J. Appl. Ent. 123, 1–8 (1999)

ISSN 0931-2048

Automating insect identification: exploring the limitations of a prototype system

P. J. D. Weeks[1], M. A. O'Neill[2], K. J. Gaston[3] and I. D. Gauld[1]
[1]Department of Entomology, The Natural History Museum, London, UK; [2]Oxford Orthopaedic Engineering Centre, Oxford, UK and [3]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

Abstract: Automated identification systems based on computer image analysis technology provide an attractive, but as yet unexploited potential solution to the growing burden of routine species identifications presently faced by a dwindling community of expert taxonomists. DAISY (the Digital Automated Identification SYstem) is a prototype novel automated identification system, developed to explore this possibility. In its pilot phase, the DAISY algorithms were developed to discriminate five species of parasitic wasp, based on differences in their wing structure. Here, again using wing form, the ability of DAISY to discriminate amongst an order of magnitude more species – 49 species of closely related biting midges is examined. In so doing an attempt is made to establish a set of basic 'benchmark' tests of the efficacy, and weaknesses, of such an automated identification system.

1 Introduction

Reliable species-level identification of organisms is a fundamental part of most biological work. Accurate identification underpins the control of agricultural pests (LATTIN and KNUTSON, 1982; HAWKSWORTH, 1994), is vital for formulating conservation legislation (MAY, 1990), assists with pharmaceutical prospecting (REID et al., 1993), and is essential for monitoring the spread of pollution and disease vectors (CHALMERS, 1996). Species names are the language of biology, and the label by which one accesses information about organisms. Notwithstanding this importance, the body of taxonomic expertise available to carry out reliable identifications of insect pests, pathogens and environmental indicators is being steadily eroded world-wide (GASTON and MAY, 1992) and demand for routine identifications now far outstrips the capabilities of the dwindling biosystematics community (HOLDEN, 1989; HOUSE OF LORDS, 1991). This steadily worsening situation has attracted considerable international attention, most recently from the intergovernmental Subsidiary Body for Scientific, Technical and Technological Advice (SBSTTA) to the Convention on Biological Diversity. The SBSTTA recognized that increasing taxonomic capacity is a *sine qua non* for the implementation of the Convention, and has recommended the strengthening of infrastructure for taxonomy in biodiversity-rich tropical countries, together with the transference from developed countries of modern technologies for taxonomy (UNEP/CBD/SBSTTA/2, 1996) in order to provide a basis for the monitoring, inventory making and sustainable utilization of biological diversity (DI CASTRI et al., 1992; JANZEN, 1993).

The problem of insufficient resources being available for the identification of arthropods is further aggravated by the taxonomic community themselves. Their efforts are often primarily focused on areas where intellectual debate can easily be conducted, such as phylogenetic reconstruction. The more mundane tasks of monographic, floristic and faunistic studies are less attractive, both to scientists, and to funding agencies. Furthermore, traditional 'applied' taxonomic products, printed keys, are often almost impossible to use without both adequate reference collections and an extensive knowledge of arcane specialist terminology. Consequently, even where the literature to identify organisms exists, many biologists cannot and do not use it (GAULD, 1986; TILLING, 1987; ALBERCH, 1993). In the jargon of the marketplace, the products of the taxonomic community are generally not appropriate for the needs of the potential user community.

In attempts to rectify this situation and overcome the resulting 'taxonomic impediment', traditional taxonomic products are beginning to be augmented by the use of computerized multi-access keys, beginning with text-based keys (e.g. PANKHURST, 1978) and culminating recently in multimedia works such as CABI-KEY (WHITE and SCOTT, 1994). Whilst undoubtedly an advance over hard copy works, computerized keys still rely on the ability of users to compare pictorial information with specimens. Such skills are honed by years of practice in taxonomists, but other biologists often experience great difficulty in appreciating the subtle differences in shape and form which discriminate taxa, particularly among many groups of invertebrates. Using computers to present taxonomic characters, while relying on users to compare specimens to images or illustrations, represents a failure fully to utilize the immense potential offered by information technology.

Image analysis techniques and technology, in particular, have seen tremendous advances in recent years, raising the possibility of automating, or at least semi-

 0931–2048/99/2301–0001 $ 14.00/0

automating, much of the process of routine taxonomic identification. However, such an approach has to date only been used in a very limited fashion (WEEKS and GASTON, 1997). Thus, for example: (i) DALY et al. (1982) used image analysis to measure 25 morphometric characters of honey bees; the origin of the bees, either European or African, was determined by discriminant analysis; (ii) ZHOU et al. (1985) used image analysis methods to describe the venation of mosquito wings by fitting the coefficients of polynomials to each vein; and (iii) YU et al. (1992) measured the wings of ichneumonids by semi-automated image analysis; discriminant analysis was used to identify five species, on the basis of differences in their wings, achieving 100% accuracy. Such studies encourage the belief that image analysis techniques may represent a way forward towards a large-scale taxonomic identification system based on computer vision, but as yet no such system exists.

In an attempt to begin to bridge this gap, we have developed a prototype novel automated identification system (DAISY – the Digital Automated Identification SYstem) the technical details of which have been fully described elsewhere (WEEKS et al., 1998), which built on recent developments in human face detection (TURK and PENTLAND, 1991). Published tests of the system thus far have revealed that it can discriminate between five species of ichneumonid wasps, on the basis of wing structure alone, with 94% accuracy in correct identifications and a reasonably predictable pattern of errors (WEEKS et al., 1997). In this paper two things are attempted. First, the ability of DAISY to discriminate amongst an order of magnitude more species of an entirely different group of organisms, biting midges, again based on differences in wing structure is examined. Second, in so doing an attempt is made to establish a set of basic 'benchmark' tests of the efficacy, and weaknesses of such an automated identification system.

2 Materials and methods

2.1 The study organisms

The study was based on wings of specimens of biting midges of the genera *Culicoides* and *Forcipomyia* (Dipt., Ceratopogonidae). Wing pattern is very important in the taxonomy of the genus *Culicoides*. However, a quantitative representation of wing pattern has only rarely been achieved (LANE, 1981). Forty-nine species were selected: *Culicoides achrayi* Kettle & Lawson, *C. aitkeni* Wirth & Blanton, *C. brevipalpis* Delfinado, *C. brevifrontis* Smatov & Isimbek., *C. brosseti* Vattier & Adam, *C. brucei* Austen, *C. brunnicans* Edwards, *C. cataneii* Clastrier, *C. circumscriptus* Kieffer, *C. citroneus* Carter, Ingram & Macfie, *C. confusus* Carter, Ingram & Macfie, *C. cornutus* de Meillon, *C. cubitalis* Edwards, *C. dekeyseri* Clastrier, *C. delta* Edwards, *C. distinctipennis* Austen, *C. duddingstoni* Kettle & Lawson, *C. dzhafarovi* Remm, *C. engubandei* de Meillon, *C. exspectator* Clastrier, *C. fascipennis* Staeger, *C. fulvithorax* Austen, *C. furcillatus* Call., Kremer & P., *C. furens* Poey, *C. gambiae* Clastrier & Wirth, *C. gejgelensis* Dzhafarov, *C. grahamii* Austen, *C. grisescens* Edwards, *C. hortensis* Khamala & Kettle, *C. imicola* Kieffer, *C. impunctatus* Goetghebuer, *C. kingi* Austen, *C. krameri* Clastrier, *C. kurensis* Dzhafarov, *C. lailae* Khalaf, *C. langeroni* Kieffer, *C. odibilis* Austen, *C. pallidicornis* Kieffer, *C. praetermissus* Carter, Ingram & Macfie, *Forcipomyia biannulata* Ingram & Macfie, *F. bipunctata* Linnaeus, *F. castanea* Walker, *F. nigra* Winnertz, *F. phlebotomoides* Bangerter, *F. psilonota* Kieffer, *F. pulchrithorax* Edwards, *F. radicicola* Edwards, *F. sphagnophila* Kieffer, *F. titillans* Winnertz.

2.2 Data acquisition

Microscope slides of the right fore wings of 20 specimens (a mixture of males and females) of each of the test species were placed on the transmitted light stage of a Zeiss Stemi SV11 Apo stereomicroscope with planachromat S 1.0× objective. A Kontron ProgRes 3000 colour CCD camera was mounted on the microscope's TV camera tube with integral 0.5× objective. A video sample from the camera (viewed on an adjacent monitor), allowed the scale and orientation of the wing image to be manipulated prior to image capture. The size of the wing was altered using the zoom control on the microscope, such that the wing image almost filled a rectangle of size 540 × 320 pixels. The wing was orientated by adjusting the microscope slide until the anterior margin (which was assumed to be straight) was parallel with the *x*-axis. Once focused, an image was captured and stored to disk on a personal computer running an image analysis software package (KS400; Kontron Elektronik GmbH, Munich, Germany). The red and blue components of the captured colour image were immediately discarded, while the green component, which when isolated is transposed to a greyscale image, was corrected for shading using a previously captured shading reference image. Greyscale images comprise integer pixel values in the range 0–255. The light intensity and other optical settings remained constant throughout. The stored images were cropped such that the wings completely filled a boundary rectangle. Images were then reduced in size, maintaining the original aspect ratio, to 150 × 100 pixels. The images were reduced in order to make the computation of principal components tractable and they were then preprocessed to caricature the venation and pigmentation patterns on the wings (WEEKS et al., 1997).

2.3 Principal component imagery

Digitized wing images were rearranged into column vectors consisting of concatenated rows of pixel intensities. Thus, any particular image k was represented by a column vector $\mathbf{a_k}$ of dimensionality $I \times 1$ (where I equals the width times the height of the wing image in pixels). A set of K wing images of the same species were arranged into an $I \times K$ matrix $\mathbf{A}$, such that $\mathbf{a_k}$ was the kth column of $\mathbf{A}$. The average wing image of the set $\mathbf{a}$ was subtracted from each image in $\mathbf{A}$. Matrix $\mathbf{A}$ was then subject to principal components analysis (PCA), which computes a set of orthogonal eigenfunctions $\mathbf{p_1}, \mathbf{p_2}, \ldots \mathbf{p_i}$ which characterize the modes of variation of the wing images in $\mathbf{A}$. Eigenfunctions are ordered such that the first eigenfunction $\mathbf{p}_1$ accounts for the largest amount of variation, the second eigenfunction $\mathbf{p}_2$ for the second largest and so on.

Any image in $\mathbf{A}$ may be reconstructed exactly as a linear combination of the I eigenfunctions and eigenvalues of $\mathbf{A}$. Furthermore, images in $\mathbf{A}$ may approximately be reconstructed using only the eigenfunctions with the largest eigenvalues, say the first K eigenfunctions

$$\mathbf{a_k} \approx \mathbf{a'} = \bar{\mathbf{a}} + \mathbf{P}\mathbf{b_k} \quad (1)$$

where $\mathbf{a'}$ is an estimate of $\mathbf{a_k}$, $\mathbf{P}$ is an $I \times K$ matrix of eigenfunctions of $\mathbf{A}$, and $\mathbf{b_k}$ is a vector of eigenfunction weights which describe the contribution of each eigenfunction in representing the image $\mathbf{a_k}$. Eigenfunction weights are calculated from the scalar product of the image $\mathbf{a_k}$ and each eigenfunction as follows:

$$\mathbf{b_k} = \mathbf{P^T}\,(\mathbf{a_k} - \bar{\mathbf{a}}) \quad (2)$$

Since the first K eigenfunctions account for the most variance within matrix **A**, the error between the original and approximated image is minimized.

2.4 Image reconstruction metric

Consider an ensemble of images of the right fore wings of species α. Applying PCA to these images, as described above, yields a set of principal eigenfunctions which best describe the variation within this ensemble. If variation due to rotation, translation and scaling is accounted for, then the set of eigenfunctions generated by PCA may be regarded as a basis set with which to describe the wing morphology of α. Now, if an input image, $\mathbf{a_1}$, of a specimen of α is reconstructed in terms of this basis set of eigenfunctions, then the approximate reconstruction will be visually very similar to the original input image ($\mathbf{a_1}$). However, if an input image $\mathbf{a_2}$, of a specimen of a different species is reconstructed in terms of these eigenfunctions, then the image will be reconstructed poorly, since its form is not well described by the basis set of α. It follows therefore that the difference between the reconstruction of $\mathbf{a_1}$ and the original image $\mathbf{a_1}$ will be small, while the difference between the reconstruction of $\mathbf{a_2}$ and the original image $\mathbf{a_2}$ will be larger.

The Kendall-τ statistic (PRESS et al., 1994) and a simple vector-difference metric were used as nonparametric tests of the difference between reconstructed and original images. The measures return a value of 1.0 for perfect correlation and 0.0 for no correlation. The Kendall-τ statistic is:

$$c_\alpha = K\tau(\mathbf{a_i}\ \text{reconstructed}, \mathbf{a_i}\ \text{original}) \quad (3)$$

where $K\tau$ is the Kendall-τ rank difference metric, c_α is a measure of the correlation of image $\mathbf{a_i}$ with the images of wings of species α, and $\mathbf{a_i}$ is the ith pixel.

The vector-difference metric is:

$$c_\alpha = |(\mathbf{a_i}\ \text{reconstructed})^2 - (\mathbf{a_i}\ \text{original})^2| \quad (4)$$

where both reconstructed and original images are vector normalized prior to this calculation.

Essentially, a species classifier may be built from the first five to 10 principal eigenfunctions of a training set of wing images of a particular species. Input wing images of the same species as those used to train a classifier, return high correlation coefficients when compared with that classifier. Input wing images of a different species, return lower correlation coefficients. Generating multiple species classifiers and comparing an input wing image with each of them facilitates identification; the classifier to which the input image correlates most strongly, i.e. that producing the highest correlation coefficient, is predicted as the 'correct' species.

This is an acceptable identification scheme providing all the input wing images are from species upon which the classifiers have been trained. Wing images of species for which no trained classifier exists still generate correlation coefficients, although these are generally low. These images may be discriminated by setting a threshold correlation coefficient below which images may be regarded as unknowns. The setting of threshold values has been discussed elsewhere (WEEKS et al., 1997).

2.5 Species identification using image reconstruction

The viability of this approach to species-level identification of biting midges was assessed using the database of wing images. For each species, 11 of its 20 specimen images were designated 'training' images with the remainder designated 'test' images. Species classifiers were generated for each of the 49 species using randomly chosen training images of the same species. The test images were correlated with each of the classifiers, with the highest correlation predicted as the correct identification.

The effect on the proficiency of identification of using different specimens to train species classifiers was assessed by using randomly picked training images to populate the training sets. The size of image training sets was also varied. The effect on identification of the number of species included was assessed by incrementing the number of classifiers (2–49) and reprocessing the test images.

In addition to this 'first-past-the-post' analysis, DAISY's identification performance was monitored in alternative ways. First, the test image was only deemed to have been identified correctly if its correlation with the 'correct' species classifier was a certain magnitude greater than its correlation with any other classifier. The number of correct identifications will tend to decrease when this disparity is larger, but confidence in a predicted identification will increase. Second, the test image was only deemed to have been identified correctly if its correlation with the 'correct' species classifier was ranked in at least the top 10 correlation values. Using this method, the number of correct identifications will increase, while the identity of a test image will have been narrowed down from 49 possible species to up to 10 species.

Finally, the mean correlation of each test image with the 10 species classifiers of the genus *Forcipomyia* and with the 39 species classifiers of the genus *Culicoides* was determined. The mean of these means was calculated with test images of the same species. This gives an indication of the degree of clustering of each species with the two genera. This may allow images to be discriminated into their respective genera prior to being identified to species.

3 Results

The highest level of accurate identification obtained for the biting midges was about 86% (fig. 1). This was achieved when each of the 49 species classifiers was trained on the largest sized training set used, 11 images. In general, accuracy was found to increase as classifiers were trained on progressively larger numbers of training images, although the magnitude of the improvement was not directly proportional to the increase in training set size (fig. 1), However, even with training sets of only three specimens more than 70% of identifications were correct.

Accuracy also depended on the precise composition of the training set, particularly for smaller sets, for which the standard deviation in the proportion of correct identifications becomes quite marked (fig. 1; note, with 11 images in the training set there is no standard deviation as all training set images are in use).

Increasing the number of species to be discriminated resulted in a decrease in the proportion of correct identifications (fig. 2). With just two classifiers, more than 98% of test images of those species were correctly identified, whilst with all 49 classifiers, the proportion declined to 86%. The form of the relationship between the proportion of correct identifications and the number of species classifiers implies that with yet more classifiers this decline would continue (fig. 2).

The identification results thus far were based on an image being correlated to each of up to 49 species classifiers, with the classifier to which the test images correlate most strongly being deemed the species to which

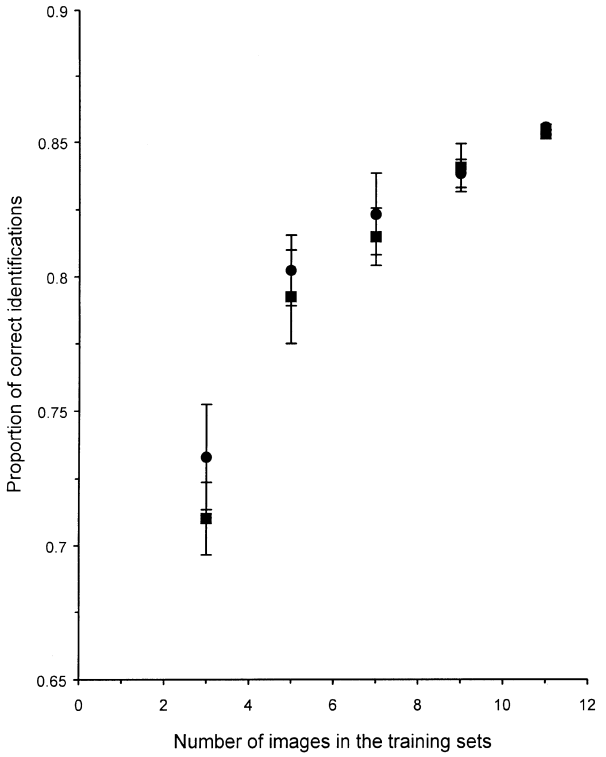

Fig. 1. DAISY identifies correctly a greater proportion of specimens as the number of images used to train the species classifiers increases. (●) *Kendall-τ metric*; (■) *vector-difference metric*

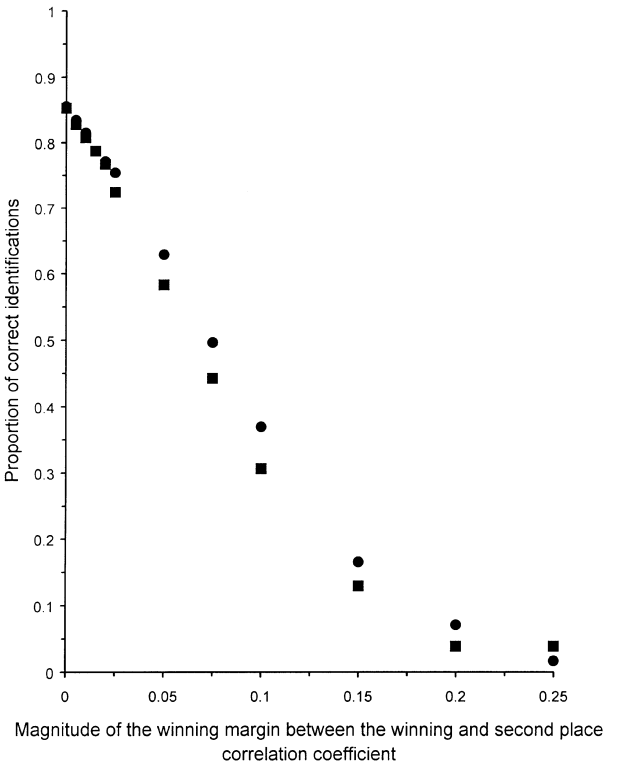

Fig. 3. DAISY identifies correctly a greater proportion of specimens as the magnitude of the winning margin stipulated between the winning and second place correlation coefficient decreases. (●) *Kendall-τ metric;* (■) *vector-difference metric*

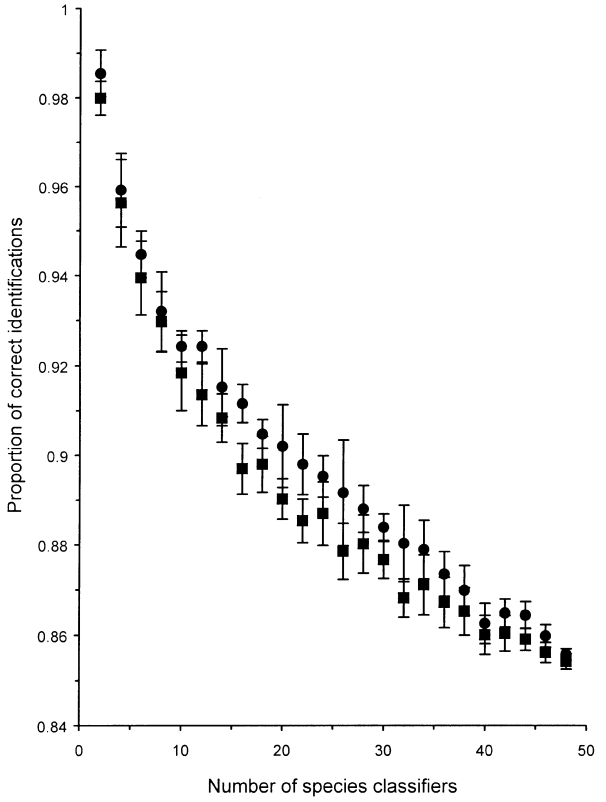

Fig. 2. DAISY identifies correctly a greater proportion of specimens as the number of species classifiers to which specimens belong decreases. (●) *Kendall-τ metric;* (■) *vector-difference metric*

a specimen belongs. This took little account of exactly how well a test image correlated with a particular classifier. Figure 3 shows how the accuracy of identification changed when a winning margin between first and second place classifier correlation coefficients was stipulated. Identification accuracy dropped from the high of 86% to approximately 60% when a winning margin of 0.05 is stipulated, confirming that the wings are extremely similar.

Figure 4 shows the effect on identification accuracy of accepting an identification as correct provided the correlation coefficient with the 'correct' classifier is ranked first, in the first two, first three and so on. Identification accuracy increased to more than 90% when one specified 'correct' as being one of three possible species. This was of practical importance as it shows there is considerable potential for using DAISY to reduce a set of possible identities from 49 to a very few.

Figures 5 and 6 show the degree of correlation of the test images, grouped within their species, with the classifiers representing the genera *Culicoides* and *Forcipomyia*. Both figures demonstrate that images of species within the genus *Culicoides* are highly correlated with classifiers representing that genus, while images of species within the genus *Forcipomyia* are highly correlated with classifiers representing that genus.

Throughout the above analyses, the proportion of correct identifications was slightly higher using the Kendall-τ metric rather than the vector-difference metric

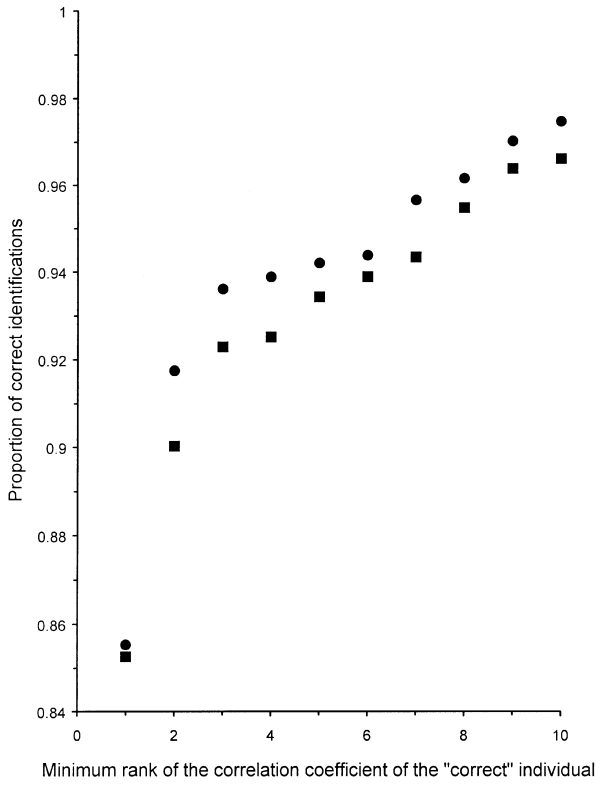

Fig. 4. *DAISY identifies correctly a greater proportion of specimens when the rank of the correlation of a specimen with its 'correct' species classifier is considered. If the correlation is ranked at least first, second or third and so on, a specimen is considered to have been correctly identified.* (●) *Kendall-τ metric;* (■) *vector-difference metric*

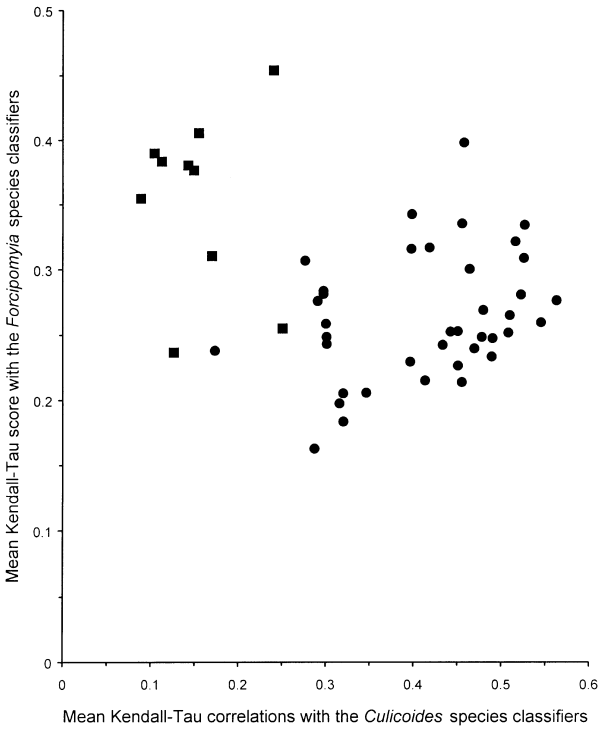

Fig. 5. *Using the Kendall-τ metric, images of specimens within the genus* Culicoides *are highly correlated with classifiers representing that genus, while images of specimens within the genus* Forcipomyia *are highly correlated with classifiers representing that genus.* (●) *Species of the genus* Culicoides; (■) *species of the genus* Forcipomyia

(figs 1, 2, 3, 4). However, the latter was substantially faster to compute.

4 Discussion

These results provide some support for the notion that the approach to automated identification of organisms embodied in DAISY is a useful one. The overall level of 86% successful identification of the 441 specimens of 49 species of biting midges is encouragingly high. This is particularly so given that whilst wing pattern has been used extensively in the taxonomy of the genus *Culicoides*, it has not provided an absolute means of distinguishing between species (LANE, 1981). Furthermore, if the data are divided into their respective genera and reprocessed, the 39 species classifiers representing the genus *Culicoides* achieve 89% successful identification (higher than predicted: fig. 2). Thus, expanding DAISY to include species of a visually similar genus served only to reduce the efficacy of identification.

This level of successful identification was achieved with species classifiers trained on only 11 specimens of each species, far less than the number of specimens usually available for many species where there is a high demand for identification. Using more specimens would better represent the phenotypic variation present in a

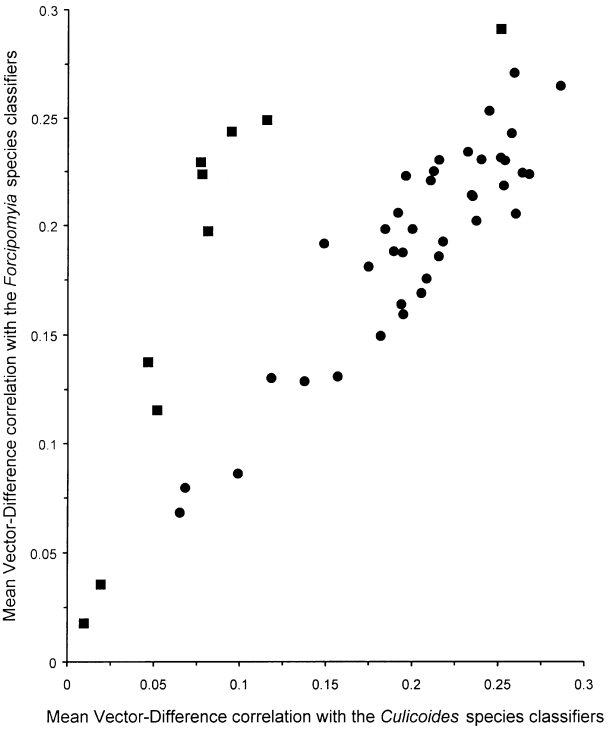

Fig. 6. *Using the vector-difference metric, images of specimens within the genus* Culicoides *are highly correlated with classifiers representing that genus, while images of specimens within the genus* Forcipomyia *are highly correlated with classifiers representing that genus.* (●) *Species of the genus* Culicoides; (■) *species of the genus* Forcipomyia

species. These results suggest that the level of accuracy would increase were classifiers to be trained on more specimens (fig. 1), although the improvements may not necessarily be dramatic. Whilst, the training of species classifiers takes longer with more training images, there is no time penalty associated within the actual identification phase. Thus, if more training images are available they should be used. Equally notable, is the relatively high frequency of correct identifications achieved even with species classifiers trained on only three specimens ($>70\%$; fig. 1). This is far fewer than the number of specimens which many taxonomists feel confident about using as a basis for discriminating species.

Moreover, these levels of correct identification can be achieved relatively quickly. The slide mounting of sufficient wings to establish a single species' training set (11 specimens) takes approximately 30 min, while imaging those wings takes less than 10 min (and generates a lasting data set). Training of a species classifier takes a few seconds. One of the significant modifications that has been made in recent versions of DAISY is in the speed of analysis. Once a specimen's wing has been detached, mounted, imaged and processed (approximately 4 min), the predicted specific identity may be determined from up to 49 species classifiers in a total of less than 3 s.

An overall figure of 86% successful identifications is obviously too low in itself to be of great practical value. Significant additional problems would also appear to be posed by two factors. First, the decline in the proportion of correct identifications as the number of species classifiers is increased (fig. 2) implies that the system has limits. This decrease in the proportion of correct identifications with increasing number of species to be discriminated results from an increasing overlap of characters. Unfortunately, the effective separation of visually very similar objects using a visual identification system will not always be successful.

Second, the narrowness of the 'winning margin' for correct identifications (fig. 3) suggests that the difference between a correct and an incorrect identification is typically very small. Whilst a reflection of the high level of similarity of many such closely related species, this is an undesirable property for an identification system to possess. However, on the positive side, the system provides great promise for very accurately reducing the set of possible identities of any test specimen, from all, to one of a very few species (fig. 4).

The challenge in developing an automated identification system is not to correctly identify specimens more often than not. Rather, it is to attain frequencies of correct identification that mean that economic pests can be recognized during quarantine inspection, that a given species is recognized by a single unique epithet throughout its range, and that nonspecialists can identify many of the most common components of their local faunas, thus contributing to knowledge of patterns of biodiversity (WHITTEN, 1996). To many systematists, little short of 100% accuracy is acceptable, but in reality, at species level, this is unlikely to be achieved – as perusal of specimens identified in the past and preserved in museum collections will show! Even when working with a well-known fauna, expert taxonomists do not attain such levels of accuracy, as deformed, undersized or probable hybrid individuals cause problems. Nor is a 100% accurate species identification necessarily what a user requires. A quarantine officer will certainly want to know if a fly infesting a cargo is a New World Screwworm, but if it cannot be identified to species, it is just as important for her or him to know it is not this pest but a member of a genus of dungflies. It is perhaps most realistic to argue that what is important is to avoid incorrect determination. Thus if a specimen to be identified closely resembles three extremely similar species, it is better to say it is a member of this species-complex than it is to wrongly assign it to one species of the three. It is noteworthy, that as DAISY presently uses only one character set, a wing, the use of other character sets may offer ways of resolving such problems. Thus DAISY may have a practical application as part of an identification system, eliminating large numbers of highly improbable species and reducing the final identification to a choice between three or four species, which may then be discriminated by the user examining other features, such as genitalia. Whatever, using the one feature, wings, DAISY has discriminatory powers that are as good as or better than many expert taxonomists. In blind tests, one of the authors (I.D.G.) with considerable experience of the taxa to be identified (GAULD, 1991), and working only with wing slides of five species of pimpline ichneumonid, achieved a lower rate of accurate identification than the 94% achieved by DAISY (WEEKS et al., 1997).

5 Future directions

Accepting that, although not without problems, DAISY potentially offers a way of circumventing the taxonomic impediment, the question arises, how might the level of correct identifications provided by DAISY markedly be improved? Several methods suggest themselves:

(i) One possible way to discriminate very similar specimens is to use local feature analysis. This method attempts to find differences between images which emerge at the local level, such as subtle differences in the shape or pigmentation of the pterostigma. Local feature analysis would identify specimens in the same way as DAISY presently does, but instead of using PCA components which are holistic, emergent local feature maps could be generated from the PCA components in the manner described by PENIO and ATICK (1996).

(ii) An alternative or perhaps complementary method would be to train a neural network on the correlation coefficients produced by specimens of the same species. In this way a specimen may not necessarily produce the highest correlation, however, the pattern of its correlations with the other classifiers may indicate its species.

(iii) The likelihood of successful identification is strongly influenced by the quality of the images of specimens captured at the outset (see also WEEKS et al., 1997). Careful re-imaging of specimens which have previously been incorrectly identified can often subsequently yield correct identifications. Obtaining

images when specimens are appropriately orientated and illuminated, in particular, is important. An objective method of capturing images in a more consistent fashion would potentially improve overall levels of correct identification markedly. One such method involves automatically extracting whole wings from captured images using active contour snakes (CURWEN et al., 1991). Once a wing is demarcated in this way, its rotation, orientation and scale may be readily recorded and standardized, thereby dispensing with the difficulties associated with manually aligning wings. Removing the necessity to precisely align specimen slides, will move the process a step closer to developing an identification system of practical value.

(iv) The decline in the level of correct identifications with greater numbers of species classifiers (fig. 2), combined with the tendency of specimens to be more strongly correlated with classifiers for the genus to which they belong than to ones for a genus to which they don't (figs 5 and 6), suggests that a structured hierarchical approach to identification may be more appropriate. If classifiers were trained on a selection of images representing species of the same genus, it may be possible to identify specimens to genera, using genus classifiers, and then, using only the appropriate species classifiers, to species. Depending on its success, this scheme could be extended to many taxonomic levels. Of course, as with traditional keys, a specimen misidentified at, say, family level would stand no chance of being correctly identified.

(v) As a last resort, if a specimen can only be narrowed down to, for example, one of three species, the user may always refer to the original specimen where a convenient or even obvious character may be used to discriminate between species. Whilst this is perhaps an undesirable end-point, providing DAISY provides the user with sufficient taxonomic information to finalize an identification then this may be deemed acceptable.

The discrimination of closely related species of most groups of organisms is not a simple task, even for expert taxonomists, and it would be foolish to expect to develop a 'perfect' automated system without many iterations of testing and modification. However, as a first step down this road, the results obtained from DAISY provide an encouraging start.

Acknowledgements

We are grateful to J. BOORMAN for the use of the slides of biting midges. This work was supported by AFRC grant 40/A01764. K. J. GASTON is a Royal Society University Research Fellow.

References

ALBERCH, P., 1993: Museums, collections and biodiversity inventories. Trends Ecol. Evol. **8,** 372–375.

CHALMERS, N. R., 1996: Monitoring and inventorying biodiversity: collections, data and training,. In: Biodiversity, science and development. Towards a new partnership. Ed. by DI CASTRI, F.; YOUNÈS, T. Wallingford: CAB International, 171–179.

CURWEN, R. M.; BLAKE, A.; CIPOLLA, R., 1991: Parallel implementation of Langrangian dynamics for real-time snakes. In: Proc. Brit. Mach. Vis. Conf. (Glasgow). Ed. by MOUFORTH, P. London, Springer Verlag, 29–35.

DALY, H. V.; HOELMER, K.; NORMAN, P.; ALLEN, T., 1982: Computer-assisted measurement and identification of honey bees (Hymenoptera. Apidae). Ann. Entomol. Soc. Am. **75,** 591–594.

DI CASTRI, F.; ROBERTSON VERNHES, J.; YOUNÈS, T., 1992: Inventorying and monitoring Biodiversity. Biol. Int, Special Issue **27,** 1–28.

GASTON, K. J.; MAY, R. M., 1992: Taxonomy of taxonomists. Nature **356,** 281–282.

GAULD, I. D., 1986: Taxonomy, its limitations and its role in understanding parasitoid biology. In: Insect parasitoids. Ed. by WAAGE, J.; GREATHEAD, D. London: Academic Press, 1–21.

—, 1991: The Ichneumonidae of Costa Rica, 1. The subfamilies Rhyssinae, Pimplinae, Poemeniinae, Acaenitinae and Cylloceriinae. Mem. Am. Ent. Inst. **47,** 1–589.

HAWKSWORTH, D. L., 1994: The identification and characterisation of pest organisms. Wallingford: CAB International. 486pp.

HOLDEN, C., 1989: Entomologists wane as insects wax. Science **246,** 754–756.

HOUSE OF LORDS, 1991: Systematic biology research. Report of the Select Committee on Science and Technology. London: HMSO. 106 pp.

JANZEN, D. H., 1993: Taxonomy: universal and essential infrastructure for development and management of tropical wildland biodiversity. In: Proc. Norway/UNEP Expert Conference on Biodiversity, Trondheim, Norway. Ed. by SANDLUND, O. T.; SCHEI, P. J. Trondheim: NINA, 100–113.

LANE, P. L., 1981: A quantitative analysis of wing pattern in the *Culicoides pulicaris* species group (Diptera, Ceratopogonidae). Zoo. J. Linn. Soc. **72,** 21–41.

LATTIN, J. D.; KNUTSON, L., 1982: Taxonomic information and services on arthropods of importance to human welfare in Central and South America. FAO Plant Protection Bull. **30,** 92–95.

MAY, R. M., 1990: Taxonomy as destiny. Nature **347,** 129–130.

PANKHURST, R. J., 1978: Biological Identification. London: Arnold. 104 pp.

PENIO, P. S.; ATICK, J. J., 1996: Local feature analysis: a general statistical theory for object representation. Network: Computation Neural Systems **7** (3)**,** 477–500.

PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P., 1994: Numerical recipes in C. Cambridge: Cambridge University Press.

REID, W. V.; LAIRD, S. A.; MEYER, C. A.; GÁMEZ, R.; SITTENFELD, A.; JANZEN, D. H.; GOLLIN, M. A.; JUMA, C., 1993: Biodiversity prospecting: using genetic resources for sustainable development. Washington, DC: World Resources Institute.

TILLING, S. M., 1987: Education and taxonomy: the role of the Field Studies Council and AIDGAP. In: Nature, natural history and ecology. Ed. by BERRY, R. J.; CROTHERS, J. H. London: Academic Press, 87–96.

TURK, M.; PENTLAND, A., 1991: Eigenfaces for recognition. J. Cogn. Neuroscience **3,** 71–86.

UNEP/CBD/SBSTTA/2 . . . , 1996: Report of the Subsidiary Body on Scientific, Technical and Technological Advice on the work of its second meeting. Montreal: Secretariat to the Convention on Biological Diversity.

WEEKS, P. J. D.; GASTON, K. J., 1997: Image analysis, neural

networks, and the taxonomic impediment to biodiversity studies. Biodiver. Conserv. **6,** 263–274.

WEEKS, P. J. D.; GAULD, I. D.; GASTON, K. J.; O'NEILL, M. A., 1997: Automating the identification of insects: a new solution to an old problem. Bull. Entomol. Res. **87,** 203–211.

WEEKS, P. J. D.; O'NEILL, M. A.; GASTON, K. J.; GAULD, I. D., 1998: Species-identification of wasps using principal component associative memories. Image and Vision Computing.

WHITE, I. M.; SCOTT, P. R., 1994: Computer information resources for pest identification: a review. In: The identification and characterisation of pest organisms. Ed. by HAWKSWORTH, D. L. Wallingford: CAB International, 129–137.

WHITTEN, A., 1996: Field guides: useful tools in environmental planning and management. World Bank Environ. Department, Dissemination Notes. Washington DC, World Bank, **51,** 1–4.

YU, D. S.; KOKKO, E. G.; BARRON, J. R.; SCHAALJE, G. B.; GOWEN, B. E., 1992: Identification of ichneumonid wasps using image analysis of wings. Syst. Ent. **17,** 389–395.

ZHOU, Y. H.; LING, L. B.; ROHLF, F. J., 1985: Automatic description of the venation of mosquito wings from digitized images. Syst. Zool. **34,** 346–358.

Authors' addresses: Dr I. D. GAULD (corresponding author), Dr P. J. D. WEEKS, Department of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD UK; Dr M. A. O'NEILL, Oxford Orthopaedic Engineering Centre, Nuffield NHS Trust, Windmill Road, Oxford OX3 7LD, UK; Dr K. J. GASTON, Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK