2008 International Conference on BioMedical Engineering and Informatics

A New Type of Image-Based Key

Dr. Bruce Kirchoff
Dr. David Remington
Department of Biology, UNC Greensboro, Greensboro, NC 27402-6170
email: kirchoff@uncg.edu

Dr. Lixin Fu
Dr. Fereidoon Sadri
Department of Computer Science, UNC Greensboro, Greensboro, NC 27402-6170
email: lfu@uncg.edu

Abstract

Keys are character based tools for plant identification. They are based on the decomposition of the plant into very small, atomistic parts. These parts are described with the technical and often arcane terminology of plant taxonomy. Even the best electronic keys (Delta, Lucid) make use of this terminology. Keys are not based on pattern recognition, the forte of visual experts. Instead they demand that the user look at the plant as if it consisted of a series of isolated parts that are classified by name. Keys would be more effective if they were visually based. They would be easier to use for visual experts because accurate perception is their providence. They would also be easier to use for novices because they would not depend on knowledge of arcane terminology. This paper proposed an innovative image-based key system for species recognition.

1. Introduction

Imagine a field botanist strolling through a wood with a botanical companion from another country. As they stroll he points out and names a number of species, to the delight of his companion. He does this without interrupting their pleasant walk. In the middle of a sentence he looks up and says “oh, you must see this,” and names a species unique to the area. The companion is able to identify some individual to genus, but knows few of the species. His eye is drawn instead to some of the characteristics that differentiate the species from those in his native country. He can pick out the differences, but species recognition is not automatic.

This description of expert and novice recognition is based on experiences learning new plants, and is supported by research in cognitive psychology. Expert pattern recognition is based on a visual processing mode that develops through experience. This mode, holistic processing, is unique to visual experts. Novices are analytic processors: they see parts when they see anything at all. The parts that they see most clearly are those that they have learned to recognize from prior experience. For instance, most children collect and play with leaves. These experiences form the basis for their later recognition of leaves as distinct plant parts. On the other hand, few people ever look closely at flowers, and have difficulty recognizing all but the most obvious floral parts. Until they gain experience with a range of flowers they have difficulty telling a stamen from a staminode, a hypanthium from a perianth tube. When they gain this experience they actually change the way that they see. They no longer see just the parts, but can attend to the whole structure in a way that is not possible for a visual novice.

In addition to being a visual expert, the field botanist in our example is also a disciplinary expert. He can recognize and name plant parts because of his long experience with the terminology of plant taxonomy. As a disciplinary expert he has the conceptual tools to use the identification aids for unknown species: keys. A plant key is designed to help user identify the *species* of plant an unknown *specimen* belongs to. This is usually done with a series of dichotomous choices based on verbal description; e.g.:

- *1A:* leaves opposite or whorled *2*
- *1B:* leaves alternate *17*
- *2A:* bud scales imbricate 3
- *2B:* bud scales valvate or missing 11
- *3A:* leaves compound *Acer negundo*
- *3B:* leaves simple 4

Keys are character based tools for plant identification. They are based on the decomposition of the plant into very small, atomistic parts. These parts are described with the technical and often arcane terminology of plant taxonomy. Even the best electronic keys (Delta, Lucid) make use of this terminology. Keys are not based on pattern recognition, the forte of visual experts. Instead they demand that the user look at the

978-0-7695-3118-2 2008

DOI 10.1109/BMEI.2008.301

plant as if it consisted of a series of isolated parts that are classified by name. Keys would be more effective if they were visually based. They would be easier to use for visual experts because accurate perception is their providence. They would also be easier to use for novices because they would not depend on knowledge of arcane terminology. The dependence of keys on the technical terminology makes it virtually impossible for novices to use them. A partial list of the terms describing leaf shape will demonstrate the problem: linear, oblong lanceolate, elliptic, oblanceolate, ovate, broadly elliptic, obovate, orbicular, reniform [1]. To make things worse, there are separate sets of terms for the shape of the leaf apex, of the leaf base, of the margin, for the texture of the leaf, its venation, its covering of hairs, not to mention the terms associated with the other parts of the plant [1-3]. Needless to say, only novices who see a direct application for this knowledge in their careers will be motivated to learn to use keys. Few members of the general public will have the required tenacity.

2. Visual key approach

To rectify above problems we will propose the development of a new type of visual key. By a “visual” we mean a key that uses images, with little or no reliance on terminology. These keys will be easier to use for both experts and novices. For experts, the use of images more closely reflects the way that they recognize plants. For novices, they key avoids problems with terminology. Plants can be identified solely by visual means. In this section we lay out our preliminary ideas on the creation of these keys.

All of the proposed keys will be computer based. Images will be displayed on the screen, and the user will make selections based on the similarity between the displayed images and the relevant parts of the plant he wants to identify. The computer will track these selections and compute Bayesian posterior probabilities for the likelihood of the unknown’s identity. These probabilities will be used to assign the unknown plant to the correct taxon. The user’s job is to select the images that are most similar to the unknown plant. The key tracks these selections and computes the most likely identity of the unknown.

Like conventional keys, the user will progress through the visual keys character by character. The difference will be that the visual characters will not be atomistic, and will not be associated with terminology. For the purpose of the visual keys a character will be a collection of images of homologous (comparable) plant parts [4]. For instance, a collection of leaf images represents one character; of bud images another; collections of images of bark, twigs, and flowers represent three more characters. The characters are always images of complex plant parts that can be easily recognized by both novices and experts. This is the strength of our method. Because the characters used to identify the unknown are images, they will be more familiar to novices than terminology-based characters. A shorter learning period will be required before the keys can be used effectively.

There are several possible methods of image display and probability calculation. The first, and most radical of these methods will be described first. In this method the images constituting one character are displayed randomly, without replacement. Random image display is used to simplify the statistical calculations. The user’s task is to select the image that matches the unknown, or to reject all of the images; i.e., to tell the program that none resemble the unknown. As images are selected the program calculates posterior probabilities for each species. The initial priors are all equal. As an image belonging to a specific taxon is selected, its probability of being the unknown increases. The random display and selection of images continues until the posterior probability of some taxon reaches a predefined level (ca. 95%). If this probability level is not reached with the first character, the program gives the user the option of moving to a second, third and so forth until the desired probability level is reached. At this point the program displays one image from each character of the selected taxon and asks the user to confirm that it is the unknown. The user will also have access to a written description of the taxon, and a summary drawing that shows the structure of parts not included in the key. If he rejects this choice, or if he is not sure, the program will give him two options. He can either start over and use the key again, or see summary displays of other likely species.

Although this approach should work well for small groups of taxa, it is unlikely to be effective in very large groups. As the number of species covered by the key increases, the chances of seeing an image that sufficiently resembles the unknown will decline precipitously. The user may have to go though 10 or more screens before he sees a leaf that resembles his unknown. This will make the key unusable.

3. Taxon identification and image display

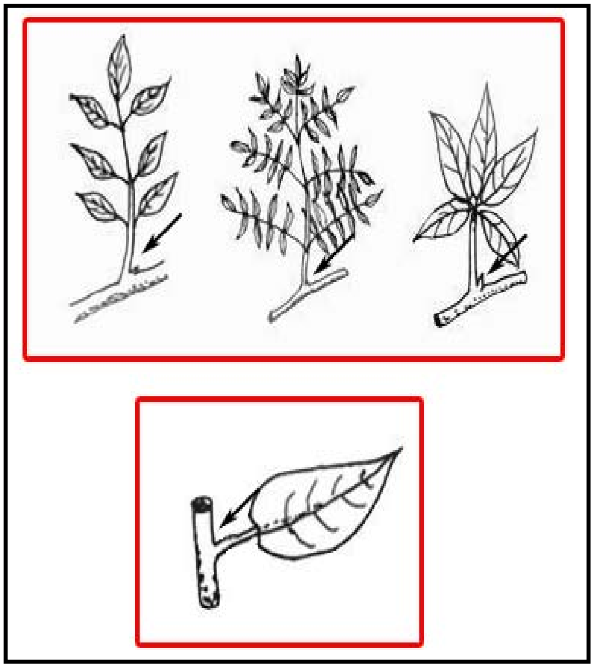

Figure 1: Compound (upper) vs. simple (lower) leaf selection. Arrows indicate bud position.

To remedy the situation with large numbers of taxa, we will use a different approach when the number of possible taxa is greater than 20. In this case we will develop separate routines to track taxon identification and image display.

Taxon identity will be tracked in basically the same way as above. The initial images displayed will be drawings that summarize major differences in character structure. For instance, the user will be asked to differentiate between compound and simple images by selecting one of two drawings that represents this dichotomy (Fig. 1). Use of summary drawings allows us to quickly eliminate large groups of images, and taxa. After the number of possible images has been reduced through display and selection of summary images, we will switch to a new method of photographic image display. Later images will be selected for display based on a character similarity matrix developed out of our work on character cladograms [4, 5]. This matrix will be constructed so that it can be easily expanded to include new taxa. Our eventually goal is to produce a key to all of the plants of the Southeastern United States. Use of an expandable similarity matrix will allow us to approach this goal in stages.

Separating taxon identification from image display means that the program will have to track two sets of probability estimates, one for individual taxa and the other for larger groups. Several possible options exist for methods to do this within a Bayesian framework. The images for display will be tracked using the similarity matrix, updated based on the previous similarity choices made by the user. This approach is expected to provide strong statistical estimates of confidence in the correct identification with a much smaller number of image display sets.

4. Bayesian statistical framework

4.1 Bayes' Theorem

To provide the user with a confidence estimate for correct taxon identification, we propose to use a Bayesian statistical framework. Bayesian parameter estimation is based on Bayes' theorem [6]:

$$P(X \mid Y) = \frac{P(Y \mid X)P(X)}{P(Y)}$$

$$= \frac{P(Y \mid X)P(X)}{P(Y \mid X)P(X) + P(Y \mid X^C)P(X^C)}$$

where X and Y are events, $P(X)$ represents the probability of X, $P(X|Y)$ represents the probablility of X conditioned on Y, and X^C is the complement of X such that $P(X) + P(X^C) = 1$.

The theorem can be applied to probability distributions:

$$f(params \mid data) = \frac{L(data \mid params) f(params)}{\int L(data \mid params) f(params) dparams}$$

Here, *f(params)* represents the *prior distributions* of a set of parameters, *L(data | params)* is likelihood for observing the data given a set of parameter values, and *f(params | data)* is the *posterior distribution* of the parameters. Thus, in a Bayesian statistical model, parameters are interpreted as belonging to a probability distribution, rather than fixed but unknown values as they are understood in conventional frequentist statistics. The sample data is used to update the prior distribution so that the posterior distribution provides point and interval estimates of the parameter values. Even though prior distributions are often defined to be uninformative (e.g. uniform distributions), the Bayesian framework is especially useful for evaluating the effect of uncertainty about the value of some parameters, including "nuisance parameters" on estimates of other parameters.

4.2 Statistical framework

In our model, the parameters represent confidence that the key user's specimen belongs to a particular species. A Bayesian framework allows us to account for the uncertainty associated with the large number of species to which the specimen could potentially belong and avoid zero probabilities. We use a flexible sampling procedure that allows for preferential sampling of species that have been selected in previous rounds to increase statistical precision for the most likely species. Data corresponding to the number of times the user has been shown a picture of a particular species (n_i) and the number of times he/she has selected that species as matching the specimen (x_i) are stored in a table as in the following example:

Species	# *observations* (n_i)	# *selections* (x_i)
A	3	0
B	4	0
C	4	1
D	9	8
E	3	0
F	7	4
G	5	1

Assume m species are represented in the key. The parameters $p_1, p_2, \ldots, p_m$ are confidences associated with each species (i.e. confidence of the species choice) and follow a *Dirichlet distribution* with a 0-1 range. Parameters of the prior distribution could be defined to make any choice equally likely. When considering species i alone, this is equivalent to *Beta distribution* with a prior mean of p_i equal to $1/m$.The likelihood for species i is a binomial distribution, which gives the probability of selecting pictures of species i x_i times out of n_i observations.

The posterior distribution of p_i is also Beta, but the posterior mean is $(x_i+1)/(n_i+m)$. The posterior mean can be thought of as the constancy with which species i is selected when it is observed. Each of the m traits has a separate posterior distribution and constancy estimate. These will add to 1, so do not represent true confidences (probabilities). To obtain the posterior confidence for species i, one can use the ratio of the constancy to the sum across all species in the key. That is,

$$\tilde{c}_i = \frac{\bar{p}_i}{\sum_{j=1}^{m} \bar{p}_j},$$

where $\tilde{c}_i$ represents the posterior confidence estimate.

5. Extensions

When multiple kinds of characters are used (e.g. leaves and flowers), the statistical framework stays the same, but the likelihood becomes compound (one binomial distribution per character for each species). Thus, the data consist of separate n_{ij} and x_{ij} columns in the data matrix for each character j.

When large numbers of species are included in the key, a similarity matrix based on previously-determined visual similarities among the species could be combined with data from the user's selections among the initial line drawings to develop the prior distribution. Thus (based on the example in Fig. 1), if the user selected a picture of a compound leaf as matching the specimen, parameters of the prior distribution would be adjusted to give more weight to plants with compound leaves and much lower (but still non-zero) weights to plants with simple leaves. Thus, the user would be shown primarily, but not exclusively, photographs of compound-leaved plants, making the identification process much more streamlined but still providing an opportunity to "jump" out of a mistaken initial choice.

6. System implementation issues

In the simplest form, a series of fixed screens are presented to the user during the identification process. An improved approach uses user's previous interactions to guide the selection of future screens that are presented. Accomplishing this improvement requires significant enhancements to the implementation as follows:

First, we store the pictures and photos in a flexible database. Each picture or photo is annotated with information, such as its taxon (or taxa, if more than one applies) and the database is sorted and indexed on the taxa to facilitate efficient access. Further, we develop selection rules, basically an expert system, that decides on the photos to include in the next display to be presented to the user. The history of user interaction is captured via the set of taxa of the photos selected so far. The "next screen selection rule" uses the interaction history, and determines the set of taxa for photos to include in the next screen. The photos are then selected randomly, or in sequential order, from the database. A history of included photos is maintained to avoid selecting the same photos repeatedly. We envision building the expert system incrementally. The initial set of rules is determined by interviewing experts in the field. Then the rule set is refined through system usage during a test-and-refine period based on the degree of success of user interaction. In this respect, we implement a "learning system" that refines the rule set as the system is being used during this period.

7. Conclusion

In this paper we point out the advantages of visual keys instead of arcane terminology for species identification, which relies on a holistic view instead of individual parts. Based on Bayesian statistical framework, we propose a computer-based, user interactive learning system that may greatly benefit biology education and research.

References

[1] Walters, D.R. and D.J. Keil.1996. *Vascular Plant Taxonomy*. 4th Updated Version ed. Dubuque, IA: Kendall Hunt. 621 pp.

[2] Judd, W.S., C.S. Campbell, E.A. Kellog, P.F. Stevens, and M.J. Donoghue.2002. *Plant Systematics: A phylogenetic approach*. 2nd ed. Sunderland, MA: Sinauer Associates. 576 pp.

[3] Simpson, M.G.2006. *Plant Systematics*. Amsterdam: Elsevierpp.

[4] Kirchoff, B.K., S.J. Richter, and D.L. Remington. 2007. *Characters as Groups: A New Approach to Morphological Characters in Phylogenetic Analysis*. Taxon **56:479-492**.

[5] Kirchoff, B.K., S.J. Richter, D.L. Remington, and E. Wisniewski. 2004. *Complex data produce better characters*. Systematic Biology **53**: 1-17.

[6] Hastie T., Tibshirani R., and Friedman J. *The elements of statistical learning, data mining, inference, and prediction*. Springer.2001.