Towards Intelligent Assembly of Media Assets for Automated Character Animation

Michael Hausenblas, Roland Mörzinger, Philip Hofmair, Werner Haas

Institute of Information Systems & Information Management
JOANNEUM RESEARCH Forschungsges. mbH
Steyrergasse 17, 8010 Graz, Austria
firstname.lastname@joanneum.at

Abstract. Creating character animations manually is an expensive and laborious task. In this work we analyse the current, manual workflow of creating character animations. We derive requirements for an automated process, and propose to utilise linked open datasets for context management, along with ontologies to assemble and reuse character animations. First experiences with the prototypical implementation of the context manager are reported.

1 Motivation

Character animations are a vital and omnipresent element of today's multimedia environments. From the infamous Clippy[1] over the blockbuster Chicken Run[2] to the many ways one is able to spend in Second Life[3]—character animations certainly have found their way into many places of our everyday's life. However, creating character animations manually is an expensive and laborious task. Further, even when the issue of creating appealing character animations was resolved, reusing them would still not be possible straight-forward.

To resolve these issues, we propose to use Semantic Web technologies along with public available datasets to handle that kind of *common knowledge* needed in the process of assembling animated characters. The goal of the presented work is to enable a faster and highly automated process of creating animated characters in the realm of the **S**emantic **A**udiovisua**L** **E**ntertainment **R**eusable **O**bjects (SALERO) project[4], a project focusing on reusable and intelligent multimedia objects for games, movies and broadcast content. In SALERO *intelligent content* for media production, consisting of multimedia objects with context-aware behaviours for self-adaptive use and delivery across different platforms is developed. The intelligent content should enable the creation and reuse of complex, compelling media by artists who need to know little of the technical aspects.

[1] The animated paper-clip, an entry point for the MS Office help system
[2] http://www.aardman.com/chickenrun/
[3] http://secondlife.com/
[4] http://www.salero.info/

This paper is structured as follows. In section 2 we discuss related approaches. Section 3 gives an analysis of the character animation workflow serving as a basis for the requirements listed in section 4. We discuss the proposed system architecture in section 5, and report on first experiences with a prototypical implementation of an important part (the Context Manager). Finally, we conclude the work and sketch future steps in section 6.

2 Related Work

Issues around character animations have been discussed in various fields. The notion of context has been studied in different areas of artificial intelligence. For example McCarthy [McC87] proposed to formalise context as a possible solution to the problem of generality. Dey et.al. discussed requirements for dealing with context in smart environments [DAS99]. However, contextual issues stemming either from *vague human user input* or *low-level semantic online data* need to be addressed properly. To the best of our knowledge there exists no other approach using linked open data sets to manage contextual issues of character animation.

Unlike a number of efforts in the realm of multimedia ontology engineering [TPMC04,ATSH07] we are not interested in using fine-grained media asset descriptions on a formal basis, per se. We rather focus on multimedia metadata [HBB$^+$07] in conjunction with domain ontologies, and issues regarding media assets and Semantic Web technologies [BH07].

The AIM@SHAPE project[5] aims to foster the development of new methodologies for modelling and processing the knowledge related to digital shapes. This knowledge is concerned with the spatial extent of the object, the structure, attributes (colours, textures), and semantics, as well as interaction with time.

3 Reusing Character Animations Today

As animated characters possess only a limited inventory of actions, the risk of boring the consumer is imminent. It hence seems obvious to suggest that the more animations a character can do, the more flexible it becomes. Flexibility in turn is one driving force for reusability.

Current practice teaches that there is no common agreement where and how media assets have to be stored. For example, there may exist facial animations, gestures, textures, audio files, etc.—all stored in different locations for every single character. The relations between characters, positions, etc. are not readily kept for further use in a uniform way. When having a large number of animations, a comprehensive categorisation becomes important. The initial work in the SALERO project was done in organising collections of concrete animations for supporting reusability as follows.

[5] http://www.aimatshape.net/

Morphology The morphology of the movement describes the proportions of the character for which the movement has been designed. More exactly, the proportion of the body against the head of the character is used;

Type Describes if the animation is related to a biped or not. In the biped case, we further define either a men, a women or a non-anthropomorphic character;

Weight Describing the movement's fluency. Three values have been assigned: stylised, normal and fat;

Acting Refers to the acting of the character, such as fast, nervous, calm, etc.;

Category (movement, gestures, facial) This element is associated with the type of bones the animation is affecting. Not all the animations must define the movement of the full skeleton, but just parts of it;

Description Within this element a rather textual description of what the animation is doing can be provided. It is not used for querying and retrieving the animation but to have a comprehensive description at hand.

4 Requirements

Based on the previous section, we have identified the following patterns in today's process in creating character animations manually. We note that for an effective and efficient management of the media assets certain practical constraints must be acknowledged:

- A rather large *set of media assets* has to be managed, including administrative metadata from low-level to higher semantics;
- *Character* properties and the actual instances are predefined and fixed. This means, the number of combinations is finite and can be assessed in advance;
- Each character consists of certain *parts*. Their orchestration must be defined in an unambiguous way;
- The actual *spatial assembly is fixed*—as a consequence, each part of a character has a predefined position;
- Real-time query parameters basically determine the final assembly of the animated character. These query parameters can (i) be triggered by a *human user* (e.g. a keyword) or (ii) stem from *online* data (for example geo-location, temperature, etc.);
- Typically the *query parameters can not directly be used for the assembly.* Depending on the source of the query parameter, it either may be too vague (in case of a human user) or semantically not rich enough (online data).

We note that certain of the above listed requirements contradict each other. For example, when one is after well-annotated media assets for being able to quickly identify a certain part, this comes normally with the cost of manual work.

5 Semantic Support for Character Animation Assembly

To enable the automated assembly and reuse of character animations, we propose a system called *Intelligent Media Object Generator* (IMOgen). In Fig. 1 the system architecture of IMOgen is depicted. The system accepts manual input (in the form of keywords) or alternatively online data, such as geo-coordinates from a GPS-enabled device. The IMOgen system comprises the following components:

1. The *Context Manager* (CM), being responsible for normalising the input queries. The output is a set of normalised query terms (URI references);
2. The *Matcher*, matching the URI references from the CM with media asset descriptions. The CM iteratively feeds the Matcher with URI references, until an appropriate media asset is found;
3. The *Assembler*, a simple metadata alignment engine that takes media asset descriptions and domain ontologies and produces the metadata necessary to assemble an animated character.

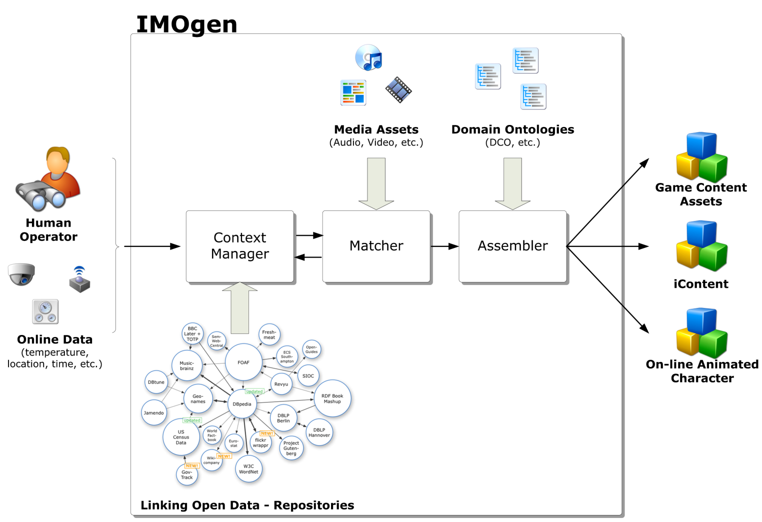

Fig. 1. The IMOgen System Architecture.

In the following sections we explain the main components of the IMOgen architecture. The current focus of the development has been set on the Context Manager, as this is the critical piece in the assembly process w.r.t. both functional and practical issues.

5.1 Context Management

Many options are available to manage contextual issues. However, we take a rather unconventional approach and utilise a huge set of already manually annotated data available on the Web. In the IMOgen *Context Manager* (CM) we utilise a mechanism entitled *linked open data*, a collaborative effort carried out in the W3C *Semantic Web Education and Outreach Interest Group* (SWEO)[6]. The declared goal of SWEO's Linking Open Data (LOD) community project[7] is to extend the Web with Semantic Web compatible versions of common datasets, such as Wikipedia, WordNet, Geonames, etc. This goal is achieved by publishing these datasets in RDF [MM04] on the Web, and adding links between data items from different data sources. The current state of the LOD is depicted in Fig. 2; at the time of writing the datasets collectively comprises over two billion RDF triples interlinked by approximately 680.000 links.

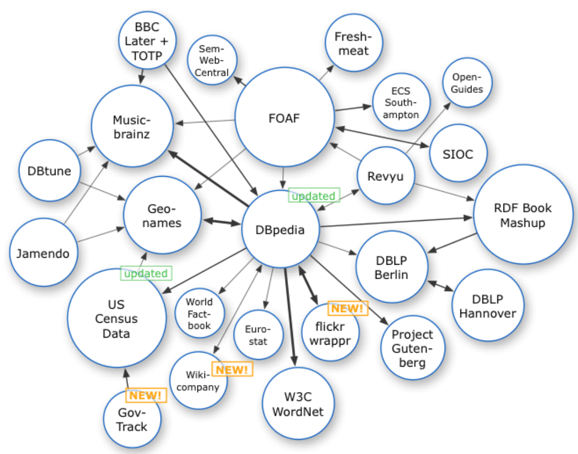

Fig. 2. The current state of the LOD dataset, as of 11/2007.

The LOD dataset acts as the common knowledge base in the CM to resolve contextual issues. Take for example a GPS-equipped sensor that is connected to the IMOgen system as an online data source. It may input the following geo-coordinates to the *IMOgen Context Manager*:

```
latitude='47.06401' longitude='15.45323'
```

(which incidentally happens to be the location of our Institute's head quarter in Graz, Austria). Using a lookup service[8] of the geoname partition of the LOD[9], the following query finds nearby Wikipedia entries:

```
http://ws.geonames.org/findNearbyPlaceName?lat=47.06401&lng=15.45323&type=rdf
```

[6] http://esw.w3.org/topic/SweoIG/
[7] http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
[8] http://www.geonames.org/export/
[9] http://www.geonames.org/ontology/

When executing the above query, the Web Service returns RDF triples that further can be used in the IMOgen Matcher to find appropriate parts of an animated character. In case the IMOgen Matcher signals to the CM that no matching media asset was available, the process is repeated. In the latter case, either an other LOD data set—for example the DBPedia[10] data set—may be selected as the seed for the query, or links (e.g., `rdfs:seeAlso`, etc.) from the initial query are followed.

The IMOgen Context Manager Web Interface (CMWI)[11] can be used to manually run such queries. The IMOgen CMWI is a debug interface used to test and evaluate contextual queries.

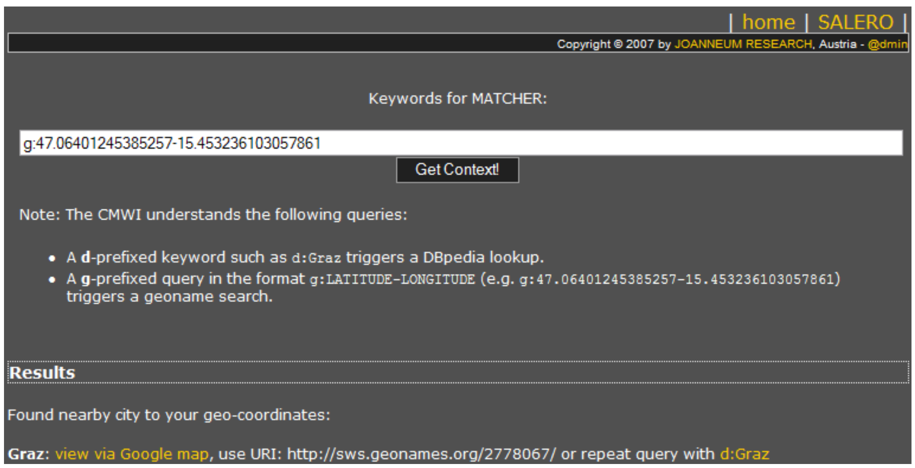

Fig. 3. The IMOgen Context Manager Web Interface: Geoname Query.

In Fig. 3 the initial query (with data from, e.g. an GPS-equipped sensor) for the city at the coordinates `47.06401-15.45323`, and the result is depicted. In a consecutive query the so found city (Graz, in this case) can be used to further query another part of the LOD data set. Continuing the example, for example from the DBPedia LOD partition it can be gathered that Graz is a city in Austria, and further that Austria is located in Central Europe. Hence, a media asset annotated with '*suitable for Central Europe*' may yield a match.

5.2 Management of Multimedia Metadata

It is a common practice that different types of media assets are spread over different locations. Amongst other storage facilities, a media asset may be a database

[10] http://dbpedia.org/

[11] Available at http://sw.joanneum.at/IMOgen/index.php

entry, it may exist as a file or a combination of files in a file system. Further, it can be composed of descriptions within a MPEG-7 [ISO01] document. As we aim at a non-disruptive solution, we do not propose to use a uniform storage management. Rather we suggest a system that facilitates enhanced retrieval of existing media assets based on semantic media annotations.

For the purpose of managing the media assets, a semantic annotation tool is needed. It may be used to described the media assets semantically or to populate a *domain ontology* by semi-automatically annotating media assets. This tool benefits from terms proposed by context and content analysis functionality. For instance, when annotating a background image content-based analysis suggests that the image contains blue sky based on colour statistics, and the context analysis infers from the task context that the user is working on an animation for a production taking place in winter. The annotation tool then provides relevant concepts from the ontology for annotating.

For the search tool this management enables full text search on concepts and attributes. An interface allows for querying the ontology given as an input some terms and queries, and getting in return suggested concepts (query expansion). Another important feature is the auto-completion suggesting concepts from the ontology to search for.

5.3 Character Domain Ontology

To enable the assembly of the parts an animated character consists, we propose to utilise domain ontologies that capture the knowledge how parts are assembled. In our case a so called *character domain ontology* (CDO) is utilised attaching a specific instantiation of a character with its constituting parts.

Based on the work of two partners in the SALERO project[12] we have formalised the character animation domain—with concepts such as `Character` and `Part`, and roles such as `hasPart`, etc.—in OWL-DL [DS04]. The CDO acts as one of the domain ontologies used by the IMOgen Assembler to finally generate the metadata necessary to assemble the animated character, as demonstrated below. Lines 1–2 in listing 1.1 constitutes a subset of the CDO T-Box. Line 3

```
1 Character, Part ; atomic concepts
2 hasPart ; atomic role
3 ted : Character ; ted is a character
4 funnyHeadT : Part ; the funny head is a part
5 ∃ ted.hasPart{funnyHeadT} ; ted has a funny head
```

Listing 1.1. An example CDO usage.

and 4 are the A-Box of the CDO and may act as an input to the IMOgen Assembler. Finally, line 5 is either the result of the IMOgen Matcher's output or the result of the IMOgen Assembler's internal reasoning.

[12] Universitat Pompeu Fabra and Activa Multimedia

6 Conclusions & Acknowledgements

In this work we have analysed the current process of manually creating character animations in the realm of the SALERO project. In order to automate the process, we have derived requirements and proposed a system architecture, the intelligent media object generator (IMOgen). IMOgen targets at generating the metadata needed for assembling animated characters, based on human user input and online data stemming from sensors. For the most novel part of the IMOgen system architecture—the Context Manager—we have implemented a prototypical implementation and demonstrated its use. Currently, we gather experiences with the Context Manager and refine the system's design. In a next step we plan to implement the whole system, extend the DCO, and integrate it into editing tools.

Acknowledgements. This work was partially supported by the European Commission under contract FP6-027122; the Semantic AudiovisuaL Entertainment Reusable Objects (SALERO) project. The authors wish to thank Werner Bailer for his useful comments and the anonymous reviewers for helping to sharpen the focus of the paper. Finally we would like to acknowledge the work done by the SALERO project partners.

References

[ATSH07] R. Arndt, R. Troncy, S. Staab, and L. Hardman. Adding Formal Semantics to MPEG-7: Designing a Well-Founded Multimedia Ontology for the Web. Technical report, The University of Koblenz-Landau, 2007.

[BH07] T. Bürger and M. Hausenblas. Why Real-World Multimedia Assets Fail to Enter the Semantic Web. In *International Workshop on Semantic Authoring, Annotation and Knowledge Markup (SAAKM07)*, Whistler, Canada, 2007.

[DAS99] A. Dey, G. Abowd, and D. Salber. A Context-based Infrastructure for Smart Environments. In *Proc.of the 1st International Workshop on Managing Interactions in Smart Environments*, pages 114–128, 1999.

[DS04] M. Dean and G. Schreiber. OWL Web Ontology Language Reference. W3C Recommendation, W3C Web Ontology Working Group, 2004.

[HBB$^+$07] M. Hausenblas, S. Boll, T. Bürger, O. Celma, C. Halaschek-Wiener, E. Mannens, and R. Troncy. Multimedia Vocabularies on the Semantic Web. W3C Incubator Group Report, W3C Multimedia Semantics Incubator Group, 2007.

[ISO01] ISO. Multimedia Content Description Interface (MPEG-7). Standard No. ISO/IEC 15938, 2001.

[McC87] J. McCarthy. Generality in artificial intelligence. *Commun. ACM*, 30(12):1030–1035, 1987.

[MM04] F. Manola and E. Miller. RDF Primer. W3C Recommendation, W3C RDF Core Working Group, 2004.

[TPMC04] C. Tsinaraki, P. Polydoros, N. Moumoutzis, and S. Christodoulakis. Integration of OWL ontologies in MPEG-7 and TV-Anytime compliant Semantic Indexing. In *16^{th} International Conference on Advanced Information Systems Engineering*, Riga, Latvia, 2004.