Where to look when identifying roadkilled amphibians?

Marc Franch1,2,*, Cristiano Silva3, Gil Lopes3, Fernando Ribeiro3, Paulo Trigueiros3, Luis Seco1,4, Neftalí Sillero1

1 CICGE (Geo-Space Sciences Research Centre), Prof. Manuel de Barros Astronomical Observatory, University of Porto, (Portugal).
* Corresponding author. E-mail: apoarmatu@gmail.com
2 Department of Animal Biology (Vertebrates), University of Barcelona, (Spain)
3 School of Engineering of the University of Minho, Algoritmi Research Center, Guimarães (Portugal)
4 ISMAI, University Institute of Maia, Maia (Portugal)

Submitted on 2015, 27th July; revised on 2015, 18th October; accepted on 2015, 18th October
Editor: Marco Sannolo

Abstract. Roads have multiple effects on wildlife; amphibians are one of the groups more intensely affected by road-kills. Monitoring roadkills is expensive and time consuming. Automated mapping systems for detecting roadkills, based on robotic computer vision techniques, are largely necessary. Amphibians can be recognised by a set of features as shape, size, colouration, habitat and location. This species identification by using multiple features at the same time is known as “jizz”. In a similar way to human vision, computer vision algorithms must incorporate a prioritisation process when analysing the objects in an image. Our main goal here was to give a numerical priority sequence of particular characteristics of roadkilled amphibians to improve the computing and learning process of algorithms. We asked hundred and five amateur and professional herpetologists to answer a simple test of five sets with ten images each of roadkilled amphibians, in order to determine which body parts or characteristics (body form, colour, and other patterns) are used to identify correctly the species. Anura was the group most easily identified when it was road-killed and Caudata was the most difficult. The lower the taxonomic level of amphibian, the higher the difficulty of identifying them, both in Anura and Caudata. Roadkilled amphibians in general and Anura group were mostly identified by the Form, by the combination of Form and Colour, and finally by Colour. Caudata was identified mainly on Form and Colour and on Colour. Computer vision algorithms must incorporate these combinations of features, avoiding to work exclusively in one specific feature.

Keywords. Roadkills, amphibians, jizz, computer vision, algorithms, survey methodology.

INTRODUCTION

Roads have multiple effects on wildlife, such as animal mortality, habitat and population fragmentation, as well as modification of animal behaviour (Trombulak and Frissell, 2000; Jaeger and Fahrig, 2004; Jaeger et al., 2005; Epps et al., 2007; Lengagne, 2008; Fahrig and Rytwinski, 2009). Amphibians in particular, due to their activity patterns, population structure, and preferred habitats, are strongly affected by traffic intensity and road density and die massively on roads (Fahrig et al., 1995; Glista et al., 2008; Gryz and Krauze, 2008; Sillero 2008; Garriga et al., 2012).

Monitoring roadkills is expensive and time consuming, and depends mainly on volunteers (Grilo and Ascensão, 2011; Garriga et al., 2012; Beebee, 2013). Surveys can be performed by car (Santos et al., 2007; Sillero, 2008; Garriga et al., 2012; Matos et al., 2012), bike (Ashley and Robinson, 1996; Collinson, 2013), or by foot (Ashley and Robinson, 1996; Collinson, 2013; Ruiz-Capillas et al., 2015). Cheap, easy to implement, and automatic methods for detecting roadkills over larger areas (broad monitoring) and along time (continuous monitoring) are necessary. In this sense, Mapping Systems provide the capacity to detect automatically the casualties of roadkills by intelligent algo-
algorithms (Chambon and Moliard, 2011; Varadharajan et al., 2014), improving current monitoring systems.

In the context of the research project Roadkills (PTDC/BIA-BIC/4296/2012), funded by Fundação para a Ciência e a Tecnologia (FCT) of Portugal, we developed a Mobile Mapping System installed on a trailer and composed of standalone power generation, computer imaging capturing, and recording under a meter GPS receiver and one linear controlled and standardized lighting. To capture images, we adapted a line scan camera, commonly used in industry and in controlled environments. This camera model has an optical resolution between 250 µm/pixel at 35 km/h, 500 µm/pixel at 70 km/h and 1000 µm/pixel at 140 km/h.

The camera acquire sequential images of 4096 pixels width and one pixel length (approx. 1.0 m × 0.25 mm). The construction of an image or a frame is obtained through the acquisition of several lines that the line scan camera captures. The obtained frame (as a result of the sum of several sequential lines) is the work unit of the algorithms with 4096 pixels width and a variable number of lines as its length.

This new methodology improves passive sampling as car speed can be higher than currently used in visual surveys (20 km/h) (Sillero, 2008; Garriga et al., 2012; Matos et al., 2012). The proposed Mobile Mapping System can be implemented on road maintenance vehicles and on surveillance or maintenance vehicles in protected areas without an expert for their correct use. Analysis may be in real time from software resident in the Mapping System itself or by transferring to a remote server.

These systems are based on robotic computer vision techniques such as object recognition or structure from motion and are widely used in many computer vision applications. Computer vision techniques usually involve three distinct steps: feature detector, description and matching (Moreels and Perona, 2007; Dernth, 2014). Most of the proposed features provide both a detector and a descriptor algorithms, which can be combined each other like in SIFT (Scale-Invariant Feature Transform), SURF (Speeded Up Robust Features), BRIEF (Binary Robust Independent Elementary Features) or ORB (Oriented and Rotated BRIEF), all implemented in the OpenCV library (Rublee et al., 2011). OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products (Yu et al., 2004).

Bag of words or Bag of features is one of the popular visual descriptors used for visual data classification. It is a sparse vector of occurrence counts of a vocabulary of local image features (Gang, 2012). Overall, the algorithms work looking for the detected patterns in the “problem” images in known patterns from images library. Initially, known patterns of the library are provided by us. Once designed the algorithm and running, it is self-sufficient in images that stock the library (with our supervision).

Computer vision algorithms try to emulate human vision when recognising an object or a part of it. In herpetology, amphibians are not recognised immediately by a unique feature, but by a set of particular and subtle characteristics, together with knowledge of the species (Arnold and Ovenden, 2003). Professionals and amateurs cannot clearly define what these particular features are in order to identify an amphibian (Ellis, 2011). The informal term among naturalists and scientists of this integrated identification form is “jizz” and is based on shape, size, colouration, habitat and location (Macdonald, 2002; Daston, 2008; Ellis, 2011). Currently there are no data about the particular features on amphibians’ “jizz” although this identification way is rather trivial.

Which criteria we follow, as experts, at the moment of identification? On which basis do we identify a road-killed amphibian? Species identification of the roadkilled amphibian may be very difficult due to the corpse state (Ashley and Robinson, 1996; Glista et al., 2008). Moreover, there is a gradient from intact corpses to completely unidentifiable carcasses depending on the violence of the impact with the vehicle, the time of permanence on the road and the weather or environmental conditions (Glista et al., 2008). Despite this, how “jizz” works in expert identification? Are all used characters subtle or do some characteristics emphasise over others? Is there a prioritisation process in the moment to generate an idea about the identification? Thus, it is necessary to understand how we prioritise these features when deciding if a road-killed amphibian is in fact a roadkilled amphibian (e.g. a roadkilled slag) or not. Computer vision algorithms will work in a similar way to human vision. They will use our prioritisation process when analysing an image.

In order to answer these questions, we asked to several amateur and professional herpetologists to answer a simple test composed of numerous pictures of roadkilled amphibians. Specifically, we aimed to determine which parts or characteristics of the roadkilled amphibian body are used by experts to identify correctly the species, when a specific part of the body is essential to identify a species, or by the contrary, when experts use all the amphibian body to obtain a final conclusion. Thus, we wanted to give a numerical priority sequence of particular characteristics to improve the computing and learning process of algorithms.

Our results will be essential to improve the ability of algorithms applied to images obtained for the Mobile
Mapping System developed for the automatic detection of amphibian roadkills on roads. When algorithms are developed, they will be tested under controlled conditions (roadkilled amphibians' 3D plastic models) and on the field (real roadkilled amphibians) with and without the obtained identification parameters.

**MATERIAL AND METHODS**

**Structure of the test/survey**

The test consisted of five sets with ten images each and an associated questionnaire. The test was built with Google Sheets, which is able to provide reports and data tables. All images were collected in the Iberian Peninsula (mainly Catalonia and Portugal). Besides amphibians, images also included some reptiles, invertebrates, birds, and mammals, in order to include false positives. Thus, we included 50 images (Table 1), 43 with roadkilled amphibians and seven with other groups (four reptiles: one *Lacerta bilineata*; three *Podarcis sp.*; two small mammals, *Talpa europaea* and *Rattus sp.*; and one invertebrate, an indeterminate caterpillar) in order to maintain the participant's attention and avoid trends in their responses. We included pictures of 29 anura (15 *Bufo bufo* and 14 caudata (seven *Salamandra salamandra*), proportion corresponding to relative species abundances (Matos et al., 2012). Two images of *B. bufo* were very difficult to identify. We obtained the test images from our own repository and the University of Barcelona. We assigned a random order to the images to avoid possible trends of subjective grouping. Test images were not possible to expand. Although for some collaborators this was a problem, our idea was that the user tries to identify in the same conditions as the computer algorithm.

We included a brief introductory letter on the front cover of the first questionnaire (with the presentation of the survey). At the beginning of each set of images, identification was requested, as anonymous or identified contributor.

The questionnaire for each image was composed of six questions (Suppl. Mat. Appendix A1-Survey). The first question (Q1): Is it possible for you to identify a roadkilled amphibian in this image? was a binary question ('yes' or 'no') with the possibility of including comments. If the user answered 'no', the questionnaire for this image was finished. If the user answered 'yes', the test proceeded to the second question (Q2): Which is the taxonomic level do you get? Taxonomical identification was possible to levels of Order, Genus and Species. Selecting an option in a higher taxonomic level did not conditioned options at lower taxonomic levels. The third question (Q3): Which features did you use for the identification? had multiple choices: Form, Colour, Others Patterns, Form & Colour, Form & Others Patterns, Colour & Others Patterns, or Form & Colour & Others Patterns. They could select one of three options or a combination of them. In the following three questions (Q4: Form: general body shape; head; forelimbs; hind-limbs; tail or other characters; Q5: Colour: general design pattern; specific colours; specific designs or other characters; and Q6: Others Patterns: recognition of bones / carcasses; recognition of skins or other characters.), the user was able to select only one option. When selecting the option “Other characters”, the user must specify the criteria in a text box, or any other comment.

Once each set was finished, the results were recorded and the link to the next set appeared. In the final set, we added a field where the user can introduce additional comments, views, or constructive reviews to the test.

**Table 1. List of species and number of test images for each species.**

<table>
<thead>
<tr>
<th>Order</th>
<th>Genus</th>
<th>Species</th>
<th>Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caudata</td>
<td><em>Salamandra</em></td>
<td>salamandra</td>
<td>7</td>
</tr>
<tr>
<td>Caudata</td>
<td><em>Triturus</em></td>
<td>marmoratus</td>
<td>6</td>
</tr>
<tr>
<td>Caudata</td>
<td><em>Pleurodeles</em></td>
<td>walli</td>
<td>1</td>
</tr>
<tr>
<td>Anura</td>
<td><em>Bufo</em></td>
<td>bufo</td>
<td>15</td>
</tr>
<tr>
<td>Anura</td>
<td><em>Bufo</em></td>
<td>calamita</td>
<td>4</td>
</tr>
<tr>
<td>Anura</td>
<td><em>Hyla</em></td>
<td>meridionalis</td>
<td>3</td>
</tr>
<tr>
<td>Anura</td>
<td><em>Discoglossus</em></td>
<td>galganoi</td>
<td>2</td>
</tr>
<tr>
<td>Anura</td>
<td><em>Rana</em></td>
<td>iberica</td>
<td>2</td>
</tr>
<tr>
<td>Anura</td>
<td><em>Alytes</em></td>
<td>obstetricans</td>
<td>1</td>
</tr>
<tr>
<td>Anura</td>
<td><em>Discoglossus</em></td>
<td>pictus</td>
<td>1</td>
</tr>
<tr>
<td>Anura</td>
<td><em>Pelophylax</em></td>
<td>perezi</td>
<td>1</td>
</tr>
<tr>
<td>Reptilia</td>
<td><em>Podarcis</em></td>
<td>hispanica</td>
<td>3</td>
</tr>
<tr>
<td>Reptilia</td>
<td><em>Lacerta</em></td>
<td>bilineata</td>
<td>1</td>
</tr>
<tr>
<td>Mammalia</td>
<td><em>Rattus</em></td>
<td>sp.</td>
<td>1</td>
</tr>
<tr>
<td>Mammalia</td>
<td><em>Talpa</em></td>
<td>europaea</td>
<td>1</td>
</tr>
<tr>
<td>Lepidoptera</td>
<td>(Indeterminate Caterpillar)</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Total: 50

**Participants**

The diffusion of the survey was carried from mailing lists of known contacts, calling on ResearchGate (scientific and social networking service), on Facebook (social networking service), from a post in Road Ecology Group (642 members at July of 2014), and through the website, Facebook and Twitter (social networking service) accounts of the Spanish Herpetological Society (AHE). Potential participants were selected for their experience in the study of roadkills. We also offered the possibility to the potential participants of spreading the test to their contacts with experience in his research field.

**Data processing**

The survey was online during six months (July 2014 to January 2015). We obtained the generated tables for each set after closing the site. We extracted proportions of each category/subcategory and applied Chi square test in order to find differences in frequencies between groups and categories.
RESULTS

Hundred and five people answered the survey, of which 63 (60%) identified themselves and 42 (40%) were anonymous. Participants’ contributions were different between sets. The answers to the first sets were higher than for the last sets: from 105 contributors in first set to 53 in the fifth one.

Although Question 1 (Q1: Is it possible for you to identify a roadkilled amphibian in this image?) only supported an affirmative answer if contained a road-killed amphibian, we admitted as affirmative answer in the group of “Fool” pictures if the participants tried to identify the image by comments. For this question we obtained 3241 responses with about 60% correctly identified a roadkilled amphibian in the image. Detailed results for taxonomical groups are found on Table 2. We found highly significant differences in the correct identification among the three groups ($\chi^2 = 54.12$, df = 2, $P < 0.001$). Anura was the group most easily identified as road-killed of the three groups (between Anura and the rest of groups; $\chi^2 = 14.72$, df = 1, $P < 0.001$). Fools were clearly identified as a no roadkilled amphibian (between Fools and the rest of groups; $\chi^2 = 43.91$, df = 1, $P < 0.001$). Caudata was the most difficult group to identify as roadkilled amphibian (no differences were found between Caudata and the rest of groups; $\chi^2 = 0.47$, df = 1, $P = 0.492$).

About 52% of 2776 possible responses in Q2 (In the case of identifying an amphibian, what is the taxonomi c level do you get?) identified the Order in the image, 28% the Genus and 20% the Species. Detailed results are found on Table 3. We found highly significant differences among taxonomic level ($\chi^2 = 855.89$, df = 2, $P < 0.001$) and also among the taxonomical levels of Anura ($\chi^2 = 984.18$, df = 2, $P < 0.001$) and Caudata ($\chi^2 = 30.88$, df = 2, $P < 0.001$), indicating an increase on the difficulty with the taxonomical level. There were significant differences between Anura and Caudata in assigning the order ($\chi^2 = 187.43$, df = 2, $P < 0.001$); assignment was easier for Anura. There were no differences between Anura and Caudata for identifying the genus ($\chi^2 = 0.13$, df = 2, $P = 0.715$). Species were identified with higher difficulty in Anura than in Caudata ($\chi^2 = 20.03$, df = 2, $P < 0.001$).

In Q3 (What have you based for identification? Form, Colour, Others patterns, Form & Colour, Form & Others patterns, Colour & Others patterns or Form & Colour & Others patterns), we have obtained 1755 responses, 1253 for Anura and 412 for Caudata (Table 4). We found highly significant differences among identification elements ($\chi^2 = 1784.60$, df = 6, $P < 0.001$).
general, roadkilled amphibians were mostly identified by the Form, in a lesser extent by the combination of Form and Colour, and finally by Colour. The other elements and combinations had much lower frequencies. Anura presented the same pattern ($\chi^2 = 1366.55$, df = 6, $P < 0.001$). For Caudata, the identification was based mainly on Form and Colour and on Colour in less proportion. The other elements and combinations had much lower frequencies. There were highly differences in frequency ($\chi^2 = 517.65$, df = 6, $P < 0.001$).

In Q4 (Form: General Body Shape; Head; Fore-limbs; Hind-limbs; Tail or Other Characters) we have obtained 1381 responses, 1025 for Anura and 278 for Caudata (Suppl. Mat. Table T1). We found highly significant differences among identification Form features ($\chi^2 = 2995.32$, df = 5, $P < 0.001$), mainly by the General Body Shape, and Hind-limbs in lesser importance. The other features had much lower frequencies. The same on pattern for Anura ($\chi^2 = 2283.24$, df = 5, $P < 0.001$) as well as for Caudata ($\chi^2 = 788.10$, df = 5, $P < 0.001$).

We obtained 883 responses in Q5 (Colour: General Design Pattern; Specific Colours; Specific Designs or Other Characters), 565 responses for images of Anura and 291 responses for Caudata (Suppl. Mat. Table T2). We found highly significant differences among identification Colour features ($\chi^2 = 611.04$, df = 3, $P < 0.001$) and for Anura ($\chi^2 = 581.33$, df = 5, $P < 0.001$), based mainly on General Design Pattern. For Caudata, we also found highly significant differences ($\chi^2 = 206.13$, df = 5, $P < 0.001$), but based mainly in Specific Colours.

In Q6 (Others Patterns: Recognition of Bones / Carcasses; Recognition of Skins or Other Characters), we obtained 391 responses, 326 for Anura and 42 for Caudata (Suppl. Mat. Table T3). We found significant differences among Others Patterns identification features ($\chi^2 = 526.89$, df = 2, $P < 0.001$), as well as for Anura ($\chi^2 = 431.35$, df = 2, $P < 0.001$) and Caudata ($\chi^2 = 71.36$, df = 2, $P < 0.001$). For all of them, participants identified roadkilled amphibians by Other Patterns based on Recognition of Skins.

Fig. 1. Radar graphic of the “jizz” sequence for the different parameters of Form, Colour and Other Patterns together for general amphibians’ roadkills, for anurans and for caudata.
We extracted the prioritisation sequence based on the percentage of all well identified images (Figure 1). General sequence for amphibians’ roadkills is:

\[ JIZZ = 0.3858 \times F + 0.2974 \times F&C + 0.1476 \times C + 0.0672 \times F&C&OP + 0.0433 \times F&OP + 0.0393 \times OP + 0.0194 \times C&OP \]

where
F (Form); C (Colour) and OP (Other Patterns)

Regarding the sequence for the different parameters of Form, Colour and Other Patterns together:

\[ JIZZ = 0.3686 \times GBS + 0.1669 \times GDP + 0.1390 \times HL + 0.0970 \times SC + 0.0725 \times RS + 0.0602 \times SD + 0.0247 \times H + 0.0205 \times FL + 0.0138 \times T + 0.0095 \times RB/C + 0.0266 \times OC \]

where
GBS (General Body Shape); GDP (General Design Pattern); HL (Hind-limbs); SC (Specific Colours); RS (Recognition of Skins); SD (Specific Designs); H (Head); FL (Forelimbs); T (Tail); RB/C (Recognition of Bones / Carcasses) and OC (Other Characters)

Detailed sequences for Anura and Caudata and their specific characters can be found in the supplementary files (Suppl. Mat. Appendix A2-Jizz Sequence).

From this sequence we extracted, in general terms, the high importance of Form (alone or in combination with colour) for a roadkilled amphibian identification. For specific patterns, the main one was GBS (General Body Shape: 36%) followed by GDP (General Design Pattern: 16.7%).

DISCUSSION

Almost 60% of the Q1 responses were correct. Anura was the group most easily identified as roadkilled (73%), because the large hind-limbs of frogs and toads. Results on body features confirmed this conclusion (see below). The correct identification of salamanders and newts was more difficult (55%) because they have similar fore and hind-limbs. In fact, the tail was the second more important feature for their recognition (see below). Participants clearly identified fools as non roadkilled amphibians (almost 90%). Therefore, participants were much more effective in identifying correctly what was not a roadkilled amphibian. This is a surprising result, because there are many features on roads that can be confounded with a roadkilled amphibian (e.g. small mammals, snails, tree sheets, rubbish). However, non-amphibian roadkills can be very difficult to identify correctly from inside a car but not from a close sight.

Participants identified correctly in 61% cases the order of the animals, but genus and species below 30% and 21%, respectively. This rank of taxonomical levels appeared also in Anura and Caudata. Therefore, the lower the taxonomic level of amphibian, the higher the difficulty of identifying them, both in Anura and Caudata. However, the correct assignment of the order and species was easiest in Anura than in Caudata, confirming the easiest identification of anurans. The difficulty in assigning the genus was similar for both groups of amphibians. This lack of differences between both groups could be due to the lower number of genera in Caudata.

Roadkilled amphibians in general and Anura group were mostly identified by the Form, in a lesser extent by the combination of Form and Colour, and finally by Colour. The correct identification of Caudata group was based mainly on Form and Colour and on Colour in less proportion. These results confirmed that people in fact uses “jizz” to identify roadkilled amphibians. There was not one unique feature to get the correct identification, but a combination of features, although the most important was the form for Anura but Form and Colour for Caudata. This last result can be due to the easy identification of salamanders because of their clear pattern of black and yellow colours. Colours alone were not sufficient as diagnostic feature as there are many species with similar or cryptic colours.

In consequence, roadkills were correctly identified by General Body Shape and Hind-limbs in Anura and Caudata groups, by General Design Pattern in Anura and by Specific Colours in Caudata, as well as by Recognition of Skins in both groups. Again, these results confirmed the use of “jizz” to recognise herps. Also, Q5 answers were coherent with Q4 ones. Here, amphibian skin properties were useful to distinguish from fools.

Not all the people answered all the sets of the test due to limitations on survey design. We should not have included anonymous people as this reduces the sample size of our analysis about the effectiveness on answering depending on the profession and skills of the participants.

Our results may be better if we had included more images of other taxa (like insects, reptiles, birds and mammals), objects, vegetables, and other variables related to images for improving the test. The obtained results probably would be more robust but it was not the goal of the survey: we aimed to obtain a sequence of features and characteristics to identify a roadkilled amphibian.

Main conclusions

“Jizz” is the method used to identify herps. There is not a unique feature for the correct diagnoses of amphib-
ians but a combination of several features, depending mainly on the shape of the body and colours. More importantly, different combinations of shapes and colours are essential to identify one species from others. Computer vision algorithms must incorporate these combinations of features to have success, avoiding to work exclusively on one specific feature.

ACKNOWLEDGMENTS

Special thanks to all respondents (Suppl. Mat. Appendix A3-Contributors / Respondents) and to all who have contributed to diffusion and promotion of the Roadkills survey: Road Ecology Group on Facebook, Asociación Herpetológica Granadina, Asociación Herpetológica Española (AHE), Herpnet and all individuals who have also believed appropriate the disclosure of the survey. Thanks to an anonymous reviewer for English linguistic advising and for the comments and suggestions that greatly improved the last version of the manuscript. M. Franch and C. Silva are funded by a grant (UMINHO/BI/175/2013, UMINHO/BI/172/2013) (from the Foundation for Science and Technology Portugal (FCT) and co-financed by FEDER through COMPETE –POFC (“Intelligent systems for mapping amphibian mortality on Portuguese roads”, PTDC/BIA-BIC/4296/2012). N. Sillero is supported by an FCT contract IF/01526/2013.

REFERENCES


