Journal of Asia-Pacific Entomology 17 (2014) 143–149

Contents lists available at ScienceDirect

Journal of Asia-Pacific Entomology

journal homepage: www.elsevier.com/locate/jape

Identification of butterfly based on their shapes when viewed from different angles using an artificial neural network

Seung-Ho Kang [a], Jung-Hee Cho [b], Sang-Hee Lee [b,*]

[a] Department of Information Security, Dongshin University, Republic of Korea
[b] Division of Mathematical Model, National Institute for Mathematical Sciences, Republic of Korea

ARTICLE INFO

Article history:
Received 2 May 2013
Revised 4 December 2013
Accepted 12 December 2013

Keywords:
Automatic species identification
Butterfly
Branch length similarity entropy
Artificial neural network

ABSTRACT

Identification of butterfly species is essential because they are directly associated with crop plants used for human and animal consumption. However, the widely used reliable methods for butterfly identification are not efficient due to complicated butterfly shapes. We previously developed a novel shape recognition method that uses branch length similarity (BLS) entropy, which is a simple branching network consisting of a single node and branches. The method has been successfully applied to recognize battle tanks and characterize human faces with different emotions. In the present study, we used the BLS entropy profile (an assemble of BLS entropies) as an input feature in a feed-forward back-propagation artificial neural network to identify butterfly species according to their shapes when viewed from different angles (for vertically adjustable angle, $\theta = \pm 10°, \pm 20°, \ldots, \pm 60°$ and for horizontally adjustable angle, $\varphi = \pm 10°, \pm 20°, \ldots, \pm 60°$). In the field, butterfly images are generally captured obliquely by camera due to butterfly alignment and viewer positioning, which generates various shapes for a given specimen. To generate different shapes of a butterfly when viewed from different angles, we projected the shapes captured from top-view to a plane rotated through angles θ and φ. Projected shapes with differing θ and φ values were used as training data for the neural network and other shapes were used as test data. Experimental results showed that our method successfully identified various butterfly shapes. In addition, we briefly discuss extension of the method to identify more complicated images of different butterfly species.

Introduction

Reliable species identification is often emphasized as the primary step for an understanding of ecology. The identification process is necessary to explore evolutionary and developmental concepts (McMillan et al., 2002; Silveira and Monteiro, 2009) as well as for monitoring the spread of pollution and disease vectors and identifying areas of biodiversity (O'Neill and Gaston, 2004; Rua et al., 2009). The identification process also has some important practical applications such as agriculture and border control, in which pests and invaders must be identified and eradicated before they become established as unwanted visitors in agricultural areas. Despite its importance, the identification process still remains a difficult, expensive, and time-consuming task, not only because of insufficient information with regard to the key descriptions for many species, but also due to the complicated shapes of species (Gaston and May, 1992; Weeks and Gaston, 1997; Weeks et al., 1999a, 1999b; Hopkins and Freckleton, 2002).

With recent developments in computer architectures as well as innovations in software technology regarding digital image processing, researchers from various backgrounds have attempted to partially overcome the problems in identification by developing computer-based systems that can automatically identify species of live plants, insects, or animals (Gaston and O'neill, 2004). These systems allowed feature extraction of image shapes (Arbuckle et al., 2001; Schroder et al., 2002; Larios et al., 2007, 2008). They were also helpful in managing the so-called "taxonomic crisis" (Dayrat, 2005).

Among animal species, butterflies are particularly important insects because they are associated with crop plants for human and animal consumption. Approximately 16,000 species of butterflies have been recorded worldwide (Zhou, 1994). Many of these have varying coloration, spots, textures, and patterns, which cause difficulties in identifying species. Classical identification methods, such as the use of taxonomic keys, as well as various modern methods such as DNA sequencing rely on manual identification and classification of butterfly species by highly trained and skilled individuals (Walter and Winterton, 2006). The current identification options available to researchers are often inefficient and time-consuming processes. Hence, there is a clear need to incorporate modern, rapid, and automatic technologies in the process of species identification. A representative identification technology, the digital automatic identification system (DAISY), was developed; this system adopted principal component analysis and linear

* Corresponding author. Tel.: +82 42 717 5736; fax: +82 42 717 5758.
E-mail addresses: sunchaos@nims.re.kr, sunchaos.sanghee@gmail.com (S.-H. Lee).

1226-8615/$ – see front matter
http://dx.doi.org/10.1016/j.aspen.2013.12.004

discriminant analysis for insect species identification using digital images (Weeks et al., 1999b). The system had accuracies of more than 90% for identifying 35 butterfly species (Watson et al., 2004). Another method, species identified automatically (SPIDA), which is based on artificial neuron networks to recognize images encoding wavelets, has also been suggested. This method has been applied to identify spiders (Do et al., 1999). Several Chinese researchers developed the web platform "Butterfly Ecology" (Chen, 2000; Hung, 2002; Lai, 2002; Liu, 2003), which uses content-based image retrieval (CBIR). Using this platform, they analyzed more than 300 butterfly specimen images, adopting 18 color patterns, seven shape patterns, and 18 texture patterns that allowed users to query butterfly images by selecting patterns. The "But2Fly" system that uses a radial basis function neural network has also been developed (Zhang, 2006; Liu et al., 2008). Although the above-mentioned systems are attractive and useful for automatically identifying butterfly species, they do nevertheless have certain limitations, which have been mentioned by Wang et al. (2012). Additionally, another limitation is that the systems use images obtained from the top-view. However, in field, obtaining target object images in top-view by using a camera is difficult due to the butterflies' alignment and viewers' positioning. To overcome the view angle problem, we suggested a new method using branch length similarity (BLS) entropy (Lee, 2010; Lee et al., 2010, 2011) for identifying the boundary shape pattern of a butterfly as the feature to automatically identify butterfly species. In our previous study, we showed that BLS entropy concept can be a powerful tool to characterize the butterfly shape (Kang et al., 2012a, 2012b). In this study, we focus on the identification of butterfly species based only on shapes, and do not take into account other information from butterfly wings (e.g., wing color and patterns). Herein, we briefly discuss a means to reduce the computational burden in the identification process.

Materials and methods

Butterfly wing images

We used images of 150 specimens from 15 butterfly species: *Celastrina argiolus*, *Cynthia cardui*, *Dilipa fenestra*, *Favonius orientalis*, *Graphium sarpedon*, *Libythea celtis*, *Luehdorfia puziloi*, *Lycaena dispar*, *Lycaena phlaeas*, *Ochlodes subhyalina*, *Papilio maackii*, *Papilio xuthus*, *Parantica sita*, *Parnassius bremeri*, and *Sasakia charonda* (see Fig. 1). For each species, there were 10 specimens. Each image had a resolution of 760 × 507 pixels and was in RGB color. We focused on the shape of butterfly wings. Thus, the left wing of each specimen was segregated from the original image and transformed into a binary image with a resolution of 379 × 567 pixels. As shown in Fig. 2, each image pixel was described by a value of 1 (black) or 0 (white) by using the image processing tool provided by MATLAB ver. R2007a. The Prewitt edge detection algorithm (Gonzalez and Woods, 2002) was then used to obtain an image of the boundary shape of the left wing from the corresponding binary image.

In order to enable visualization of wing shapes from different viewer angles, we transformed the original images obtained from the top-view by using the orthogonal projection method. The projection method is shown in Fig. 3. According to the viewer's positioning, a square shape is shown as a diamond shape, which can be characterized as two angle variables, θ for vertically rotated projection and φ for horizontally rotated projection. In this study, the projection angle was considered as $\theta = 0, \pm 10°, \pm 20°, \ldots, \pm 60°$ and $\varphi = 0, \pm 10°, \pm 20°, \ldots, \pm 60°$.

Fig. 3(a) shows how a square shape is visible to a viewer when it forms an angle of θ. (m, n) represents the coordinates of a point on the boundary of the original square shape. L represents the distance from the viewer to the square. We arbitrarily set the value as 200. From the viewpoint of the viewer, (m, n) is seen as (m', n') coordinates. The mathematical relationship between the two points can be simply represented using triangle proportionality (Fig. 3(b)) as follows:

$$(m', n') = \left(\frac{mL}{L + n\cos(90-\theta)}, \frac{nL\sin(90-\theta)}{L + n\cos(90-\theta)} \right)$$

To obtain images that were horizontally rotated at an angle, φ, we rotated the original images at 90° on the same plane using two-dimensional rotational matrix as follows:

$$(p', q') = \begin{pmatrix} \cos(90^\circ) & -\sin(90^\circ) \\ \sin(90^\circ) & \cos(90^\circ) \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix}$$

where, (p, q) and (p', q') represent the coordinates of a boundary pixel of the original shape and the rotated shape, respectively. Next, the rotated image was vertically projected according to the relationship between (m, n) and (m', n'). This is the same with the case when the shape is projected such that there are changes in both θ and φ. Fig. 4 shows the original shape of *Papilio maackii* and its projected shapes

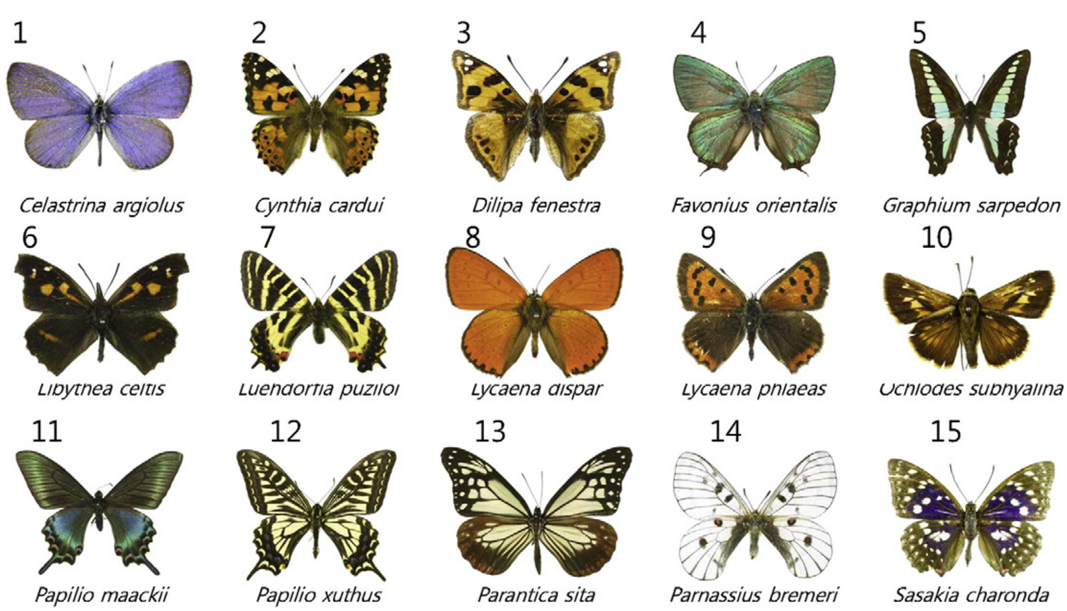

Fig. 1. Representative images for 15 butterfly species used for automatic species identification.