value reflects the local gray correlation of image size.⁽⁶⁻⁷⁾

$$COR = \sum_{i=1}^k \sum_{j=1}^k \frac{i \cdot j \cdot G(i, j) - \mu_i \mu_j}{s_i s_j} \quad (6)$$

where

$$\mu_i = \sum_{j=1}^k i \cdot G(i, j)$$

$$\mu_j = \sum_{i=1}^k j \cdot G(i, j)$$

$$s_i^2 = \sum_{j=1}^k G(i, j) (i - \mu_i)^2$$

$$s_j^2 = \sum_{i=1}^k G(i, j) (j - \mu_j)^2$$

2.3 Co-occurrence matrix characteristic values

Table 1~3 show the four characteristic value of co-occurrence matrix for each fish. We capture one part of each fish. And we can find different fishes have different values.

Table 1: Statistical results of sunfish body texture

Name	ENT	ASM	COR	CON
Sunfish1	6.610716	0.216256	0.72475	0.30549
Sunfish2	6.682494	0.221526	0.758724	0.287059
Sunfish3	6.567916	0.222732	0.692758	0.34549
Sunfish4	5.958765	0.347993	0.672726	0.196863
Sunfish5	6.613791	0.215509	0.74275	0.320784
Sunfish6	6.464438	0.261066	0.680734	0.326667
Sunfish7	6.884852	0.148534	0.734582	0.417647
Sunfish8	6.662622	0.241086	0.786118	0.249804



Fig 1: Body texture pattern of sunfish

Table 2: Statistical results of sunfish body texture

Name	ENT	ASM	COR	CON
Crap1	7.179772	0.12919	0.862893	0.349412
Crap2	7.231622	0.0834	0.649161	0.900784
Crap3	7.07714	0.113908	0.676716	0.736078
Crap4	7.157248	0.100221	0.657833	0.921569
Crap5	6.513132	0.220037	0.619326	0.389804
Crap6	6.661643	0.203593	0.725724	0.313333
Crap7	7.100701	0.107043	0.738975	0.555686
Crap8	7.305372	0.082896	0.840736	0.673333
Crap9	7.328107	0.081013	0.69722	1.030196

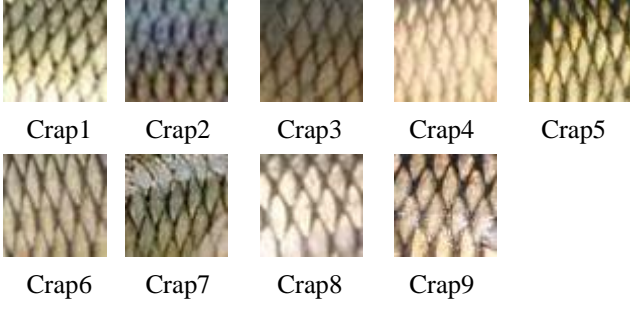


Fig 2:Body texture pattern of crap

Table 3: Statistical results of bass body texture

Name	ENT	ASM	COR	CON
Bass1	6.381112	0.309316	0.739679	0.20627
Bass2	6.211198	0.348428	0.806412	0.14823
Bass3	6.182986	0.337062	0.667224	0.20745
Bass4	6.572694	0.242015	0.764562	0.27841
Bass5	7.028404	0.106752	0.602086	0.72784
Bass6	6.412045	0.266177	0.666517	0.27764
Bass7	5.779659	0.510017	0.64922	0.13176
Bass8	6.407621	0.300795	0.764388	0.20039
Bass9	6.857924	0.117146	0.476464	0.77843
Bass10	6.834359	0.128222	0.551338	0.641176
Bass11	6.566195	0.244329	0.701262	0.33215
Bass12	6.569027	0.274583	0.700281	0.38745
Bass13	6.414443	0.260578	0.465271	0.46352
Bass14	6.412397	0.214351	0.572996	0.4201
Bass15	6.547202	0.216138	0.525349	0.53725
Bass16	6.751274	0.138794	0.600531	0.56784

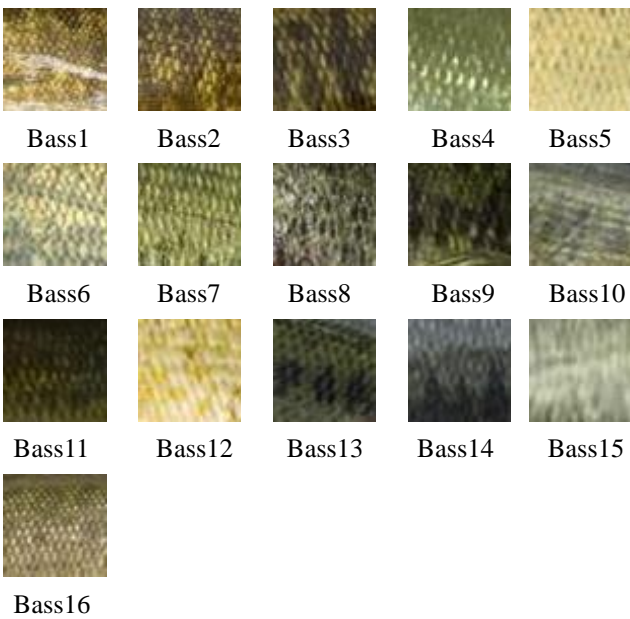


Fig3:Body texture pattern of bass

3. Simulation

By adopting the co-occurrence matrix in this experiment, we can quantify three kinds of fish body texture. According to the values of co-occurrence matrix, we adopt least square method to fit coefficient. After setting a fixed output figure of the same fish, we can fit coefficient through four statistical results of every single fish species. Calculate the actual value, after getting the coefficient and compare it with the setting figure. According to the results of comparison and analysis, we can use confusion matrix to deal with. In image accuracy assessment, it is mainly used for comparative classification results and actual measured value. It shows the result in confusion matrix about the precision of the classification.

Table 4:Cofusion matrix

Name	Sample1	Sample2	Sample3	Sum
Sample1	a	b	c	A
Sample2	d	e	f	B
Sample3	g	h	i	C
Sum	D	E	F	G

where

$$\begin{aligned}
 A &= a + b + c & D &= a + d + g \\
 B &= d + e + f & E &= b + e + h \\
 C &= g + h + i & F &= c + f + i \\
 G &= A + B + C
 \end{aligned} \tag{7}$$

Kappa: a kind of calculation method of classification accuracy. A kappa calculation result ranges from -1 to 1, but usually kappa falls between 0 to 1. It can be divided into five groups to represent different levels of consistency.

$$Kappa = \frac{Pa - Pe}{1 - Pe} \tag{8}$$

where

$$Pa = \frac{a + e + i}{G}$$

$$Pe = \frac{A * D + B * E + C * F}{G * G}$$

Table 5:The result of kappa

Interval	Results
0.0-0.20	low consistency(slight)
0.21-0.40	general consistency (fair)
0.41-0.60	medium consistency (moderate)

0.61-0.80	a high degree of consistency (all)
0.81-1	almost perfect

Table 6: Statistical results of Confusion matrix

Name	Crap	Sunfish	Bass	Sum
Crap	7	1	1	9
Sunfish	0	6	2	8
Bass	0	4	12	16
Sum	7	11	15	33

$$Pe = \frac{9*7 + 8*11 + 16*15}{33*33} = 0.359045$$

$$Pa = \frac{7+6+12}{33} = 0.757576$$

$$\text{kappa} = 0.621777$$

From Table 6, we calculate the value of accuracy. According to Table 5, the result of kappa is in a high degree of consistency.

4. Conclusions

In this study, we propose a particular method to identify specified fish species by co-occurrence and confusion occurrence matrix. Due to the limited space, this study just selects three kinds of fishes. It may not applicable to all fishes, but the four characteristic value of co-occurrence are effective parameters to describe the body texture. This conclusion can be used as a reference for the classification. But the method based on co-occurrence also has its own disadvantages. It requires a great amount of calculation, not to mention that it's time-consuming. In future, we want to increase the amount of fishes and verify the study in a practical condition.

References

- (1) Bai xuebing, Wang keqi, Wang hui: "Research on the classification of wood texture based on gray level co-occurrence matrix", Journal of harbin institute of technology, Vol.37, No.12, pp.1167-1170, 2005
- (2) Haralick R M, Shanmugam K: "Texture features for image classification", IEEE Tmns on Sys, Man, and Cyb, Vol.3, No.6, pp.610-621, 1973
- (3) GAO Cheng-Cheng, HUI Xiao-Wei: "GLCM Based Texture Feature Extraction, BEIJING SURVEYING AND MAPPING ", Vol.19, No.6, pp.195-198, 2010
- (4) C.S.Lu, P.C.Chung, C.F.Chen : "Unsupervised texture

segmentation via wavelet transform ", Vol.1997, No. 05, 2010

- (5) Ulaby FT, Kouyate F'Briseo B, et al: "Textural information in SAR Images", IEEE Transactions on Geo science and Remote Sensing, Vol.24, No.2, pp.235-245, 1986
- (6) Bharati M H, Liu J J, MAC G: "Image texture analysis: Methods and comparisons", Chemo Metrics and Intelligent Laboratory Systems, Vol.72, No.1, pp.57-65, 2004
- (7) Yao H Y, Li B C : "Generalized co-occurrence matrix method for content-based image retrieval[J]. Computer Engineering Applications", Vol.34, No.6, pp.98-101, 2004