

High-Level Description Tools for Humanoids

Víctor Fernández-Carbajales¹, José María Martínez¹, and Francisco Morán²

Grupo de Tratamiento de Imágenes

¹ Escuela Politécnica Superior, Universidad Autónoma de Madrid
Cantoblanco, Madrid, Spain

{Victor.Fernandez, JoseM.Martinez}@uam.es
<http://www-gti.ii.uam.es>

² E.T.S. Ing. Telecomunicación, Universidad Politécnica de Madrid
Madrid, Spain

fmb@gti.ssr.upm.es
<http://www.gti.ssr.upm.es>

Abstract. This paper presents a proposal for description tools, following the MPEG-7 standard, for the high-level description of humanoids. Given the almost complete lack of high-level description tools for 3D graphics content in the current MPEG-7 specification, we propose descriptions aimed at describing virtual humanoids, both for indexing and query support (no extraction tools are presented here), and also for the generation of personalized humanoids using high-level descriptions via a simple GUI instead of complex authoring tools. This later application, which is the focus of the work presented here, is related with the Authoring 744 initiative that targets the creation of content from descriptions that are authored in a user friendly (natural) way. This work is under development within the EU-funded research project OLGA, where the description tools should provide the means for the creation and modification of humanoids inside an on-line 3D gaming environment, but our description tools are generic enough to be used in the future in many different applications: robot portrait, indexing/searching, etc.

1 Introduction

Low-level descriptions of audiovisual content are usually easy to extract automatically and are being used in some applications, but users like and need high-level descriptions to search and browse through large repositories in an efficient way. Therefore, high-level descriptor tools have a special interest nowadays, and besides the work trying to bridge the semantic gap in automatic indexing (automatically inferring high-level descriptions from also automatically obtained low-level ones), there is also a need for additional specification or improvement of high-level descriptions, mainly for specialized characterisation of content.

High-level descriptions are not only useful for searching and browsing, but also for other applications, like the generation of synthetic content, or its reduced storage or transmission. The Authoring 744 framework [1] already permits to generate synthetic content from descriptions in a user-friendly way and avoiding the need of complex

authoring tools. Besides, the EU-funded research project OLGA (a unified scalable framework for On-Line GAMing) is exploring the advantages of storing and transmitting descriptions, and synthesising the 3D content at the client terminal.

Focusing in the description of humanoids, it will be possible to perform searches for one-legged and tall humanoids, to send a description of a person shown in a surveillance camera instead of the whole video, and to create a robot portrait (identikit picture) via high-level descriptions of the associated humanoid. In any of these cases, the advantages (less representation size, faster queries, user-friendly interfacing, etc.) are clear, but unfortunately so are the disadvantages: analysis resources for generating the descriptions, fidelity of the reconstructions, etc.

The motivation of this work is thus to have tools for high-level description of humanoids, allowing further development of the abovementioned applications and associated base technologies. MPEG-7, the standard for multimedia content description [2] lacks many specialized and high-level description tools, but also provides extensibility mechanisms allowing the creation of new MPEG-7 compliant description tools. The availability of such specialized and high-level description tools may allow an easier adoption of MPEG-7 for final applications and therefore by industry.

As there are no description tools for humanoids in MPEG-7, we propose in this paper a set of them. Our descriptions, which are mostly high-level ones, could be used in the future for the abovementioned applications; currently, we are focusing specifically on building an user-friendly GUI authoring tool allowing to create or modify an already existing virtual humanoid for personalized avatars for 3D games [3].

The rest of the paper is structured as follows: Section 2 provides a brief overview of the state of art regarding 3D humanoid authoring, representation, description and rendering. Section 3 presents the proposed description tools in its different subsections. Section 4 provides a very-short presentation of the current work in authoring tools of these descriptors, before raising some short conclusions in Section 5.

2 State of the Art

For the creation of 3D content, there exist powerful authoring tools, that can be classified as “design-driven”. Design-driven authoring tools are proprietary applications, like 3ds Max [4] or Maya [5], that use proprietary representation formats, which in some cases have become *de facto* standards. These programs are very complex for the average users, with long curves of understanding and learning; the users need a lot of knowledge in computer graphics theory and additional training in the use of each particular authoring tool.

Besides the proprietary representation formats, there are different standards for the representation of 3D content, which are mainly focused on representation accuracy within the degree of compression required for efficient storage and delivery. Focusing on 3D humanoids representation there are two main standards: H-Anim and MPEG-4. Both use a standard skeleton, that was first standardised by the H-Anim group and later adopted by MPEG in 2004.

H-Anim [6] is targeted to the representation of an abstract model of human figures. This international standard describes a way of representing humanoids allowing, e.g.,

to animate using motion capture data and animation tools from one vendor the humanoids created with modelling tools from another vendor.

MPEG-4 [7] has greatly extended the 3D graphics assets of VRML97 [8] and, regarding humanoids, two of its specifications are relevant: FBA and BBA. FBA (Face and Body Animation) tools are to be found in the two first versions of MPEG-4's Part 2, Visual, and define control points for animating the humanoid's face and body, providing knobs on the face for expressing emotions by moving all its important parts (eyes, eyebrows, lips, ears, etc.), and on the body joints (articulations). The animation of humanoids, or any other virtual character, with the BBA (Bone-Based Animation) tools, specified in MPEG-4's Part 16, AFX (Animation Framework eXtension) [9], is based on the creation of a skeleton (i.e., a group of hierarchically organised bones) together with their associated muscles and skin. Each bone has its own axis system and is then integrated in the common skeleton axis. With this integration, also kinematic constraints are imposed. All bones can suffer rotation, scaling and translation, allowing the deformation and animation of the "base" humanoid as wished, the corresponding modifications propagating to the skin associated to the bones. The BBA specification is more powerful than FBA because it offers more usage functionalities for completely generic virtual characters, and because the control points of FBA are not as useful as the bones of BBA for the generation and animation of humanoids. Besides, BBA improves the quality of the specified graphics, as the system of muscles coupled to the bones generates more realistic movements.

Another related research project was EMM [10], whose system aimed at the description of complete scenes using primitive objects and actions that were pre-stored in the system databases and knowledge engine. A script allowed to combine the primitive objects and actions in order to create animated movies with them.

Besides the representation of content, description of content is also of importance, not only for indexing and searching of 3D content, but also for understanding, and reasoning, lightweight delivery, etc., and, as already mentioned, for authoring [1]. The AnyShape project [11] proposes a set of ontologies for the description of humanoids, as well as for their animation the terms they propose allow to be identify parts of the humanoid (hands, head, arm, etc.) but don't describe the characteristics of each (thickness, width, length, high, etc.). Our current works is related to theirs but they propose a framework aiming at the modification and animation of humanoids based on shape parameters adjustment, whereas our proposal, besides being based in the skeleton and associated muscles and skin, includes also the creation of new content and not only modifications as in the AnyShape project.

As started above there aren't description tools for detailed description of humanoids that may allow to characterize each part, not only identifying it. Therefore we propose a set of them within the framework of MPEG (see section 3).

Regarding the rendering of 3D content, there exist different standards (*de jure* and *de facto*): OpenGL [12], DirectX [13], etc. These standards specify APIs (Application Programming Interfaces) for writing applications to create, manage and visualize 3D content via "methods" of that API, requiring to have the appropriate version of the corresponding library or drivers installed in the rendering terminal. Another possibility are the rendering programs based on some representation format, that interpret and render all the primitives objects of the selected representation format standard, e.g., an MPEG-4 player or a VRML engine.

3 Structure of the Proposed Description Tools

In this section, we will explain the structure of the proposed description tools for humanoids. Although there are several standards (see previous section) for humanoid representation, we have not found in the literature any other work related to the description of humanoids. Each subsection explains a branch of the description tools tree.

3.1 Humanoid DS

The **Humanoid DS** (see Fig. 1) is the root of the proposed description tools for humanoids. The Humanoid DS is composed by the GeneralCharacteristics DS, the CorporalSubdivisions DS, the Handicaps DS, the NormalExtras DS and the FantasticExtras DS.



Fig. 1. Humanoid DS

3.2 GeneralCharacteristic DS

The **GeneralCharacteristic DS** (see Fig. 2) provides the description tools for the basic main characteristics of humanoids: sex, age, skin colour, race (including the degree of mixture of racial features), and corporal measures (e.g., height, weight, corporal mass).

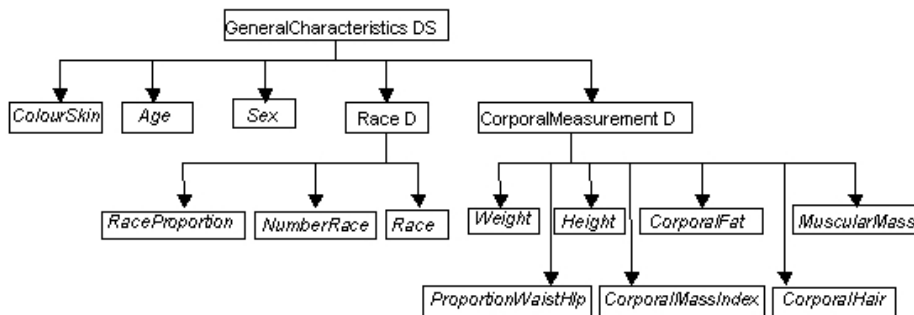


Fig. 2. GeneralCharacteristic DS

3.3 CorporalSubdivision DS

The **CorporalSubdivision DS** (see Fig. 3) provides the description tools for the subdivisions of the different parts of the body, allowing the user to describe the

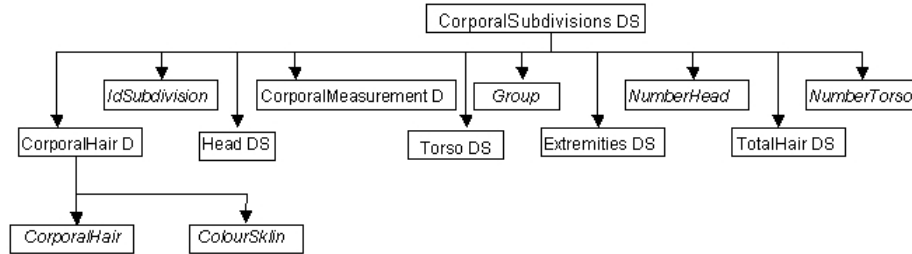


Fig. 3. CorporalSubdivision DS

principal divisions of the body at different levels of detail. The first subdivision creates description tools for the head, the torso, the extremities and the corporal hair (colour and quantity).

3.4 Head DS

The **Head DS** (see **Fig. 4**) is composed by description tools for a generic description of the head (contour, high, width, etc.) and additional description tools for a more detailed description: Face DS, Skull DS and Neck DS. The **Face DS** includes description tools for a detailed description of ears, eyes, mouth, nose and jaw, as well as the dimension of face and other generic features. The **Skull DS** and **Neck DS** provide generic description tools (height, width, etc.) together with other, targeted to more detailed descriptions of the skull and neck respectively.

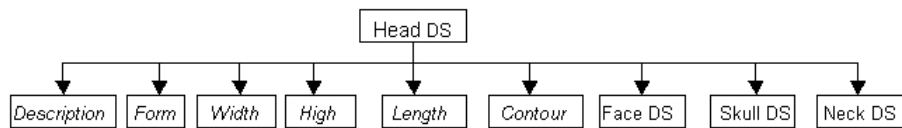


Fig. 4. Head DS

3.5 Torso DS

The **Torso DS** (see **Fig. 5**) includes generic features (high, width, length, etc.) as well as specialized description tools for a detailed description of the chest, abdomen, pelvis and back of the humanoid.

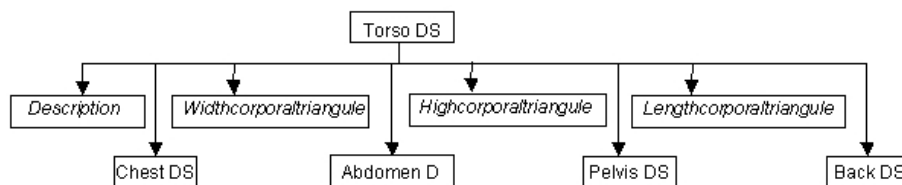


Fig. 5. Torso DS

3.6 Extremities DS

The **Extremities DS** (see **Fig. 6**) specifies lots of description tools for the detailed description of the extremities (limbs), from the complete arm to a phalanx detail level, including the possibility of accepting fantastic extremities as wings or tails. In every extremity, we can describe, at the same time, his joints (shoulder, elbow, knee, etc.) and not-articulated parts (hand, arm, thigh, etc.).

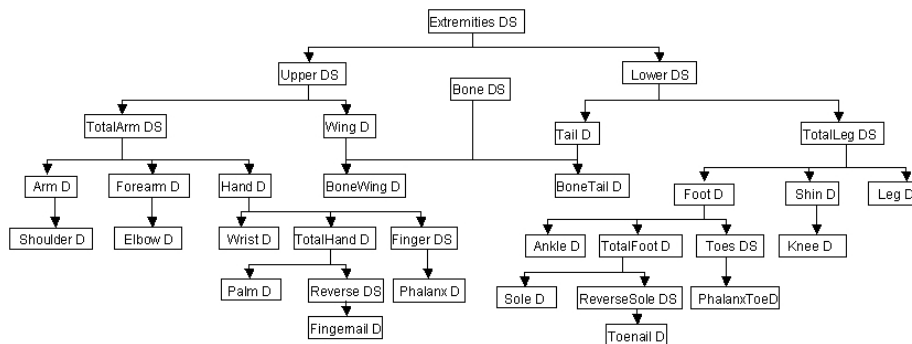


Fig. 6. Extremities DS

3.7 HeadHair DS

The **HeadHair DS** includes description tools for the different hairy parts of the head: hair **HairD** (indicating the colour, the hairdo, etc.), **Moustache D**, **Beard D**, **Sideburns D** and **Eyebrow D**. All these description tools allow to specify the colour of the hair, its shape, its density, etc.

3.8 Handicap DS

The **Handicap DS** specifies description tools for handicaps or disadvantages of the humanoid. These description tools cover anatomical handicaps (e.g., amputation of an arm, leg or the nose, the presence of a hump, etc.) as well as functional ones (e.g., paralysis, visual handicaps, etc.).

3.9 Extras

In order to fully describe the humanoid (but not including clothes), there are two description tools for description add-ons or extras, split in two categories: the **NormalExtras D** specifies description tools for tattoos, scars, (beauty) spots, glasses, etc.; the **FantasticExtras DS** includes description tools aimed at fantastic humanoids, such as **Wing DS**, **Tail D**, **Horn D** and **Claw D**.

4 Current Work

We are currently fine-tuning an authoring tool for the creation and modifications of humanoids that is based in the “writing” of descriptions in an easy to use GUI

(see **Fig. 7**). The tool creates MPEG-7 descriptions following the description tools presented above. The descriptions are transformed afterwards in MPEG-4 files that are used for visualization of the humanoids under creation or modification.

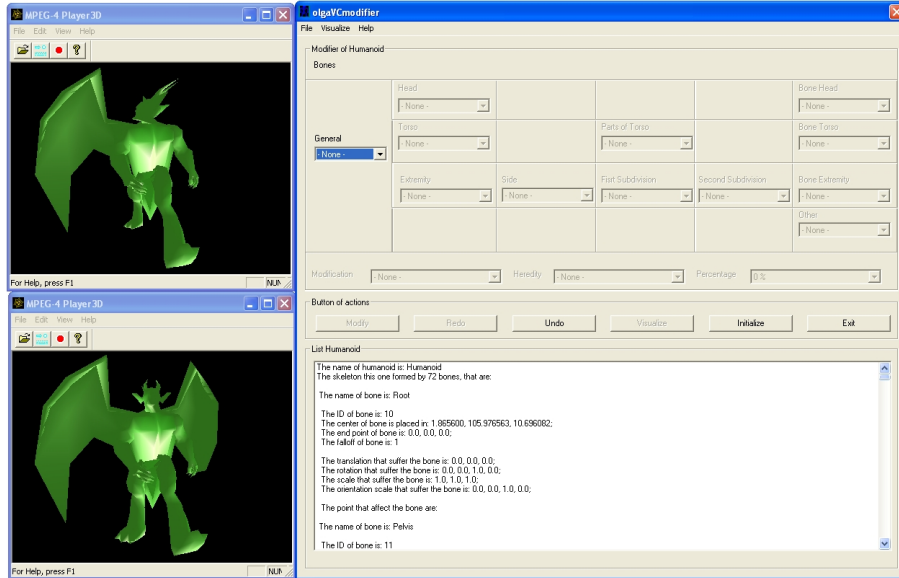


Fig. 7. Screen shot of our authoring tool

5 Conclusions

The use of high-level descriptions of any kind of content allows a better searching, browsing, delivery, storage, etc., of the described content. Besides, descriptions can also be used for authoring. Current standardized description tools are mainly focused on low- and mid-level description tools, being mainly automatically extracted. The adoption of high-level descriptions require not only new analysis and reasoning technologies for bridging the semantic gap for automatic indexing, but also new tools specifying such high-level descriptions.

This paper has presented a set of tools allowing the high-level description of humanoids, paving the ground for different applications like searching for people in archives, robot portraits, etc. Currently we are applying these description tools in an easy GUI authoring tool for the creation and modification of 3D humanoids via descriptions. The descriptions are used for the creation of 3D avatars in the MPEG-4 AFX representation format, either at the server or at the terminal (in this case reducing the transmission requirements, but increasing the computing resources at the terminal). Extraction tools for the proposed description schemes are currently out of the scope of our work.

Acknowledgments

This work has been partially supported by the 6th Framework Programme of the European Commission, within its research project FP6-IST-1-507926, OLGA (a unified scalable framework for On-Line GAMing), and by the *Ministerio de Ciencia y Tecnología* of the Spanish Government, within its research project TIN2004-07860, MEDUSA. The authors wish to thank Marius Preda from INT (Institut National des Télécommunications) for his support regarding AFX.

References

1. Martínez, J.M., Morán, F.: *Authoring 744: Writing Descriptions to Create Content*. IEEE Multimedia October/November 2003, 94-98.
2. Manjunath, B.S., Salembier, P., Sikora, T. (eds.): *Introduction to MPEG-7*. John Wiley and Sons, 2002.
3. Fernández-Carbajales, V., Martínez, J.M., Morán, F.: *Description-Driven Generation of 3D Humanoids within Authoring 744*. EWIMT 2005, 227-232.
4. Murdock, K.: *3D Studio Max 7 Bible*. Ed. Anaya Multimedia-Anaya Interactiva, 2005.
5. Fyvie, E. : *Introducing Maya 6: 3D for beginners*. Sybex Inc., D. Brodnitz, 2004.
6. Sims, E. (ed.): *H-Anim 200x Humanoid Animation*. (FCD of ISO/IEC 19774), 2004.
7. Pereira, F., Ebrahimi, T. (eds.): *The MPEG-4 book*. Prentice-Hall, 2002.
8. ISO/IEC JTC1/SC24: *ISO/IEC 14772 1, Virtual Reality Modelling Language v2.0*, 1997.
9. ISO/IEC JTC1/SC29/WG11 (a.k.a. MPEG): *ISO/IEC 14496-16, MPEG-4 Part 16, AFX (Animation Framework eXtension)*, 2003.
10. Shen, J., Aoki, T., Yasuda, H., Miyazaki, S.: *E-Movie Creation by Rule-Based Reasoning from the Director's Viewpoint – E-Movie: Computer Animation & Real Images*, EWIMT 2004, 252-259.
11. Garcia-Rojas, A., Thalmann, D., Vexo, F., Mocozet, L., Magnenat-Thalmann, N., Mortara, M., Spagnuolo, M., Gutiérrez, M.: *An Ontology of Virtual Humans: Incorporating Semantics Human Shapes*. EWIMT 2005, 7-14.
12. Leech, J., Brown, P. (ed.): *The OpenGL Graphics Systems: A Specification Version 2.0*, 2004.
13. Gaines, B.: *Managed DirectX 9 Graphics and Game Programming*, Kick Start. Smas, 2003.