Guest Editorial: Computer Vision for Animal Biometrics

Biometric Computer Vision that detects, tracks, identifies, describes, and classifies animal life from captured image and video data, is an emerging subject in machine vision. It is an exciting moment for this field of study. For the first time a myriad of real-world systems and applications are becoming integrated into the practice of the biological sciences. Indeed, Computer Vision systems have also started to assist work in a variety of allied scientific areas including field ecology, agricultural research, animal welfare, conservation, public health and the behavioural sciences.

This Special Issue brings together a selection of eight timely papers in this field, submitted by researchers of institutions from across four continents. The works presented here include the analysis of animals ranging from insects and fish, to birds and mammals, but also life during embryonic development. The presented papers showcase some of the current diversity in this research domain: methods span the whole spectrum from traditional feature-based approaches to Deep Learning solutions.

Robota et al. introduce “Tracking of group-housed pigs using multi-ellipsoid expectation maximization” describe a system that utilises depth images accurately to estimate the position and orientation of individual pigs in a group-housed environment over significant periods of time. By applying expectation maximization as a policy for ellipse fitting, their method is able to exploit consistent shape and fixed target numbers to aid tracking. Practical results demonstrate that the system can track 15 group-housed pigs for an average of 19.7 minutes between failure events.

Xie et al. in their paper “A novel open snake model based on global guidance field for embryo vessel location” present a framework for blood vessel region extraction and accurate snake-based localisation in imagery of animal embryos. Their open snake model utilises a global guidance field and is initialised by a deformation template. Experimental results on a specific embryo vessel database demonstrate that the proposed algorithm can robustly locate the embryo’s blood vessels and obtain orientations of the vessel branches. Comparisons with traditional methods illustrate the effectiveness and competitiveness of their proposed model.

Bakkay et al. in their paper “Automatic detection of individual and touching insects from trap images by combining contour-based and region-based segmentation” introduce a method for the detection of insects from camera trap images in difficult conditions by employing an innovative region merging algorithm and an adaptive k-means clustering approach, operating on the object contour’s convex hull. Quantitative evaluations show that the proposed method can detect insects with higher accuracy than most widely used approaches.

Erola et al. in their paper “Automatic individual identification of Saimaa ringed seals” describe a method for the automatic image-based individual identification of endangered Saimaa ringed seals (Phoca hispida saimensis) that exploits the species’ permanent and individually unique visual pelage patterns. The proposed framework performs segmentation of the seals from the background, as well as post-processing and classification steps required for identification. Two existing individual identification methods are compared to the presented work using a challenging data set of Saimaa ringed seal images. The results show that the proposed segmentation and post-processing steps are effective and can provide increased identification performance against a generic baseline.

Akkaya et al. in their paper “Mouse face tracking using a convolutional neural network” present a convolutional neural network (CNN) tracker called MFTN for following a mouse’s face in video footage. Notably, in the proposed architecture, target information is extracted from a combination of low- and high-level features by a particular sub-network to achieve a more robust and accurate tracker. Experiments show that the particular MFTN/c tracker achieved an accuracy of 0.8, a robustness of 0.67, and a throughput of 213 fps on the GPU-powered testing workstation.

Beyan et al. in their paper “Extracting statistically significant behaviour from noisy fish tracking data” describe an approach to the cleaning of a large and noisy visual tracking dataset to allow for the extraction of statistically sound results from the underlying image data. In particular, the paper presents an analysis of a dataset of 3.6 million underwater trajectories of a species of fish, which are also labelled with the water temperature at the time of acquisition. By a combination of data binning and robust estimation methods, the authors demonstrate reliable evidence for an increase in fish speed as water temperature increases. Several statistical tests applied to the data confirm that results are statistically significant.

Ardo et al. in their paper “A CNN-based cow interaction watchdog” introduce an automated video analysis system for the processing of cow footage that can select or discard recorded video material based on user-defined criteria commonly required in behavioural research to reduce the amount of time experts have to spend on watching video. A CNN architecture is proposed and then evaluated in a pilot study. It is shown that 38% (50% with additional filler parameters) of the recordings in the test dataset could be correctly and successfully removed, while only losing 1% (4%) of the potentially interesting video frames.

Finally, Silla Junior et al. in their paper “Bird and whale identification using sound images” describe a novel approach for the automated identification of birds and whales from calls. The visual features of audio used are constructed from different spectrograms and from harmonic and percussion images. These images are then divided into sub-windows from which sets of texture descriptors are extracted for classification. The experiments reported in this paper use a dataset of bird vocalizations targeted for species recognition and a dataset of right whale calls targeted for whale detection, as well as three well-known benchmarks for music genre classification. The authors demonstrate that the fusion of different texture features, as well as texture and audio features together can enhance performance.

As is clear from the above content, this Special Issue highlights the great breadth of research in Visual Animal Biometrics today, and the even greater potential for this area of Computer Vision in the future. One may argue that the field is indeed on its way to realising another facet of Jim Gray’s 4th scientific paradigm, ever more intricately binding together biological research questions with Computer Vision engineering.

In any case, we hope that readers will find the papers put forward here inspiring and informative; and we would like to extend our sincere thanks to all authors and reviewers of the works before us.

Guest Editors

Tilo Burghardt, University of Bristol, (tilo@cs.bris.ac.uk)