

Classification of Live Moths Combining Texture, Color and Shape Primitives

Gustavo E. A. P. A. Batista*[†], Bilson Campana*, Eamonn Keogh*

*Department of Computer Science & Engineering
University of California Riverside
Riverside, CA, USA

[†]Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo
São Carlos, SP, Brazil
{gbatista, bcampana, eamonn}@cs.ucr.edu

Abstract—Each year, insect-borne diseases kill more than one million people, and harmful insects destroy tens of billions of dollars worth of crops and livestock. At the same time, beneficial insects pollinate three-quarters of all food consumed by humans. Given the extraordinary impact of insects on human life, it is somewhat surprising that machine learning has made very little impact on understanding (and hence, controlling) insects. In this work we discuss why this is the case, and argue that a confluence of facts make the time ripe for machine learning research to reach out to the entomological community and help them solve some important problems. As a concrete example, we show how we can solve an important classification problem in commercial entomology by leveraging off recent progress in shape, color and texture measures.

Keywords-Live insect classification, distance measures combination, k -nearest neighbor

I. INTRODUCTION

The impact of insects on human health and well-being are immeasurable, both in a positive and negative sense. However these complex interrelationships are not static, but changing as humans change the environment. As a simple example, consider the Spruce Bark Beetle (*Dendroctonus rufipennis*). Due to global warming the average temperature in Anchorage has increased 3.9 degrees Fahrenheit over the last century. This has allowed the more Beetles to survive the normally harsh winters, exploding their population and allowing them to devastate and estimated 30 million trees per year over the last few years [1]. A similar story can be told in all the Northern US states and in Canada.

These problems have attracted the attention of entomologists, chemists, geologists etc. However we are unaware of any computer science efforts, apart from some GIS data management. Does computer science, in particular machine learning, really have nothing to offer here? Most of the techniques to mitigate insect's impact, be they mosquitoes carrying malaria, beetles destroying forest or moths ruining crops, rely on knowing the locations of outbreaks early on. This is typically achieved by using traps (pitfall traps, flight interception traps etc) to capture some insects, which are manually counted every few days. The obvious problem with this is that it is labor intensive and expensive, and a few days

are a long time in the life span of many insects. Thus by the time the infestation has recognized, the insects have already gotten a head start. An obvious solution to this problem would be to use computers to count insects, yet this has not happened, why?

In an influential paper Gaston & O'Neill ask "*Automated species identification: why not?*" [2]. They go on to answer their own question by listing the most commonly stated objections. The most important of these is suggested as "*the simplest explanation for why automated identifications have not become the norm for routine identifications is that such an approach is too difficult.*" We believe the difficulties lie in the two areas, the need for preprocessing of the data (cropping, edge extraction, histogram equalization and rectification, dimensionality reduction etc) and the need for careful setting of many parameters. For example, they note "*Artificial neural networks (ANNs) are the most commonly employed computerized pattern recognition tool*" however, ANNs can have more than a dozen parameters (initial weights, transfer functions, architectures and learning rules etc).

In this work we propose to mitigate these difficulties by removing the need for complex preprocessing and using a simple approach based on a k -nearest neighbor classifier that combines distance measures for texture, color and shape. We show that our approach is conceptually simple but still competitive with more sophisticated approaches previously published in literature¹.

The rest of this paper is organized as follows. Section II discusses how the texture, color and shape primitives are extracted from images and codified as distance measures. Section III explains how our approach combines distance measures using a simple grid search procedure. Section IV presents our experimental results and compares with previously published work. Section V gives conclusions and directions for future work.

¹There are results in the literature that are slightly superior to our results. However, it appears that the authors went through multiple cycles of adjusting the parameters on the test set. In contrast we have scrupulously avoiding changing (our very few) parameters after seeing the test set.

II. TEXTURE, COLOR AND SHAPE PRIMITIVES

In this work, we classify moths according to their species combining color, shape and texture information, encoded as distance measures among objects. Identification of live insects from images imposes several challenges. For instance, light conditions may vary according to time, weather and equipment available, imposing difficulties in recognizing colors. Live insects are not easily photographed in the exact same pose and orientation, meaning that the same species of insects might present considerably different shapes. Additionally, cameras might be slightly out of focus, causing blurred images that make texture identification difficult. Although these issues might impose difficulties for a classifier that relies in only one of these features, we show that a combination of these features can produce an accurate and robust classifier.

Our approach leverages off a recently introduced texture distance measure, CK-1 [3], by combining it with color and shape distances using a simple linear combination. A nearest neighbor classifier uses the combined distance to classify new images. This approach is evaluated in a collection of live moth images with 35 different species found in the British Isles [4]. In order to make this paper self-contained, we start by reviewing the distance measures used in this work.

A. Texture Similarity

CK-1 is a compression-based distance for texture inspired by the theoretical concept of Kolmogorov complexity. Although Kolmogorov conditional complexity gives rise to a distance measure that is optimal in the sense that it subsumes other measures [5], such distance is uncomputable in the general case. Therefore, several researchers have proposed approximations to this distance using compression algorithms [6], [7] and many others have evaluated these approximations in diverse domains [8], [9].

CK-1 extends the applicability of compression-based distances to image textures by using video compression. Given two images x and y , CK-1 is defined by Equation 1.

$$d(x, y) = \frac{C(x|y) + C(y|x)}{C(x|x) + C(y|y)} - 1 \quad (1)$$

where $C(a|b)$ is the size of a synthetic MPEG-1 video composed by two frames a and b , in this order.

MPEG-1, and most video encoding algorithms, achieve compression by finding recurring patterns within a frame (intra frame compression) and/or between frames (inter frame compression). When x and y are two similar images, inter frame compression step should be able to exploit that to produce a smaller file size, which can be interpreted as significant similarity. As digital video is an important commercial application, many efforts have been made to achieve high compression rates in video encoding, making it a good approximation of the Kolmogorov conditional complexity.

Another observation is that there is no necessity to “hack” the video encoding algorithm, since no internal modification is necessary. In order to reinforce its simplicity, [3] shows that this measure can be implemented in just one line of Matlab code.

Notice that as CK-1 uses video encoding, there is no necessity to linearize images. Therefore, no spatial information is lost. This feature makes CK-1 well suited for measuring texture similarity. In order to achieve color invariance, all images are converted to gray scale intensity values. However, in the case of moth classification, color and shape might also present useful information to classify at least some classes. We extract this information in the form of distance measures, as explained next.

B. Color Similarity

Color histograms are a simple representation to capture color information from images. Although color histograms are insensitive to position and rotation of an insect in an image, different species might share the same color, making color an useful secondary information. For instance, several moth species have similar shades of gray or brown, such as *Agrotis exclamationis* and *Ochropleura plecta*, two common European moths. However, color is useful in discriminating at least some species, such as *Campaea margaritata* also known as the Light Emerald moth because of its distinctive green hue, and *Cabera pusaria* also known as the Common White Wave.

In our experiments, we extracted histograms for red, green and blue color channels independently, and concatenated these histograms in a unique “time” series. Since light conditions vary for different images, we reduced color variation by grouping pixels in bins of size 4, resulting in $256/4 = 64$ observations for each color channel. We also did not count white pixels since virtually all of them were present in the images background, and therefore do not present any useful information for classification purposes. We employed the standard Euclidean distance to measure distances between pairs of color histograms. We chose Euclidean distance since it proved to be competitive to more sophisticated measures, and it is parameter-free.

C. Shape Similarity

Shape is one of the most relevant features to determine similarity among objects. Although there are dozens of different shape similarities measures [10], we are interested in one that is simple to compute and does not require tuning many parameters. In this work, we convert a two-dimensional shape to a single-dimensional “time” series, by calculating the distance between a central point and the object contour. This approach is simple to implement and does not require any parameters to be tuned.

Once each moth shape was converted to a time series, several distance measures can be used to estimate similarity

between each pair of specimen. In this work, we use the Euclidean distance due its simplicity and competitiveness with more complex measures on shape matching problems [11].

Shape is an useful feature for classifying insects, since species frequently present morphological differences. In the case of moths, although several species can roll their wings close to their bodies when resting, and therefore having a similar “triangular” shape, such as *Anaplectoides prasina*, other species do not have the same ability and present a very different shape when resting, such as the odd-looking *Laothoe populi*. Nevertheless, some individuals are difficult to classify using shape similarity only. Unlike photographs of dead insects that can are precisely posed, images of live moths might present a large variety of shapes.

III. COMBINING DISTANCE MEASURES

Linear combination is a simple approach to compose distances measures. In the case of composing texture, color and shape, it assumes the form shown in Equation 2.

$$d_{comb}(x, y) = w_t \times d_{texture}(x, y) + w_c \times d_{color}(x, y) + w_s \times d_{shape}(x, y) \quad (2)$$

where x and y are two images, $d_{texture}$, d_{color} and d_{shape} are distance measures for the texture, color and shape primitives, respectively, and w_t , w_c and w_s are weights for texture, color and shape, respectively. In order to make the distance measures commensurable, we normalized the three distance measures in the interval $[0, 1]$. A problem is how to determine values for w_t , w_c and w_s that would maximize accuracy or other measure of success.

Notice that certain values of distance weights produce the exact same combined distance measure, but with different scales. For instance, $d_{comb1} = (w_t = 0.1, w_c = 0.1, w_s = 0.1)$ and $d_{comb2} = (w_t = 0.2, w_c = 0.2, w_s = 0.2)$ have the same distance information, since $d_{comb2}(x, y) = 2 \times d_{comb1}(x, y)$ for any images x and y . However, for a classification procedure such as k -nearest neighbor, the scale is not relevant. We simplify the search space eliminating one degree of freedom by making:

$$w_t = w \quad (3)$$

$$w_c = (1 - w) \quad (4)$$

This choice of w_t and w_c is arbitrary. However, we verified that choosing any other combination of two parameters produce similar results.

Therefore, Equation 2 can be rewritten as:

$$d_{comb}(x, y) = w \times d_{texture}(x, y) + (1 - w) \times d_{color}(x, y) + w_s \times d_{shape}(x, y) \quad (5)$$

$(w - 0.1, w_s - 0.1)$	$(w - 0.1, w_s)$	$(w - 0.1, w_s + 0.1)$
$(w, w_s - 0.1)$	(w, w_s)	$(w, w_s + 0.1)$
$(w + 0.1, w_s - 0.1)$	$(w + 0.1, w_s)$	$(w + 0.1, w_s + 0.1)$

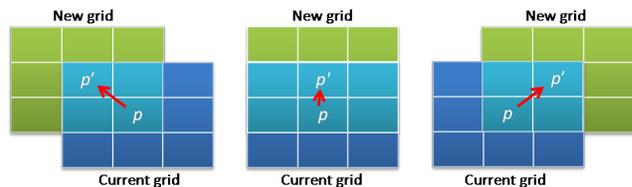


Figure 1: Our greedy search algorithm uses a 3×3 grid (top). The central cell, p , stores the current values for w and w_s parameters. The highest accuracy in grid, p' , determines the new values for w and w_s (bottom).

with $w \in [0, 1]$. Parameter w_s can vary independently from w .

Several approaches can be used to search for the w and w_s parameters, ranging from a simple and computationally intensive brute-force search to more sophisticated approaches such as heuristic and stochastic search procedures.

We designed a simple greedy search procedure, using a 3×3 grid. The central cell of this grid stores the classifier accuracy for current values of w and w_s . Remaining grid cells store accuracies for possible combinations of parameter values using a single increment or decrement. In our experiments, a 0.1 increment was used for both w and w_s . Figure 1 illustrates this grid.

The greedy search algorithm simply chooses the highest accuracy in grid and updates the values of w and w_s accordingly. The new values for these parameters are stored in the middle of a new grid, and the accuracies for the new grid are updated. The search ends when, for a given grid, the current parameter values, i.e. the central cell, has the highest accuracy in the grid.

IV. RESULTS AND DISCUSSION

The collection used in this paper consists of 774 images of live moths belonging to 35 different species found in United Kingdom [12]. Each image is 1024×960 pixels in resolution with 24-bit RGB color. As moths usually occupy about 10% of the total image area, we cropped each image by finding the nearest bounding rectangle around each moth, resulting in images of approximately 500×800 pixels each. In addition, the background was deleted with a semi-automatic technique.

We used accuracy as the main method to assess our results. All results were obtained using the 1-nearest neighbor classifier with leave-one-out cross-validation. We measured the accuracy for 1-nearest neighbor for each distance individually, and obtained 70.98% for texture, 50.13% for color and 49.87% for shape. Combining the three distance

functions with our greedy grid search algorithm with 10% of the training set used to evaluate the parameters values (validation set), the 1-nearest neighbor classifier obtained 79.53% accuracy.

The work of Mayo & Watson [4] used the same data set, but with different features and pre-processing techniques. Their approach uses a much more data intensive algorithm that extracts a total of 11,300 numerical features, including global statistics and local features extracted from images “patches” of 30×30 pixels each. In addition, they used several pre-processing techniques including smoothing by averaging color intensity in a 3×3 pixel neighborhood grid, edge detection, background removal and conversion to binary images. Features were extracted from the black and white images as well as color images. In the case of color images, features were extracted from the RGB and HSB color spaces.

With the same nearest neighbor classifier used in this work, Mayo & Watson obtained accuracies of 71.6% for one neighbor and 65.36% for five neighbors. Their best approach was a Support Vector Machine classifier that obtained 84.8% accuracy. Beyond the much greater time complexity of their method, the authors acknowledge (personal communication) that their results may be optimistic given that they adjusted parameters, settings and choice of pre-processing techniques by looking at test data.

V. CONCLUSION AND FUTURE WORK

In [2], Gaston & O’Neill discuss the possible reasons why automated classification is not the norm for routine insect identification, and they point out that a commonly stated objection is that such approach is too difficult. As a concrete example, we present in this paper a conceptually simple scheme to classify live moths from images based on texture, color and shape features. The algorithms employed to extract those features require the set up of few or no parameters, and can be implemented in few lines of source code. Nevertheless, our results are competitive with more sophisticated techniques previously published in literature.

We believe that simple techniques with few parameters can boost the adoption of our ideas by researchers of diverse areas such as biology, entomology and lepidopterology. The application of parameter laden algorithms have many problems, one of them is that with several parameter to set up it is difficult to avoid overfitting. An additional problem of parameter-laden algorithms is that they make it exceptionally difficult to reproduce published experimental results, and to truly understand the contribution of a proposed algorithm [7].

ACKNOWLEDGMENT

This work was funded by NSF 0808770 and FAPESP 2009/06349-0. The authors would like to thank Anna Watson

and Michael Mayo for donating the collection of moth images.

REFERENCES

- [1] G. P. Juda, V. Barber, P. Duffy, H. Linderholm, S. Rupp, S. Sparrow, E. Vaganov, J. Yarie, E. Berg, R. D’Arrigo, O. Eggertsson, V. Furyaev, E. H. Hogg, S. Huttunen, G. Jacoby, V. Kaplunov, S. Kellomaki, A. Kirilyanov, C. E. Lewis, S. Linder, M. Naurzbaev, F. Pleshikov, U. T. Runeson, Y. Savva, O. Sidorova, V. Stakanov, N.M.Tchebakova, E.N.Valendik, E.F.Vedrova, and M. Wilmking, “Climate change and insects as a forest disturbance in the arctic,” *Encyclopedia of Earth*, 2010, eds. Cutler J. Cleveland (Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment). International Arctic Science Committee (Content Partner); Sidney Draggan (Topic Editor).
- [2] K. J. Gaston and M. A. O’Neill, “Automated species identification: why not?” *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 359, no. 1444, pp. 655–667, 2004.
- [3] B. Campana and E. Keogh, “A compression based distance measure for texture,” in *SIAM International Conference on Data Mining*, 2010, pp. 850–861.
- [4] M. Mayo and A. T. Watson, “Automatic species identification of live moths,” *Knowledge-Based Systems*, vol. 20, no. 2, pp. 195–202, 2007.
- [5] M. Li and P. Vitanyi, *An introduction to Kolmogorov complexity and its applications*, 2nd ed. Springer Verlag, 1997.
- [6] M. Li, X. Chen, X. Li, B. Ma, and P. Vitanyi, “The similarity metric,” in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2003, pp. 863–872.
- [7] E. J. Keogh, S. Lonardi, C. A. Ratanamahatana, L. Wei, S. Lee, and J. Handley, “Compression-based data mining of sequential data,” *Data Mining and Knowledge Discovery*, vol. 14, no. 1, pp. 99–129, 2007.
- [8] N. Krasnogor and D. A. Pelta, “Measuring the similarity of protein structures by means of the universal similarity metric,” *Bioinformatics*, vol. 20, pp. 1015–1021, 2004.
- [9] A. Bratko, G. Cormack, B. Filipic, T. Lynam, and B. Zupan, “Spam filtering using statistical data compression models,” *Journal of Machine Learning Research*, vol. 7, pp. 2673–2698, 2006.
- [10] R. C. Veltkamp and L. J. Latecki, “Performance of shape similarity measures,” in *Proceedings of IFCS 2006 Conference: Data Science and Classification*, 2006.
- [11] E. Keogh, L. Wei, X. Xi, S. H. Lee, and M. Vlachos, “Lb_keogh supports exact indexing of shapes under rotation invariance with arbitrary representations and distance measures,” in *Proceedings of Very Large Databases*, 2006, pp. 882–893.
- [12] A. T. Watson, M. A. O’Neill, and I. J. Kitching, “Automated identification of live moths (macrolepidoptera) using digital automated identification system (daisy),” *Systematics and Biodiversity*, vol. 1, no. 3, pp. 287–300, 2003.