Zootaxa 2217: 43–55 (2009)
www.mapress.com/zootaxa/

Forum

ISSN 1175-5326 (print edition)
ZOOTAXA
ISSN 1175-5334 (online edition)

Accelerating taxonomic discovery through automated character extraction

JOHN LA SALLE[1], QUENTIN WHEELER[2], PAUL JACKWAY[3], SHAUN WINTERTON[4], DONALD HOBERN[5], DAVID LOVELL[6]

[1] *CSIRO Entomology, GPO Box 1700, Canberra, ACT, 2601, Australia; E-mail: john.lasalle@csiro.au*

[2] *International Institute for Species Exploration, School of Life Sciences, Arizona State University, PO Box 876505, Tempe, Arizona, 85287-6505, USA; E-mail: Quentin.Wheeler@asu.edu*

[3] *CSIRO Mathematical and Information Sciences, Queensland Bioscience Precinct, 306 Carmody Road, St Lucia, Queensland, 4067, Australia; E-mail: Paul.Jackway@csiro.au*

[4] *University of Queensland Insect Collection, School of Biological Science, University of Queensland, St Lucia, Queensland, 4072, Australia; E-mail: wintertonshaun@gmail.com*

[5] *Atlas of Living Australia, CSIRO Entomology, GPO Box 1700, Canberra, ACT, 2601, Australia; E-mail: Donald.Hobern@csiro.au*

[6] *CSIRO Mathematical and Information Sciences, GPO Box 664, ACT, 2601, Australia; E-mail: David.Lovell@csiro.au*

Abstract

This paper discusses the following key messages. Taxonomy is (and taxonomists are) more important than ever in times of global change. Taxonomic endeavour is not occurring fast enough: in 250 years since the creation of the Linnean *Systema Naturae*, only about 20% of Earth's species have been named. We need fundamental changes to the taxonomic process and paradigm to increase taxonomic productivity by orders of magnitude. Currently, taxonomic productivity is limited principally by the rate at which we capture and manage morphological information to enable species discovery. Many recent (and welcomed) initiatives in managing and delivering biodiversity information and accelerating the taxonomic process do not address this bottleneck. Development of computational image analysis and feature extraction methods is a crucial missing capacity needed to enable taxonomists to overcome the taxonomic impediment in a meaningful time frame.

Key words: taxonomy, taxonomic impediment, automated character extraction, image analysis, feature extraction, pattern recognition

Introduction

There is nothing new about taxonomy: it is the oldest of all the sciences. However, there is a growing feeling that taxonomy is now more important than ever, particularly as we need to understand enough about ecosystem function to make informed natural resource management decisions in an era of global change. There is clearly a need to shift taxonomic endeavour into the digital era to improve the pace at which we can supply taxonomic products and information. Recent volumes and papers have stressed these points, and offered compelling suggestions about how to accelerate taxonomic productivity (Godfray 2002a,b; Godfray & Knapp 2004; Wilson 2004; MacLeod 2007; Wheeler 2008).

These works recommend critical steps forward for improving the taxonomic process, and the pace of species description has increased dramatically as a result. Roughly 1.8 million species have been described in 250 years, giving an average rate of about 7200 species/year. Today we average somewhere between 16,000 and 20,000 per year which is more than double the historical average (SOS Report 2009). This is a clear indication that taxonomists are cognizant of the need to increase the rate of species discovery and description, and are working towards that goal.

Accepted by Z.-Q. Zhang: 31 Jul. 2009; published: 2 Sept. 2009