Sinimbu - Multimodal queries to support biodiversity studies

Gabriel de S. Fedel[1], Claudia Bauzer Medeiros[1], and Jefersson A. dos Santos[1]

Institute of Computing, University of Campinas, Brazil
`gabrielfedel@gmail.com,cmbm@ic.unicamp.br,jsantos@ic.unicamp.br`

Abstract. Typical biodiversity information systems can only solve a small part of user concerns. Available query mechanisms are based on traditional textual database manipulations, combining them with spatial correlations. However, experts need more complex computations – e.g., using non-textual data sources. This involves a considerable amount of manual tasks, to obtain the needed information. This paper presents the specification and implementation of Sinimbu – a framework to process multimodal queries that support both text and images as search parameters, for biodiversity studies, thus providing support for subsequent complex simulations. Sinimbu was validated with real data from our university's Zoology Museum, which houses one of the largest zoological museum collections in Brazil. Not only can users interact with the system in several modes, but query possibilities (and answers) vary according to the user's profile. Query processing in Sinimbu combines work in database management, image processing and ontology construction and management.

Keywords: Biodiversity data management, CBIR, Ontologies

1 Introduction

The goal of biodiversity information systems is to help experts manage information on live organisms and their relationships with each other and the environment. This forms the basis for complex computational processes and simulations. From a high level point of view, one can say that there are two main kinds of primary data handled by such systems – (a) data on observations of species, and (b) data about the (geographical) environment in which these species were observed. Observation data contains information about *what* species were observed *where*, *when*, *how*, and *by whom*.

Queries and models in biodiversity systems require correlation of these data, using many kinds of knowledge on geographic, biologic and ecological issues. Available biodiversity systems can only cope with part of these factors, and scientists must perform several manual tasks to derive the desired information, e.g., because of semantic mismatches among data sources, or lack of appropriate operators. Data extraction for model construction, on such systems, is made via queries that are basically of two kinds: (a) those that require textual input (e.g., species' taxonomic classifications) and (b) those that allow direct interaction with some kind of cartographic representation (usually resulting from a textual-based query). There are however many other kinds of data that are used by scientists in their analyses, but which are not taken advantage of in biodiversity systems. Examples include images (in particular, photos of animals or plants), sound recordings (of animals), sketches, etc.

This paper presents *Sinimbu* – a multimodal query processing framework that extends queries on observation data by combining standard text-based queries with ontology manipulation (for semantic enhancement) and query by image content. Sinimbu means *chameleon* in tupi-guarani[1], indicating that the system supports several interaction modes and accommodates distinct user profiles – both scientists and non-experts. Once relevant data are identified and retrieved, scientists use the corresponding records in their simulations – e.g., to study species' interactions in a given region,

Sinimbu uses real data, and was implemented within the BioCORE project[2], a joint research effort from computer scientists and biologists to design tools to support biodiversity studies. Previous results of ours within this effort include the construction of an ontology service that supports several kinds of operations on ontologies, the development of a set of tools that allow querying [8], the development of a set of tools that allow querying inter-species relationships [9], a query expansion toolkit for species occurrence data, also based on ontological knowledge [19], and an infrastructure for content-based image retrieval, tested on fish image collections [7].

This work is being developed at the Laboratory of Information Systems, Institute of Computing at the University of Campinas, Brazil. The main contributions of this paper are: (1) it shows how to combine distinct query processing techniques and special purpose data structures to explore multiple modalities, for biodiversity purposes, thereby filtering and restricting relevant data for biodiversity applications; (2) it presents design details and implementation choices, thereby helping those who want to implement similar functionality; (3) it discusses design and implementation challenges associated with these choices.

The rest of this paper is organized as follows. Section 2 presents the concepts needed to understand the text. Sections 3 and 4 present Sinimbu's architecture and implementation. Section 5 discusses related work, and reviews our contributions. Section 6 presents conclusions and ongoing work.

2 Basic concepts

2.1 Observations and occurrence records

Observations of live organisms are at the core of biodiversity studies. In biodiversity systems, such information is stored in so-called *occurrence* or *collection* records. The terms designate any digital record that describes "when, where

---

[1] South America native group of languages

[2] http://www.lis.ic.unicamp.br/projects/biocore

and how" a species (or a set thereof) was observed or collected, and by "whom". One record may refer to several individuals, or even several species, if different organisms were observed at a certain geo-spatial location (e.g., when insects are collected by means of a trap).

Besides biodiversity systems, data on such records can also be obtained from portals, in which institutions publish information about their collections. Such portals play an important role in contributing to biodiversity studies (even though they cannot be considered as biodiversity information systems). Some of these portals belong to museums, where records are frequently connected with a "physical" object, in the sense that the corresponding organism (a "specimen") was actually collected from the field and preserved. Some experts distinguish between "occurrence" (in which observations do not require collecting an organism) and "collection" (associated with some kind of catalog structure). For this reason, from now on this paper will use the term "collection" to denote data on species.

2.2 Content-based image retrieval and Multimodal queries

Image data can be processed and retrieved in several ways – e.g., considering metadata (such as device used), image captions (and hopefully semantics) or, in our case, image content. In the so-called "Content-based image retrieval" (CBIR), the query predicate is an image, and the result is a set of images that are computed to be "most similar" to the input image. CBIR relies on the use of algorithms that can describe and distinguish images based on their content via *descriptors*. A descriptor [7] extracts feature vectors that represent image visual properties (e.g., color, texture, shape) and defines a distance function that is used to determine how similar two images are, given the distance of their feature vectors. From a high level point of view, a descriptor is a pair $< f, v >$ where v is the feature vector, and f the distance function used to compare two vectors. Given a query pattern (usually an image), descriptors can be used to rank sets of images according to their similarity to the query pattern.

Multimodal queries are those that involve more than one mode for defining query parameters, and normally require processing distinct kinds of data types. The goal is to improve the quality of the result, through the combination of distinct kinds of information. Research in multimodal interfaces is increasing, with a conference entirely dedicated to this subject, already in its 13th edition [14]. This conference covers a wide range of subjects, involving multidisciplinary research on multimodal human-human and human-computer interaction, interfaces, and system development. This paper concentrates on system development issues, considering the modalities of textual and image parameters.

3 Multimodal Search in Sinimbu

3.1 Overview

Sinimbu parameters combine text and/or images. Textual input can be of the following kinds: (a) species' common names, and (b) fields from occurrence records, e.g., scientific name, location. The latter are retrieved using standard database queries, whereas common names are processed through navigating in an ontology especially constructed for the system. Images themselves can be query parameters. In this case, the input image is processed using special purpose descriptors, under CBIR mechanisms.

Sinimbu supports two kinds of user profiles: biodiversity researchers, and non-experts; the query interface and results vary according to the profile. This increases the kinds of users (and uses) of the system.

3.2 Data structures

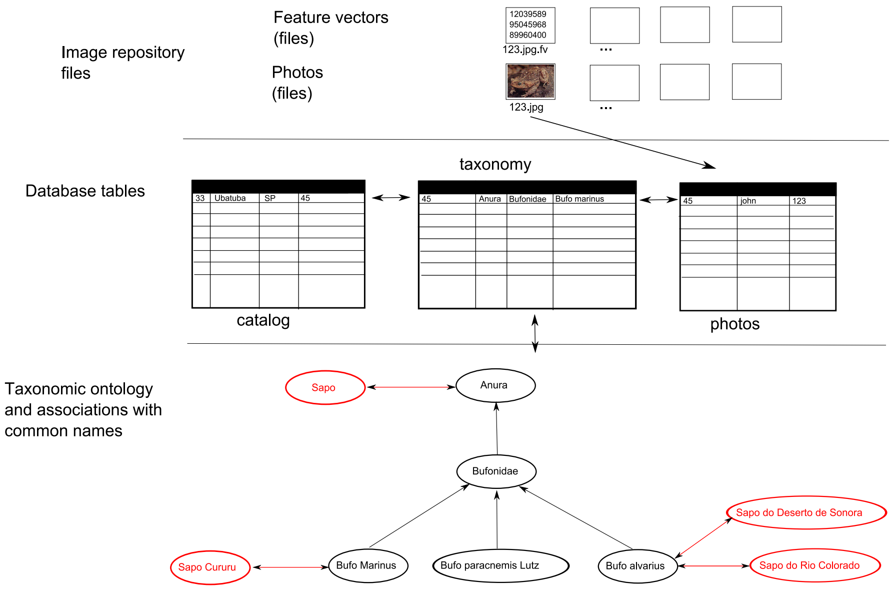


Fig. 1. Data structures in Sinimbu – some database tables and attributes are ommitted

The data used by Sinimbu is organized according to distinct structures, as shown in Figure 1. These structures are basically of three types: tables (using the PostgreSQL database system), an image repository (files of images and their descriptors), and a taxonomic ontology. Tests were restricted to museum curated data.

Database tables used by Sinimbu are the following:

- **taxonomy:** Taxonomic nomenclature of the animals whose occurrence records appear in the database, as well as the year and author of the scientist originally responsible for naming that animal. Several fields may be missing, due to incomplete identification. Its key, *id_taxa* plays a major role in linking all data used by Sinimbu.

- **catalog:** Entries of the museum catalog, containing occurrence records, with information of where, when, how, and who collected the specimen. These records are linked to the *taxonomy* table, uniquely identifying the corresponding animal.
- **photos:** Information on images (photos of animals), also linked to the *taxonomy* table. Contains metadata such as the author of the photo, file name and whether it is public or not (non-public photos are used by content-based retrieval queries, but are not displayed to the general public, only to researchers with appropriate privileges).
- **responsible:** Data on researchers responsible for the field trip in which observations were made. It is linked to the *catalog* table.
- **location:** Textual information about some geographic region, such as country, county or state, being linked to the *catalog* table.

The identifier (primary key) of the records of table *taxonomy* is called *id_taxa*. It is the unique database identifier of some level of taxonomic description for a given species (i.e., the seven basic levels recognized internationally in the Linnaean taxonomic hierarchy for nomenclature – kingdom, phylum/division, class, order, family, genus, species). The values of the *id_taxa* attribute are artificially created, and are unique in the entire database. It is the basis for linking all information about all data associated with a given taxonomic classification – images, occurrence records, etc. For instance, Figure 1 shows there is a unique identifier (45) for the frog "Bufo marinus"; there are 6 such frogs in the University's museum (and thus 6 occurrence records), and 2 photos. All records and photos are linked to *id_taxa*=45. In the *taxonomy* table, record with key 45 contains attribute values "Chordata (Phylum), Amphibia (Class), Anura (Order), Bufonidae (Family), Bufo marinus (the scientific name, Genus + species)".

Sinimbu's structures allow navigating from an image to its descriptors, and vice-versa, and from an image to the corresponding species, and to the associated occurrence records.

3.3 The taxonomic ontology

A taxonomic ontology [5] is a hierarchy of terms, under generalization/specialization relationships. Common names, in Sinimbu, are stored in a taxonomic ontology, associated with the Linnaean scientific taxonomic classification.. A given common name can be associated with distinct taxonomic levels, and a taxonomic term at any level can be linked to multiple common names.

Figure 1 (at the bottom) gives an example of this association: the *Anura* order corresponds to the *sapo* (frog) common name, while species *Bufo alvarius* is associated with *Sapo do Deserto de Sonora* and *Sapo do Rio Colorado*. From a common name, queries can retrieve species' taxonomic classifications (and thus *id_taxa*), and from these find additional information (photos, occurrence data, etc). Also in the example, if a query asks for "Sapo", then all descendants of "Anura" will be selected, because of ontological inheritance.

3.4 Architecture

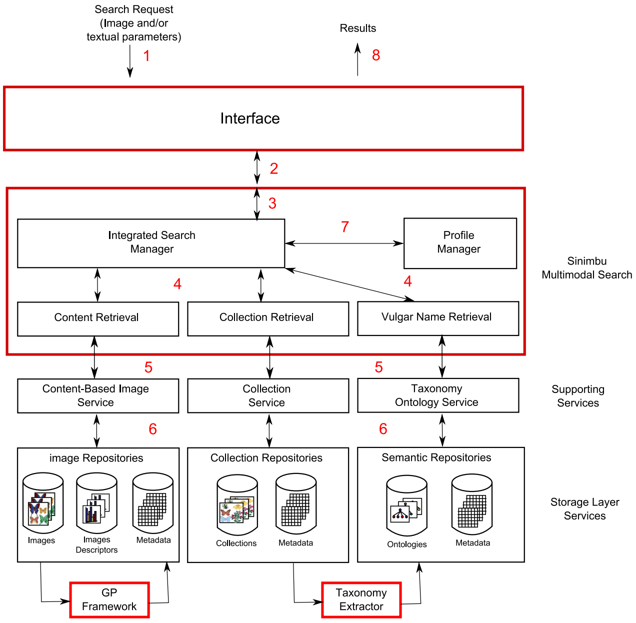


Fig. 2. Sinimbu - main blocks

Figure 2 presents a high level view of Sinimbu's architecture, and its interaction with some of BioCORE's services: Storage services (handle access to repositories), Supporting services (basic operations).

Besides the modules for Multimodal Search, Sinimbu includes the Taxonomy extractor (to build and maintain the taxonomic ontology) and our GP framework [6], used in extracting image descriptors and CBIR processing.

A query is processed as follows: end-users provide input parameters via the interface (images and/or text) – arrow numbered (1) in the figure. This is forwarded to Sinimbu's Multimodal Search - arrow (2), where it is first processed by the Integrated Search Manager (3). This module examines the request, and interacts with distinct Retrieval modules (4), each of which for a specific modality, forwarding query requests to BioCORE's supporting services (5). Once the results are returned, the Integrated Search Manager interacts with the user Profile Manager (7), in which the output visualization is built, forwarding this to presentation via the Interface (8). There are three kinds of data repositories accessed in such queries (6): Image, Collection and Semantic (ontologies).

Sinimbu's Retrieval modules implement specific algorithms that respectively construct requests for CBIR, queries on common names and queries on occurrence records.

3.5 Module invocation

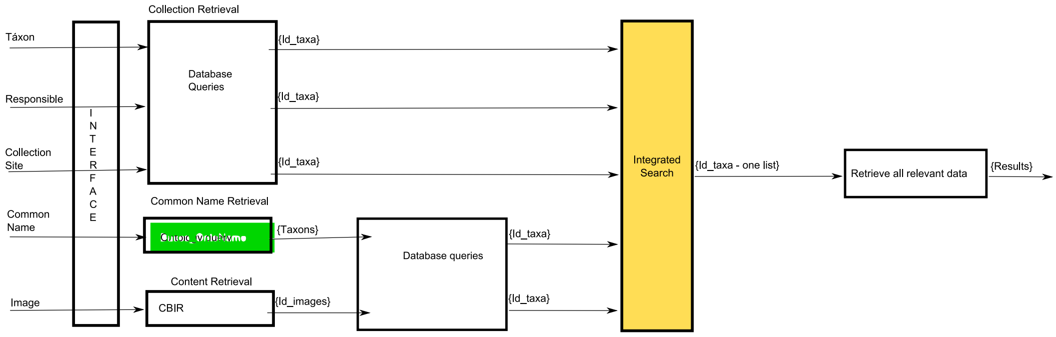


Fig. 3. Module execution flow

Figure 3 shows the invocation of modules (and data flow) within Sinimbu. It shows users can enter images and four types of textual parameters (Taxon, name of person Responsible for a field trip, Collection Site and Common name). These textual parameters were indicated by our biology partners as being the most relevant for their present needs.

The first three parameters are forwarded to the Collection Retrieval Module, where they are processed by functions that construct and invoke SQL commands. SQL queries retrieve the *id_taxa* values of occurrence records whose attributes are substrings of the corresponding textual input parameters. The value of the Taxon parameter will be checked against all taxonomic levels of an occurrence record. For instance, a query with “Taxon = Bufo” will look for records in which the ”Bufo” substring appears as part of attributes Order, Family etc – and thus all species of family=“Bufonidae” will be selected.

Common names are processed by the Common Names Retrieval module. It implements queries on the ontology, and returns a list of Taxon strings. The Content Retrieval module invokes CBIR functions that return a list of identifiers of images similar to the input image, ranked by similarity. Taxons and image identifiers are then directed to the Collection Retrieval module, producing as result a list of *id_taxa* values that correspond to that taxon or that image.

At this point, there are several lists of *id_taxa* values, one for each input parameter. These lists are forwarded to the Integrated Search Module, which will combine these lists using weights defined by the user profile, generating a single list of *id_taxa* values, ranked by weight. This final list is then processed, returning all available information on the corresponding species, according to user profiles.

3.6 Content-based image retrieval module

CBIR processing in Sinimbu uses the GP Framework developed by us, being based on genetic programming principles. Genetic programming (GP) is a machine learning technique which tries to solve problems based on biology evolutive principles. The basic structure in GP is the individual, which represents a possible solution for a given problem. The individuals are programs that during the evolutive process undergo successive recombinations and disturbances and are refined. At the end, better solutions for the identified problems are found. Genetic programming can be understood as a search in the space of all possible solutions for the individuals which best solve the problem.

The GP Framework is a CBIR infrastructure that combines an arbitrary number of descriptors $< f, d >$ to perform CBIR in an image base, combining distances from these descriptors using functions that were previously found to be adequate to discriminate amongst them. To find the best individuals, the framework requires a set of categorized images for training. In our case, the GP Framework was used as follows. First, it preprocessed all images in the image repository using three color descriptors (JAC [20], BIC [16], ACC [13]) and two descriptors that encode texture (LAS [18] and QCCH [12]). Next, it constructed a global distance function D using genetic programming. The training set was a subset of the image repository. The GP individuals were functions that combined the distances among the images, for each descriptor – i.e., the new distance function D is a combination of the distances computed by the five descriptors used. This function is used to compare input (query parameter) images to the images in the repository. Figure 4 shows an example of GP individual as a function to combine descriptors from two images I_j and I_k. This individual corresponds to the function $f(d_{1I_jI_k}, d_{2I_jI_k}, d_{3I_jI_k}) = \frac{d_{1I_jI_k} - d_{3I_jI_k}}{d_{2I_jI_k}} + \sqrt{d_{2I_jI_k} * d_{3I_jI_k}}$.

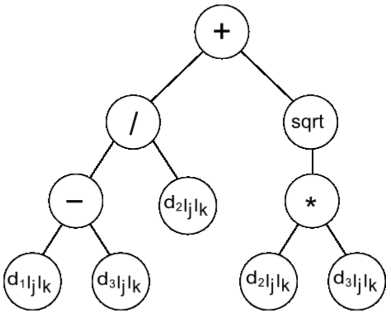


Fig. 4. Example of GP individual.

When a new image is entered as a query parameter, it is processed against all descriptors, extracting its feature vectors under the different descriptor algorithms. Next, its distance to all other images is computed using function D. The result of this comparison is forwarded to the Content Retrieval module, which creates the list of image identifiers ranked according to the similarity with the input image.

3.7 Ontology management modules

The Common Name Retrieval module builds SPARQL queries to process common name requests on the ontology. The Taxonomy Extractor constructs the ontology from data stored in the (*taxonomy*) table (i.e., it is not an exhaustive list of species names, but just those of animals recorded in the database). This occurs every time the database is updated with new species names, upon user request. Common names are next linked to it, using OWL `EquivalentClass` relationships. Building the initial ontology with the common names was a very time-consuming task. Sinimbu common names are in Portuguese, and there is no authoritative source for such correspondences, so that most common names had to be inserted manually (less than 1% of the common names were available from the occurrence records).

4 Implementation

4.1 Data sets and technological issues

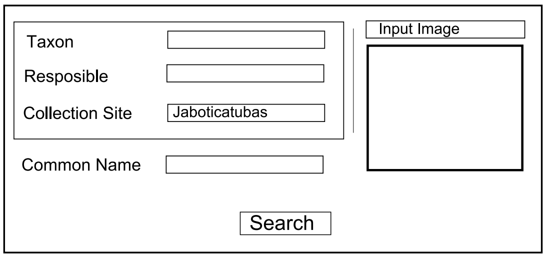
The main data source for Sinimbu are records of the University's Zoology museum, one of the largest physical zoological collections in Brazil. It houses 17 scientific collections – both terrestrial and aquatic animals – with approximately 500 thousand animals (of which roughly 400 thousand invertebrates). Primary data sources to create the occurrence records include field notes, field tags, and notes taken during a field trip. Additional data sources include photos, annotations made by curators, and information on the environment. Data are still being catalogued into the database – at present, there are approximately 60 thousand curated records, plus 1200 curated animal photos, which were used in testing Sinimbu. Though this is not a large image base, its richness lies in its variety of animal species, and on the painstaking curation process each image undergoes to ensure retrieval reliability. The ontology covers 5 thousand species, being generated from the collection repository using the Taxonomy Extraction module. Over 200 common names were inserted, using some sites and documents on Brazilian species.

Occurrence data are stored in the PostGreSQL database system, with the PostGIS extension for geographic processing. This is the database management system chosen for BioCORE, thus allowing integration with the entire system. The ontology was defined in OWL (the standard), and common names were added using the Protégé ontology tool – one of the most widely used ontology management tools. Images are stored in .jpg format, due to its compactness and ease of display; descriptors are stored in binary format, being generated by the GP Framework.

Sinimbu's multimodal search modules were developed in JAVA, again for compatibility with BioCORE. The Content Retrieval module uses JNA (Java Native Access) to process feature vectors and distance functions (themselves implemented in C by the GP Framework). Access to the PostGreSQL database uses JDBC (Java Database Connectivity). The ontology is processed using the JENA ontology framework, which facilitates access to the corresponding structure.

4.2 User session - example

Suppose that a user wants to retrieve data on species that were collected at a place called "Jaboticatubas", as shown in Figure 5. There are 1379 records in the database satisfying this criterion. Figure 6 shows the first three results, corresponding to a hummingbird and two bats. Only one photo is displayed – the first record does not have any photos associated, and the third has photos, but they cannot be shown to non-authorized users. Images in the image repository can be private or public, and only the latter can be displayed as part of a non-expert query result. However, all images are used in CBIR processing.



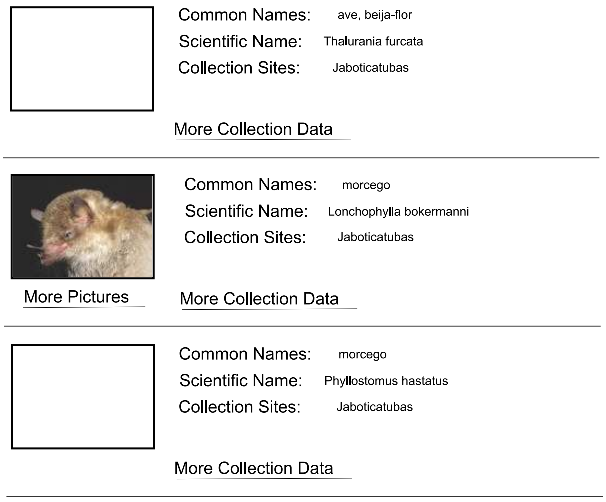
Collection Site: Jaboticatubas

Fig. 5. First query - Animals collected in a given location - "Site=Jaboticatubas"

Refining this search, now the user wants records associated with "Jaboticatubas" and common name "Sapo" (frog) - this query retrieves 1086 records, the first three being shown in Figure 7, now restricted to frogs only. Notice that now the first two records have public images associated.

Suppose that, finally, in this query sequence, the user also includes a query image of a frog – see final query on figure 8. The first three records appear in figure 9, showing that the query image is more similar to that of "Proceratus Cururu". Notice that in this case the first two records of the second query (Figure 7) are inverted, and Proceratus Cururu appears in the first place, since it is more similar to the input image than the second record.

In this sequence of queries, we point out that the result is progressively refined, thereby filtering the number of records to be used in subsequent modeling and simulation computations. Both the second and the third queries return the same amount of records – 1086. However, the order is different (because of image similarity), and for most purposes only the topmost records need to be considered by the experts that are conducting this research. The use of images in


Common Names: ave, beija-flor
Scientific Name: Thalurania furcata
Collection Sites: Jaboticatubas
More Collection Data

Common Names: morcego
Scientific Name: Lonchophylla bokermanni
Collection Sites: Jaboticatubas
More Pictures
More Collection Data

Common Names: morcego
Scientific Name: Phyllostomus hastatus
Collection Sites: Jaboticatubas
More Collection Data

Fig. 6. Partial result of the first query - "Site=Jaboticatubas"

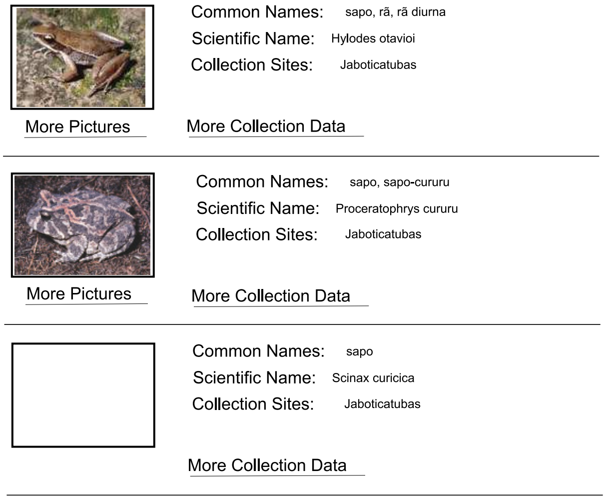
preprocessing data, combined with textual parameters, provide results that are closer to what experts require in order to conduct their studies – images alone, or text alone are not as effective in filtering.

Once the experts determine the most relevant records from the final result, then they can proceed, if desired, to subsequent steps in checking models of habitat conditions, or interactions among species of interest.

5 Related work

Most modern biodiversity information systems now provide Web interfaces, and vary greatly in range of functions supported, geographic regions covered, species groups and goals. Nevertheless, as said in the introduction, they concentrate in supporting access to occurrence data and correlating it to environmental variables. They vary from actual information systems (i.e., having a large set of analysis tools, and distinct kinds of results, such as graphs, tables, maps, charts) to portals from which scientists can download records of interest. GBIF [10] is perhaps the ultimate example of a portal, making millions of occurrence records available, mediating the access to thousands of provider sites.

Multimodality is being increasingly used to improve image retrieval systems – e.g., [4, 1, 2, 15, 17]. In such papers, authors show that a combination of textual and image parameters improves precision and recall of results, as opposed to strictly textual or strictly content-based queries. For instance, the work of Atnafu [4] presents a new data model for images in which retrieval is based on combining descriptors and metadata. Query parameters support combination of image content, and contextual and semantic information. A similar solution is presented by Addis [1], for images in art galleries. Here, users can either pose


Common Names: sapo, rã, rã diurna
Scientific Name: Hylodes otavioi
Collection Sites: Jaboticatubas

Common Names: sapo, sapo-cururu
Scientific Name: Proceratophrys cururu
Collection Sites: Jaboticatubas

Common Names: sapo
Scientific Name: Scinax curicica
Collection Sites: Jaboticatubas

Fig. 7. Partial result - second query - "Site=Jaboticatubas and Common name=sapo"

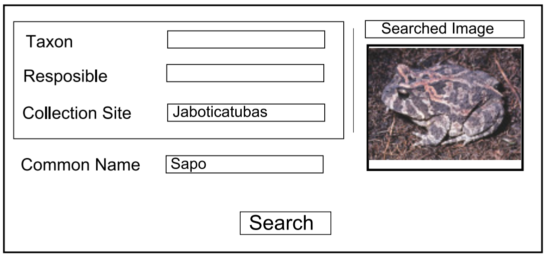
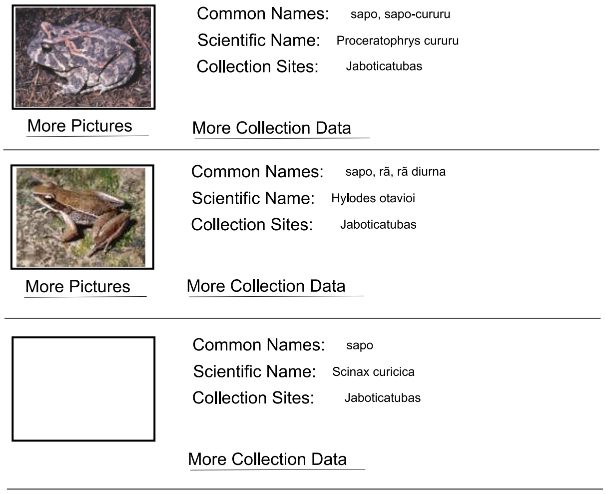

Taxon:
Resposible:
Collection Site: Jaboticatubas
Common Name: Sapo

Fig. 8. Third query - Providing input to Site, Common name and image parameters

queries or navigate interactively among concepts. Multimodal queries can also be processed on videos – e.g., the work of Ammir [2], in which relevance feedback is adopted to improve the results, or on music – e.g. the work of Zhang [21].

Most multimodal systems, however, are not directed towards biology-related data. One of the main problems of such data, when images are involved, is the kind of image available (for instance, when a photo is taken in a natural environment, animals are often partially hidden, and moreover identifying characteristics are seldom available). Thus, most such systems concern images produced in closed environments or laboratories. For instance, C-DEM [11] mines genetic data from fruit flies, in what they name "image bioinformatics". Query parameters combine data from genes (expressed as strings of characters), images of gene expressions and keywords that annotate the images. Users can request images that are similar to a given image, and/or related to the keywords provided, and/or to specific genes. i

Arpah's multimodal system [3] is also geared towards management of biological data, using images from flatworms. Queries combine taxonomic information



Common Names: sapo, sapo-cururu
Scientific Name: Proceratophrys cururu
Collection Sites: Jaboticatubas

Common Names: sapo, rã, rã diurna
Scientific Name: Hylodes otavioi
Collection Sites: Jaboticatubas

Common Names: sapo
Scientific Name: Scinax curicica
Collection Sites: Jaboticatubas

Fig. 9. Partial result of 3rd query - Site, image and common name parameters

on species, stored in an ontology, and image annotations. Over 900 images were annotated, discriminating among species' morphological characteristics. Queries can either access the annotations themselves, or the ontologies, using SPARQL. Here, there is no image processing, but the multimodality refers to a combination of queries on ontologies and text.

Since queries in Sinimbu combine textual, ontological and image information, for biological data, it can be considered closest to C-DEM in modalities and nature. The parameters and goals are, however, very different (since the latter concentrates on genetic information). Moreover, C-DEM images are homogeneous, whereas Sinimbu's repository is very heterogeneous, contemplating all kinds of animals (e.g., vertebrates and invertebrates) and photos taken in controlled (museum) and non-controlled (field trips) environments. Hence, no single descriptor is suitable – and thus we use a combination thereof.

6 Conclusions and Ongoing work

There are many challenges in managing and querying data on species' observations, ranging from the heterogeneity of the observations to the kinds of user that want to access these data. This hampers effective computational modeling and simulation in biodiversity studies. Multimodal query systems increase the flexibility in handling these data, thereby allowing users to better specify their needs and interact with query results. This paper presented the Sinimbu multimodal query framework, in which query predicates combine text and image content, and are processed by a combination of database queries, ontology manipulation and CBIR. Though Sinimbu was validated only on zoological data, its implementation and architecture are general, so that there is no constraint in adding, e.g., botanical data. This would of course require extending the ontology and new tests on the GP Framework to identify a more adequate descriptor combination. Sinimbu has been validated with distinct kinds of user profiles, showing that it can help both experts and non-experts in querying occurrence records.

Visual (image) identification of animal species varies widely – e.g., some can be recognized by shape, others by color, and so on. Sinimbu tried to overcome this limitation by using multiple descriptors, and combining them using genetic programming. This is also a novelty in terms of systems that support queries on images of biological collections, which favor a single descriptor.

To the best of our knowledge, Sinimbu is the only framework that allows multimodal search on collection data, for biodiversity purposes, combining queries on images, text, and ontologies, and for distinct user profiles. Ongoing work includes, among others, enhancing the base of (Brazilian) common names, via Web mining of specific sites, which presents the challenge of homegeneizing such names and finding out the corresponding scientific names. Another direction is increasing customization possibilities based on user contexts.

Acknowledgements The research described in this paper was partially funded by Brazilian agencies CNPq, FAPESP and CAPES, and by the INCT in Web Science.

References

1. M. Addis, M. Boniface, S. Goodall, P. Grimwood, S. Kim, P. Lewis, K. Martinez, and A. Stevenson. Sculpteur: Towards a new paradigm for multimedia museum information handling. In *Proceedings of Semantic Web ISWC 2870*, pages 582–596, 2003.
2. Arnon Amir, Marco Berg, and Haim Permuter. Mutual relevance feedback for multimodal query formulation in video retrieval. In *MIR '05: Proc of the 7th ACM SIGMM international workshop on Multimedia information retrieval*, pages 17–24, 2005.
3. A. Arpah, S. Alfred, L.H.S. Lim, and K.K.S. Sarinder. Monogenean image data mining using Taxonomy ontology. In *Int. Conf. on Networking and Information Technology (ICNIT)*, pages 478 –481, 2010.
4. S. Atnafu, R. Chbeir, and L. Brunie. Efficient content-based and metadata retrieval in image database. *Journal of Universal Computer Science*, 8(6):613–622, 2002.
5. N Cullot, C Parent, S Spaccapietra, and C Vangenot. Ontologies: A contribution to the DL/DB debate. In *Proc. of the 1st International Workshop on the Semantic Web and Databases, 29th VLDB Conf.*, pages 109–129, 2003.
6. R. da S. Torres, A. X. Falcão, M. A. Gonçalves, J. P. Papa, B. Zhang, W. Fan, and E. A. Fox. A genetic programming framework for content-based image retrieval. *Pattern Recognition*, 42(2):283–292, 2009.
7. R. da S. Torres, C. B. Medeiros, M. A. Goncalves, and E. A. Fox. A Digital Library Framework for Biodiversity Information Systems. *International Journal on Digital Libraries*, 6(1):3–17, 2006.
8. J. Daltio and C. B. Medeiros. Aondê: An Ontology Web Service for Interoperability across Biodiversity Applications. *Information Systems*, 33:724–753, 2008.

9. J. Daltio, C. B. Medeiros, L. C. Gomes Jr, and T. Lewinsohn. A Framework to Process Complex Biodiversity Queries. In *Proc. ACM Symposium on Applied Computing (ACM SAC)*, March 2008.
10. GBIF. Global Biodiversity Information Facility Portal. http://data.gbif.org/welcome.htm, accessed June 2011, 2011.
11. Fan Guo, Lei Li, Christos Faloutsos, and Eric P. Xing. C-dem: a multi-modal query system for drosophila embryo databases. *Proc. VLDB Conference*, 1(2):1508–1511, 2008.
12. Chao-Bing Huang and Quan Liu. An Orientation Independent Texture Descriptor for Image Retrieval. In *Int. Conf. on Communications, Circuits and Systems, ICCCAS*, pages 772 –776, 2007.
13. J. Huang, S. R. Kumar, M. Mitra, Wei-Jing Zhu, and R. Zabih. Image Indexing Using Color Correlograms. *IEEE Conf. Computer Vision and Pattern Recognition*, 0:762, 1997.
14. ICMI. International Conference on Multimodal Interaction. http://www.acm.org/icmi/2011/, 2011.
15. Haiyu Song, Xiongfei Li, and Pengjie Wang. Multimodal image retrieval based on annotation keywords and visual content. In *Proc. Int. Conf. on Control, Automation and Systems Engineering*, pages 295 –298, 2009.
16. Renato O. Stehling, Mario A. Nascimento, and Alexandre X. Falcão. A compact and efficient image retrieval approach based on border/interior pixel classification. In *Proc. 11th international Conf. on Information and knowledge management*, CIKM '02, pages 102–109, 2002.
17. Ja-Hwung Su, Bo-Wen Wang, Tien-Yu Hsu, Chien-Li Chou, and Vincent S. Tseng. Multi-modal image retrieval by integrating web image annotation, concept matching and fuzzy ranking techniques. *International Journal of Fuzzy Systems*, 12(2):136–149, 2010.
18. B. Tao and B. W. Dickinson. Texture recognition and image retrieval using gradient indexing. *Journal of Visual Communication and Image Representation*, 11(3):327 – 342, 2000.
19. B. Vilar, J. Malaverri, and C. B. Medeiros. A Tool based on Web Services to Query Biodiversity Information. In *5th International Conference on Web Information Systems and Technologies - WEBIST*, pages 305–310, 2009.
20. Adam Williams and Peter Yoon. Content-based image retrieval using joint correlograms. *Multimedia Tools and Applications*, 34:239–248, 2007.
21. Bingjun Zhang, Qiaoliang Xiang, Ye Wang, and Jialie Shen. CompositeMap: a novel music similarity measure for personalized multimodal music search. In *Proc of the 17th ACM international conference on Multimedia*, MM '09, pages 973–974, 2009.