Visual Comput (2007) 23: 207–218
DOI 10.1007/s00371-006-0093-4

ORIGINAL ARTICLE

Mario Gutiérrez A.
Alejandra García-Rojas
Daniel Thalmann
Frederic Vexo
Laurent Moccozet
Nadia Magnenat-Thalmann
Michela Mortara
Michela Spagnuolo

An ontology of virtual humans

Incorporating semantics into human shapes

Published online: 23 January 2007

M.A. Gutiérrez A. (✉)
ITESM Campus Toluca
m.gutierrez@itesm.mx

A. García-Rojas, · D. Thalmann · F. Vexo
VRlab – EPFL
{alejandra.garciarojas, daniel.thalmann, frederic.vexo}@epfl.ch

L. Moccozet · N. Magnenat-Thalmann
MIRALab – University of Geneva
{moccozet, thalmann}@miralab.unige.ch

M. Mortara · M. Spagnuolo
IMATI – CNR
{michela.mortara, spagnuolo}@ge.imati.cnr.it

Abstract Most of the efforts concerning graphical representations of humans (Virtual Humans) have been focused on synthesizing geometry for static or animated shapes. The next step is to consider a human body not only as a 3D shape, but as an active semantic entity with features, functionalities, interaction skills, etc. We are currently working on an ontology-based approach to make Virtual Humans more active and understandable both for humans and machines. The ontology for Virtual Humans we are defining will provide the "semantic layer" required to reconstruct, stock, retrieve, reuse and share content and knowledge related to Virtual Humans.

Keywords Human shape reconstruction and synthesis · Semantic annotations · Ontologies

1 Introduction

Virtual Humans, as 3D graphical representations of human beings have a large variety of applications. Within inhabited virtual environments, Virtual Humans (VHs) are a key technology that can provide virtual presenters, virtual guides, virtual actors, and be used to show how humans behave in various situations [20].

Creating Virtual Humans is a complex and time consuming task which involves several computer science areas: artificial intelligence, computer graphics, geometric modeling, multi-modal interfaces, etc. In this article we give an overview of the state-of-the-art on analysis and synthesis of human shapes. We present recent advances and underlaying difficulties on the creation of VHs.

With the successful development of massively multiplayer online role playing games or MMORPG and their persistent virtual worlds, it becomes now common to exchange or buy 3D objects such as virtual weapons and tools or even 3D Virtual Human characters used as avatars. For instance, Sony is proposing Station Exchange[1] (see Fig. 1), a website to exchange virtual characters. On the other hand, the recent advent of efficient and accurate acquisition systems and reconstruction platforms has potentially increased the availability and dissemination of complex digital shapes such as Virtual Humans. As a consequence, the offer and demand for Virtual Humans will probably rapidly expand. However, the infrastructures that would be required for sharing large databases of Virtual Human models and populating Virtual Environments

[1] http://stationexchange.station.sony.com/

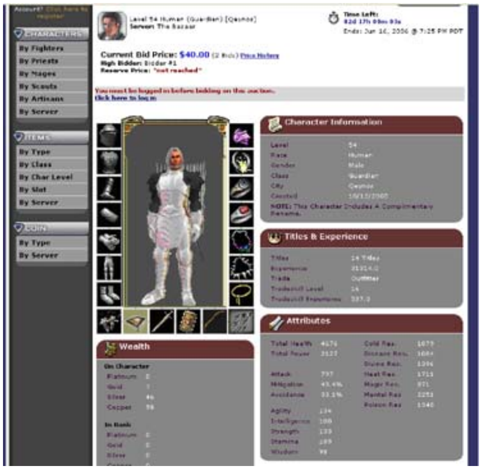

Fig. 1. Snapshot from a Virtual Character proposed for auction on the Sony Station Exchange website

are still immature; moreover a common understanding to share these models and their information does not exist.

Our contribution is a semantics-based method for organizing the various types of data that constitute a Virtual Human, in order to foster a common understanding and sharing of such complex 3D entities. The knowledge related to the synthesis, animation and functionalities of VHs is formally specified in the form of an ontology. Ontology development is a continuous process and as such the results we present here are works in progress. Nevertheless, the current ontology provides a good starting point towards the creation of a more versatile and reusable representation of Virtual Humans.

An ontology representation of a Virtual Human must be closely linked to the associated graphical one. It is particularly required to be able to go from the graphical representation to the ontology-semantic one: with the analysis of the 3D graphical representation in order to query the 3D models for semantic information. It is also required to be able to go from the ontology description to the graphical representation: with the integration of the semantic descriptors in the modeling and animation process, which means that we need to construct the graphical representation of a Virtual Human from the semantic descriptors.

The next section of this article is dedicated to the synthesis of human shapes based on real data. These are particularly suited to derive semantic data based on human shape synthesis methods. The third section surveys the requirements for modeling active virtual humans, which means turning Virtual Humans into entities able to inhabit Virtual Environments and interact with it and with other virtual inmates. The fourth section demonstrates that the morphological analysis of human body shapes combined with anthropometric knowledge allows to extract accurate low-level and intermediate semantic features. In section five we propose an ontology for Virtual Humans and we give an application scenario. Last section presents the conclusion and future perspectives of this work.

2 Synthesis of human shapes

Recent advances in scanning technology made it possible to easily model Virtual Humans (VH) from real humans [7], which is particularly important for producing realistic looking VH. However, acquisition devices do not provide "ready-to-use" results and post-processing is required in order to obtain accurate shapes.

Two strategies are possible to manage the acquisition post-process: either relay only on geometric information, such as curvature or normal, to correct the data, or consider that each object is an instance of a family of objects that share similarities and therefore use semantic-based templates to correct and complete the acquired data. This first strategy is general but does not catch the specificities of the reconstructed object. It can miss important features of the object or complete it in an inconsistent manner. The second strategy is specific to one family of objects but ensures to have as a result a shape that includes all the features of the object's family.

The integration of pre-existing knowledge in automatic reconstruction process should greatly improve the accuracy of the resulting instance to capture all the features of a family of objects. Any standard human shape is expected to have specifically located parts, e.g., five fingers on one hand. It is clear that this requires a robust (re)construction of the human shape incorporating knowledge of morphology and anthropometry.

The shape of the human body is a typical example family of an articulated physical object: it does not have only one shape but many, corresponding to all the possible postures that the underlying articulated skeleton can take. When acquiring scan data, we only obtain a single static snapshot of the body shape. As such and for many range applications, this static snapshot is not sufficient as it does not capture all the possible degrees of flexibility of the human shape.

To mimic the flexible and dynamic behavior of the human shape, the traditional approach uses skeleton-driven deformations, a classical method for the basic skin deformation that is among the most widely used techniques in 3D character animation. It binds a 3D shape to an articulated control skeleton. Binding information is then used to deform the body shape according to control skeleton motion (Fig. 2).

Therefore constructing an appropriate and accurate model of the human shape should not only reconstruct a 3D shape that matches the body of the scanned volunteer but should also reconstruct all the related information that makes it possible to reproduce the human shape in any of

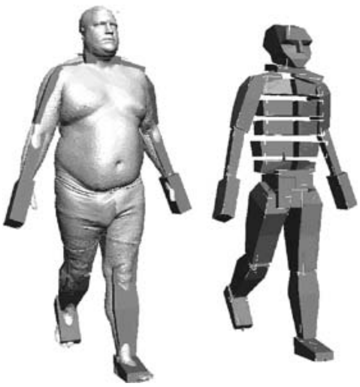

Fig. 2. Human shape and its associated control articulated skeleton

its possible postures. This aspect includes a high degree of integration of the semantics of the human body within the construction process.

From this short introduction, we can understand that human shape (re)construction involves:

– Reconstructing more than only the static shape information (and particularly the control animation structure: control skeleton and skin binding)
– Taking into account more than only the available input geometrical information in the reconstruction process

Algorithms making use of domain knowledge are more accurate because they prevent the surface from erroneously being corrected in under-sampled areas. Examples are template-model-based fitting strategies, where scan data is repaired with the geometry from a template surface. These approaches are gaining more and more interest. Examples include the methods proposed by Kähler et al. [14] for faces, using the template-to-data correspondence found using easily identifiable landmarks and Allen et al. [1, 2] who reconstruct bodies using correspondence markers attached to the scanned people.

In [16], Moccozet et al. proposed a full reconstruction pipeline that produces a close approximation of the scanned data of a human body. It is based on fitting a human template model defined in [18] by Seo et al. (which includes both the skin surface and the animation control information) to the scanned data. It captures both the variability of individual human bodies and their common structure. The correspondence between the template and the scanned data is set semi-automatically with a tool called Tailor, which is based on a multi-scale morphological analysis method [17]. Some of the landmarks can be automatically extracted and a visualization of the morphological regions helps the user localize the other required landmarks.

Beside the reconstruction of human shapes from acquired data, template-based methods can be extended to synthesize new body shapes. The most common approach consists in acquiring a set of complete models that are expected to cover the space of human body shapes. The human body space is then expressed with a few parameters extracted from statistical analysis of the template models database. One difficulty of this kind of approach lies in the correlation of the statistical parameters with morphological descriptors.

In [3], Ben Azouz et al. proposed an alternate approach for extracting the variations of the human shape from a 3D anthropometric database using a volumetric representation of human shapes. More recently, Wang [22] described a feature-based parametrization of the human body for constructing mannequins from body scan data. However, none of these methods integrates the semantic information related to the animation structure required to control and animate the human shape.

In [18], Seo et al. developed a framework for collecting and managing range scan data that automatically estimates the intrinsic articulated structure of the body from user-tagged landmarks. By framing the captured and structurally annotated data so that statistic implicit is exploited for synthesizing new body shapes, the technique supports time-saving generation of animatable body models with high realism. A user can generate a new model or modify an existing one simply by inputting a number of sizing parameters. High level semantic descriptors such as body measurements are derived from low-level semantic data such as landmarks.

The semantic description of the human body shape requires defining the common features between human shapes. Alternatively, this should also bring another open question: what makes each body shape different from each other or in other terms, how far is it possible to characterize the individualization of human shape morphologies?

A typical semantic-based shape synthesis scenario is the identikit one, where an approximation of a human shape would be directly built from high-level descriptors. The input would be a description combining quantitative and qualitative information such as: this person is big, approximately 1.8 m, young, fat, with blue eyes, dark hair.

No two persons with the same weight and same size have the same body shape. Therefore, an important issue is to determine the size of the shape descriptors space required to capture the distinctiveness of human shapes or to rely on captured data to incorporate individual features within the reconstruction scheme.

3 Active human body representation

Once a virtual reconstructed object will be immersed inside a virtual environment, it will have to be able to act as its real counterpart. Therefore it cannot be limited to a soup of triangles without higher level information. This is particularly true for Virtual Humans. Whenever they are included within a virtual environment, they are expected to move and interact with this environment. Obviously, the 3D body shape and even the control animation structure do not include the required level of information for modeling an "active" human shape.

3.1 First level of interaction: accessories

Accessories are objects such as clothes, jewels, hats, glasses that are attached to the human shape. Their motion and animation depends on the motions and animations of the human shape itself. Attaching accessories to a human shape involves locating where they should be placed on the body shape and extracting measurements information in order to fit the accessories to the body shape. At first glance, correctly placing an accessory intuitively relies on a morphological segmentation of the human shape. For example a watch will be placed at the border between the hand and the forearm; a shirt will cover the trunk and the arms.

In [5], Cordier et al. automatically resized the garments worn on a 3D body model as the body changes its dimension. They first pre-compute the attachment of the pre-defined cloth mesh to the body surface by defining attachment data of the garment mesh to the skin surface. Morphological segmentation of the human shape can also be used to optimize the simulation process when flexible accessories such as clothes are simulated according to the human shape animation such as in [6]. A more general approach would consist in defining feature lines on the human body shape. These feature lines would allow placing the accessories at the correct place and determining measurements in order to adjust them to the human morphology.

3.2 Second level of interaction: manipulation of objects

The necessity to model interactions between an object and a virtual human appears in most applications of computer animation and simulation. Such applications encompass several domains, as for example: virtual autonomous agents living and working in virtual environments, human factors analysis, training, education, virtual prototyping, and simulation-based design.

Commonly, simulation systems perform agent-object interactions for specific tasks. Such an approach is simple and direct, but most of the time, the core of the system needs to be updated whenever one needs to consider another class of objects.

Smart objects are an interesting way to model general agent-object interactions based on objects containing interaction information of various kinds: intrinsic object properties, information on how to interact with it, object behaviors, and also expected agent behaviors. The smart object approach, introduced by Kallmann and Thalmann [15] extends the idea of having a database of interaction information. For each object modeled, we include the functionality of its moving parts and detailed commands describing each desired interaction, by means of a dedicated script language. A feature modeling approach [19] is used to include all desired information in objects.

4 Features extraction and morphological decomposition

In the previous sections we have depicted a scenario where VHs are created starting by an acquisition process; then, the acquired models must undergo post-process reconstruction phases; last, available accessories must be defined, and virtual objects that can be manipulated (and how they can be manipulated) must be specified.

Having an ontology for Virtual Humans would be interesting only if we can automate first the process of acquiring the data related to Virtual Humans and second, to be able to exploit data and knowledge repositories organized according to such ontology. In this sense, we have been working on different algorithms designed to automatically or semi-automatically acquire data from existing Virtual Humans (3D shapes) and populate the ontology. This mainly results in decomposing the shape into meaningful segmented parts or in locating anthropometric landmarks over the body model.

The most common features involved in human shape synthesis are landmarks. Landmarks and segments provide low-level semantic descriptors from which it is possible to derive higher-level ones. Landmarks are points of correspondence on each object of the same kind that match between and within populations [4]. The widely adopted landmarks structure for the human shape is the one proposed in the H-ANIM standard [13] description, as shown in Fig. 3 (top). Two common approaches are currently applied to label a human shape with landmarks: either by sticking markers to the volunteer prior to the scan session or by manually labeling the scan data shape. Few attempts have been proposed to automatically or semi-automatically extract some of the landmarks or features such as in [23] where the authors apply fuzzy logic to extract five features (armhole, crotch, neck, chest, and belly button) from body scan data.

Low-level semantic descriptors must be:

- Morphologically meaningful, as they have to be used to describe the body shape in a natural way: the semantic descriptors must correspond to usual natural body descriptors if we expect anyone to be able to describe a human shape with them.
- Computationally extractable: it is obviously required that the semantic body descriptors can be systematically extracted from human shape data.
- Computationally embedded into the reconstruction process: it is important that the reconstruction process can be parameterized with the semantic descriptors.

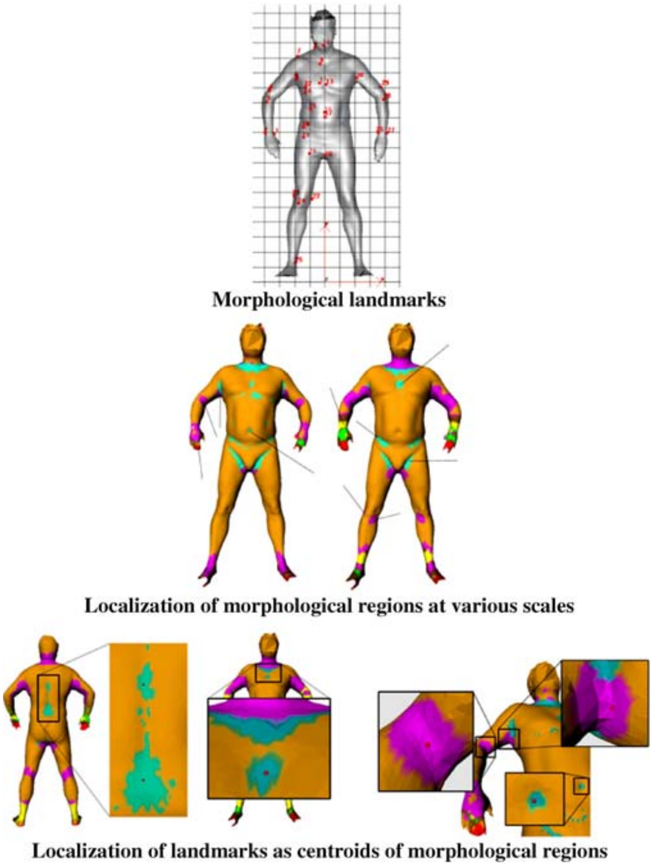

Fig. 3. Morphological landmarks

For example, it would be useless to define a size parameter for the breast if it is not possible to integrate the breast size in the reconstruction pipeline.

The computational methods involved in the extraction of features such as landmarks or shape segmentation must comply with the following constraints:

– Landmarks extraction and morphological segmentation results must be anthropometrically consistent. Extracted features and segmentation must be associated to anthropometric features and segments.
– Landmarks extraction and morphological segmentation results must be consistent and almost invariant from one data set to another.

In Fig. 3 (middle and bottom), we show some examples of landmarks extraction based on a multi-scale morphological analysis of the human shape. These features are extracted with a tool called Tailor, which is based on a multi-scale morphological analysis method [17].

Tailor detects different kinds of features (sharp protrusions, wells, dips, and branching parts) at different scales on a given shape (Fig. 4). Vertices are labeled with differ-

Fig. 4. Detected areas on a scanned body at all scales; in the last picture, the centroids corresponding to these regions are located by the circles and they globally correspond to the anatomical joints

ent colors according to the feature they are associated to. We performed tests on a set of bodies with different corpulence. For each labeled area of vertices, its centroid is defined as the mean of triangle's barycenters. We compute them across different scales and filter these features points according to their position invariance. We did experiments on the average centroid displacement across scales and noticed it stays approximately located at a same position considering a set of bodies, regardless of fatness, size and morphology, which confirms our hypothesis: the anatomic joints of a human on a scanned mesh correspond to features that can be detected by the Tailor algorithm. Some of the detected points do not correspond to real joints: we can filter them by taking into account only the blends and concave features. The Plumber algorithm, which builds on the vertex classification provided by Tailor, identifies tubular features together with their axes (Fig. 5).

Our strategy relies on the observation that shape variations correspond to areas corresponding to skeletal joints. Intuitively, we expect that the areas where we detect small

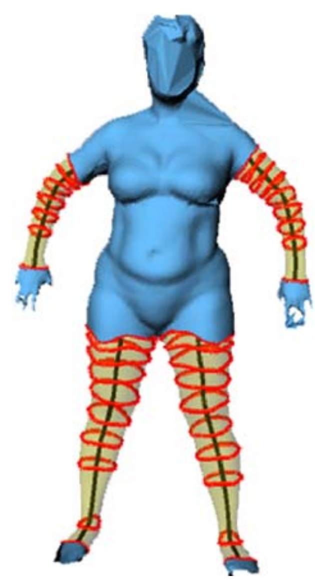

Fig. 5. Tubular features with their axes

variations lay in the same location with bigger variations. Since the skin is thinner where we have natural articulations, it produces small concavities and variations at small scales. There are also less fat tissues and no muscles around joints, so it induces variations at bigger scales among body parts. Our proposed approach is based on two main steps:

1. An initial segmentation is applied on 3D scanned data with Plumber to extract tubular (limbs) and non-tubular parts of the body (trunk).
2. A multi-scale shape analysis is further applied with Tailor onto tubular and non-tubular segments to extract potential location of joints.

These results are combined with a priori knowledge on human anatomy in order to select the best candidates. The first segmentation provides arms and legs, which are tubular. Each part contains only two joints: in the case of arms, we have elbows and wrists. In the case of legs, we have ankles and knees. Tailor centroids are then filtered in order to extract for each limb the two best candidates corresponding to the joint location. For the trunk, Tailor allows for extracting the spine curve. This curve is used to estimate the height of the shoulders and hip joints. Extracted landmarks are further projected onto the axes of the Plumber tubes in order to estimate the joints locations.

In order to correct possible incoherencies due to bad scan data, we use the symmetry property: since the human skeleton is symmetric, the lengths of symmetric bones segments should be the same. We also use anatomic proportions, according to statistical data: proportions between long bones of the skeleton are very stable with a negligible variation. We present in Fig. 8 three different scanned bodies with different morphotypes.

The experiments on various morphotypes demonstrate that the resulting method is able to handle noisy 3D input data for variable 3D body shapes and postures and to extract with a reasonable accuracy the main 16 joints and the 5 tip joints corresponding to a simple H-ANIM hierarchy (Fig. 6).

Having finished with this overview of the algorithms and techniques used for features extraction and morphological decomposition, we proceed to describe how this large body of data and associated knowledge can be organized and structured in a formal way through an ontology.

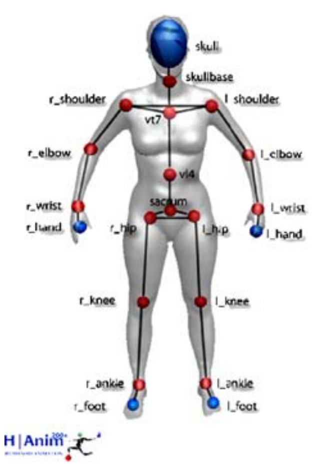

Fig. 6. Detectable set of joints on a scanned body, and their H-ANIM correspondences: in red detectable joints, in blue tip joints

5 An ontology for Virtual Humans

According to Gruber [8], an ontology is a formal specification of a shared conceptualization. Virtual Humans are complex entities composed by well defined features, and functionalities. Concepts and techniques related to the creation and exploitation of VHs such as those described in previous sections are shared by the research community. Our effort targets at unifying such concepts and representing them in a formal way. A formal representation refers to the fact that VH representations and their associated semantics shall be both human and machine readable: this is achieved by means of an XML-based representation.

Ontological principles are well recognized as effective design rules for information systems [10, 21]. This has

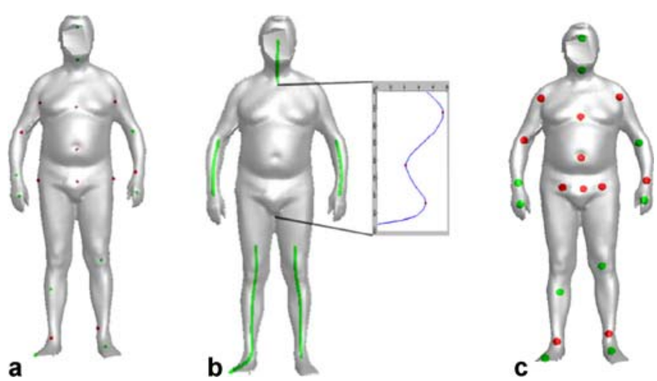

Fig. 7. **a** Extracted landmarks corresponding to the 2 steps shape analysis **b** Tubular axes and spine curve **c** Estimated joints location from **a** and **b**

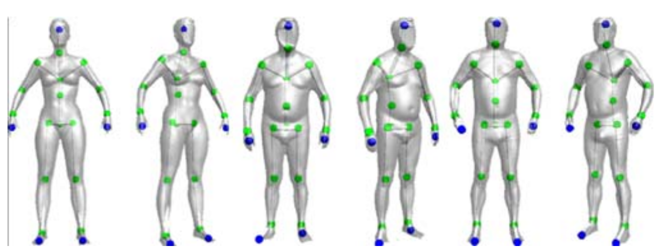

Fig. 8. Skeleton extraction for different morphotypes

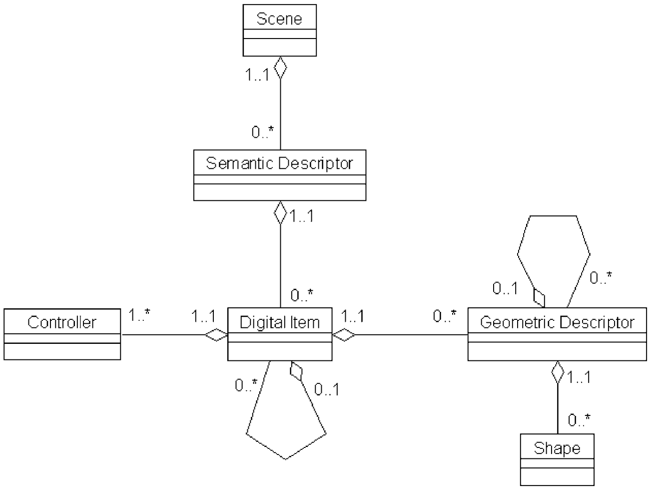

Fig. 9. Semantic representation of an interactive virtual environment

led to the notion of "ontology-driven information systems" which structural and temporal dimensions [10]. Our objective is to support the creation and exploitation of VHs by such a system. The structural dimension concerns a database containing the information (semantic descriptors). The temporal dimension is related to the interface (visual programming) that gives access to the information at run-time.

Associating semantic information to the components of a virtual environment has proved to be useful in terms of component reuse, content adaptation, etc. In [12], Gutiérrez et al. defined an object representation based on the semantics and functionality of interactive digital items – virtual objects – within a Virtual Environment (VE), see Fig. 9. Every object participating in a VE application is a dynamic entity with multiple visual representations and functionalities. This allows for dynamically scaling and adapting the object's geometry and functions to different scenarios. In [11], the semantic model presented in [12] was complemented with an ontology of objects that allowed for expressing the relationships between interaction devices and virtual entities in a VE.

The present work builds upon the acquired experience and focuses on a single type of virtual entity: Virtual Humans.

5.1 Developing the ontology

The development of an ontology usually starts by defining its domain and scope. That is, answer several basic questions known as *competency questions*. Competency questions (CQs) are one of the best ways to determine the scope of the ontology. CQs consist on a list of questions that a knowledge base based on the ontology should be able to answer [9].

Competency questions

The proposed ontology should be able to answer the following categories of competency questions:

Model history

- Is this model obtained by editing another model?
- What features have been changed on model X?
- What tools where involved in the synthesis/modification of this VH?
- Who performed the task T on the model X?

Features listing

- What is the height of the model?
- Is the model male or female?
- Is the model European?
- What are the features of this model?
- Is this model obtained artificially or does it represent a real person?
- Which VH have a landmark description?
- Which are the available structural descriptors for a particular VH?
- Which aspects of the shape are described by the structural descriptor related to a particular VH?
- Which are the standing (seating, walking) VH?
- How is the body model represented? (a mesh/ a point set/...)
- Is the VH complete? (does it have a skeleton/ a hierarchy of body parts/ a set of landmarks attached to it?)

Questions whose answer is a function of low/high level features

- Most of the answers to these questions cannot be directly answered by the ontology – at least not in the current version. Answers will be provided by external algorithms which will take as input the data retrieved through the ontology.
- Which are the VH that are fat/slim/short? Is this VH a child or an adult?
- Does it have a long nose?
- Does it miss any body part?
- Does this VH match another VH (or how much do they match)?, and in particular: Are they in the same posture? Do they have the same structure? Do they have similar parts? (same arm length/same fatness/similar nose?)
- Do they have similar anthropomorphic measures (in terms of landmarks?)
- Is the model suitable for animation?
- How will this VH look after 20 years? With 20 kg more? With another nose?
- Does this model fit this cloth? (Clothes haven't been considered in the current version of the ontology. How-

ever, they could be considered as a special case of smart object or as a geometry with particular landmarks).
– What VH do I get if I put the head of VH1 on the body of VH2?

Animation sequences

– What model does this animation use?
– What are the joints affected by this animation sequence?
– Are there any animation sequences lasting more than 1 minute suitable for this VH?
– Are there any "running/football playing" animation sequences for this kind of VH?
– Can the animation sequence X be applied to the VH Y? (in the case of key-frames for skeleton-based animation this would basically depend on the possibility to match the key-frame data to the skeleton of the VH).

Animation algorithms

– What are the input and output channels of a particular behavior controller (animation algorithm)?
– What are the models suitable to be animated with this algorithm?
– Does this VH have a vision sensor attached?
– Can this VH react to sound events in its virtual environment?

Interaction with objects

– What capabilities does an object provide?
– What are the actions the human can execute on the object?
– What are the characteristics of an object? (structure, physical properties, etc.)
– How can the object be grasped?

5.2 Ontology components

We have defined a first version of the VH ontology based on the competency questions listed before. The ontology for Virtual Humans aims at organizing the knowledge and data of three main research topics and applications involving graphical representations of humans:

– Human body modeling and analysis: morphological analysis, measuring similarity, model editing and/or reconstruction.
– Animation of Virtual Humans: autonomous or pre-set animation of VH.
– Interaction of Virtual Humans with virtual objects: virtual smart objects that contain the semantic information indicating how interactions between Virtual Humans and objects are to be carried out.

Once we have a data structure and data concepts we can specify how these interrelate with each other and build a formalization of the knowledge: an ontology for Virtual Humans. A diagram of the ontology is presented in Fig. 12. The following are the concepts we have defined to express in a formal way the information and knowledge associated to Virtual Humans:

Geometry. The geometry is the physical visual representation of the Virtual Human and is composed of two parts, primary: body shape, and secondary: accessories, garments, etc. Therefore, we can create a general description for any kind of geometry shape and inherit its properties to specialized subclasses, specifically for Virtual Humans or objects. This common class contains the geometric properties: number of vertex, edges, scale, material and texture. The Virtual Human has a geometrical representation where we can know the kind of geometry it has (e.g., surface mesh).

Animation. Virtual Human's animation should distinguish between facial and body animation because the way of animating each part is different. Body animation can be KeyFramed or Motion Captured. There are standardized formats to specify animation: MPEG-4, VRML, etc. In order for an animation to take place, the character should have a skeletal structure. This structure is defined in the Structural Descriptor. We have based this structure on the H-ANIM specification as it is the adopted standard in most of the animation formats, but any other structures can be also represented.

Morphology. Morphological Descriptor contains information like: age, weight, height, gender.

Behavior. Individual Descriptor and behavior controller are for describing the behavior. The Individual Descriptor contains the constant definition of the behavior of the Virtual Individual like his personality or cultural identification, background. The behavior controllers are algorithms that drive the behavior of the character considering the emotional state and its individuality.

Figure 12 presents how the concepts are connected to each other in the ontology. The concepts that are in a circle are subclasses of the concept Resource, in the bottom of the diagram. As a consequence, those concepts have inherited the properties of the Resource: author, version, file, etc. This means that the Resources are the product that the user wants to get when searching for a component. Some of the resources can be found in separated files or more than one in the same file. For the case of behavior controller the user can get an algorithm.

This is a simplified version of the ontology with the most important definitions we have explained in this paper; however, much more information has been considered, for example the structure of H-ANIM (joints and segments), the animation formats of MPEG-4 (FAP, BAP), etc.

Finally we can see in the diagram that there is a lot of information around one concept, the Virtual Human. There are physical concepts and abstract concepts that confer meaning to the Virtual Human. In the next section we will present an application scenario where a user wants to build its own virtual character from scratch by searching the desired components in the ontology.

5.3 Ontology application scenarios

The following are some of the main application scenarios where the ontology for Virtual Humans can play an essential role:

Virtual characters data repository. A search engine for retrieving VHs and smart objects with particular features/functionalities related to animation. As animation is tedious and resource consuming process, re-usability is highly desirable. This could be implemented inside a repository where users can search for animations with specific conditions, e.g., a mature woman running, and using the morphological descriptors and attributes of the animation. The result will be different types of files corresponding just to the animation, or to the animation with the model.

Modeling data repository. A place where a modeler/animator could find VH shapes (whether full or partial bodies) and use them to model new VH, improve or reconstruct existing ones. The ontology could help with the management of an on-line shop to build custom 3D mannequin and select and fit clothes and accessories, with real-time pre-visualization. It would help to decide on the most suitable clothes and accessories for an specific VH by matching the corresponding landmarks.

Fig. 10. An application scenario: looking for a male character in the VH repository

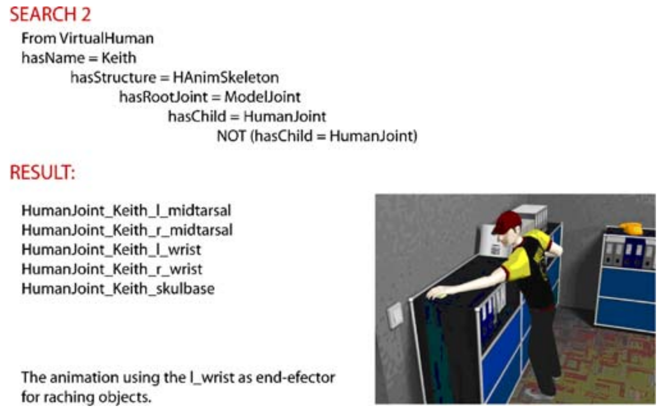

Fig. 11. An application scenario: looking for IK end-effectors of a VH

Shape recognition/extraction/analysis. A knowledge base able to answer competency questions linked to low-level features of the VH shape (landmarks, topological graphs, and so on). Main users would include researchers working on algorithms for recognizing features on a shape representing a virtual/real human. Data would be used on ergonomics studies, computer vision algorithms, etc. Let us consider a case where a user wants to create a game with a male virtual character. This character should be an animatable character, which means it should have a skeleton. The user could perform the search shown in Fig. 10.

From the search for male characters with an skeletal structure, the user could chose Keith (see Fig. 10).

To make the animation of Keith, one may want to use Inverse Kinematics. We need to know the end effectors of Keith. The end-effectors are Human Joints that are localized at the end of the skeletal structure, and as a consequence do not have children. The user could perform in the ontology following query shown in Fig. 11. This query consists on searching in Keith's skeletal structure for all the joints that are children of Root Joint and do not have more children. The results are the joints that will be used in the inverse kinematics as end effectors.

6 Conclusions

The current version of the ontology for Virtual Humans is work in progress. As stated before, there are still missing components which are required to fulfill all the needs of a complex and multi-disciplinary task such as the creation and use of Virtual Humans. However, we believe this is an important step towards a formal representation of Virtual Humans.

In this paper we described some of the main issues to be solved in order to effectively model VHs. We have presented our advances on an ontology-based approach. This is a promising alternative for modeling and managing the

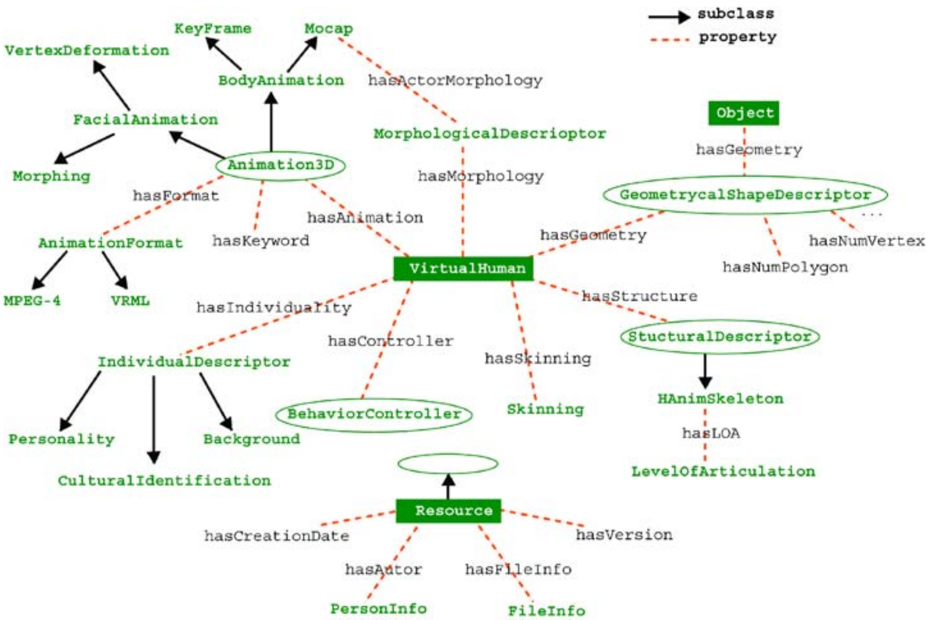

Fig. 12. Main components of an ontology for Virtual Humans

knowledge related to Virtual Humans. Taking advantage of an ontology for VHs depends on a two-way process: labeling graphical representations with semantic information and being able to extract semantic information from graphical representations. This will be achieved through shape analysis and segmentation combined with anthropometric knowledge and large sets of acquired data. We are currently focusing our efforts on advancing the state of this research.

Acknowledgement This work has been supported by the Swiss Federal Office for Education and Science and the European Union in the framework of the project IST – AIM@SHAPE Network of Excellence (http://www.aimatshape.net/).

References

1. Allen, B., Curless, B., Popović, Z.: Exploring the space of human body shapes: Data-driven synthesis under anthropometric control. In: Proceedings of Conference on Digital Human Modeling for Design and Engineering. SEA International (2004)
2. Allen, B., Curless, B., Popović, Z.: The space of human body shapes: reconstruction and parameterization from range scans. ACM Trans. Graph. **22**(3), 587–594 (2003)
3. Azouz, Z.B., Rioux, M., Shu, C., Lepage, R.: Analysis of human shape variation using volumetric techniques. In: Proceedings of International Conference on Computer Animation and Social Agents, pp. 197–206. IEEE, Washington, DC (2004)
4. BookStein, F.L.: Morphmetric Tools for Landmark Data: Geometry and Biology. Cambridge Press, Cambridge (1997)
5. Cordier, F., Seo, H., Magnenat-Thalmann, N.: Made-to-measure technologies for online clothing store. IEEE Comput. Graph. Appl. (special issue on Web Graphics) **23**(1), 38–48 (2003)
6. Cordier, F., Magnenat-Thalmann, N.: A data-driven approach for real-time clothes simulation. In: Proceedings of the 12th Pacific Conference on Computer Graphics and Applications, pp. 257–266 (2004)
7. Daanen, H.A.M., van de Water, G.J.: Whole body scanners. Displays **19**(3), 111–120 (1998)
8. Gruber, T.: The role of a common ontology in achieving sharable, reusable knowledge bases. In: Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning, pp. 601–602 (1991)
9. Gruninger, M., Fox, M.: Methodology for the Design and Evaluation of Ontologies. In: Proceedings of the Workshop on Basich Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal (1995)
10. Guarino, N.: Formal ontology and information systems. In: Proceedings of FOIS ’98, (Trento, Italy, June, 1998). IOS, Amsterdam pp. 3–15 (1998)
11. Gutiérrez, M., Thalmann, D., Vexo, F.: Semantic virtual environments with adaptive multimodal interfaces. In: Proceedings of the 11th International Conference on Multimedia Modelling (MMM2005), pp. 277–283 (2005)
12. Gutiérrez, M., Vexo, F., Thalmann, D.: Semantics-based representation of virtual environments. Int. J. Comput. Appl. Technol. “Models and methods for representing and processing shape semantics” **23**(2–4), 229–238 (2005)
13. H-ANIM Working Group: ISO/IEC FCD 19774:200x, Humanoid Animation, Annex B, Feature points for the human body. http://www.h-anim.org/
14. Kähler, K., Haber, J., Yamauchi, H., Seidel, H.P.: Head shop: generating animated head models with anatomical structure. In: SCA ’02: Proceedings of the 2002 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 55–63. ACM, Boston (2002)
15. Kallmann, M., Thalmann, D.: Direct 3d interaction with smart objects. In: VRST ’99: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, pp. 124–130. ACM, Boston (1999)
16. Moccozet, L., Dellas, F., Magnenat-Thalmann, N., Biasotti, S., Mortara, M., Falcidieno, B., Min, P., Veltkamp, R.: Animatable human body model reconstruction from 3d scan data using templates. In: Proceedings of

Workshop on Modelling and Motion Capture Techniques for Virtual Environments, CAPTECH'2004, pp. 73–79 (2004)
17. Mortara, M., Patane, G., Spagnuolo, M., Falcidieno, B., Rossignac, J.: Blowing bubbles for multi-scale analysis and decomposition of triangle meshes. Algorithmica **38**(1), 227–248 (2003)
18. Seo, H., Magnenat-Thalmann, N.: An automatic modeling of human bodies from sizing parameters. In: SI3D '03: Proceedings of the 2003 symposium on Interactive 3D Graphics, pp. 19–26. ACM, Boston (2003)
19. Shah, J., Mantyla, M.: Parametric and Feature-Based CAD/CAM. Wiley, New York (1995)
20. Thalmann, D.: The virtual human as a multimodal interface. In: AVI '00: Proceedings of the Working Conference on Advanced Visual Interfaces, pp. 14–20. ACM, Boston (2000)
21. Uschold, M., Gruninger, M.: Ontologies: Principles, methods and applications. Knowledge Eng. Rev. **11**(2), 93–155 (1996)
22. Wang, C.C.L.: Parameterization and parametric design of mannequins. Comput. Aided Des. **37**(1), 83–98 (2005)
23. Wang, C.C.L., Chang, T.K.K., Yuen, M.M.F.: From laser-scanned data to feature human model: a system based on fuzzy logic concept. Comput. Aided Des. **35**(3), 241–253 (2003)








MARIO GUTIÉRREZ graduated from Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM) in Computer Science. He received his Ph.D. degree from École Polytechnique Fédérale de Lausanne in 2005. Since August 2005 he has been an assistant professor of the Computer Science Faculty of ITESM in Mexico. His research interests include: semantics for Virtual Environments, computer animation and interaction paradigms. He has been involved in several European research projects.

ALEJANDRA GARCÍA-ROJAS graduated from Instituto Tecnológico y de Estudios Superiores de Monterrey in Computer Science. She received her M.T.D.I degree from the University Complutense of Madrid in 2004. Since January of 2005 she has been a Ph.D. student at the Virtual Reality Lab (VRlab) of the École Polytechnique Fédérale de Lausanne.

PROFESSOR DR. DANIEL THALMANN is Professor and Director of The Virtual Reality Lab (VRlab) at EPFL, Switzerland. He is a pioneer in research on Virtual Humans. His current research interests include Real-time Virtual Humans in Virtual Reality, Networked Virtual Environments, Artificial Life, and Multi-media. He is co-editor-in-chief of the Journal of Visualization and Computer Animation, and member of the editorial board of the Visual Computer and three other journals. Daniel Thalmann is a member of Eurographics since 1980. He was the initiator of the Eurographics working group on Animation and Simulation which he co-chaired for more than 10 years. He is now a member of the Stearing Committee of the Symposium on Computer Animation, a joint Eurographics-Siggraph initiative. He received his Ph.D. in Computer Science in 1977 from the University of Geneva and an Honorary Doctorate (Honoris Causa) from University Paul-Sabatier in Toulouse, France, in 2003.

DR. FREDERIC VEXO is currently senior researcher and project leader at the Virtual Reality Laboratory of the École Polytechnique Fédérale de Lausanne, Switzerland. His research interests are multi-disciplinary and include Inhabited Virtual Worlds, Haptic Interfaces, Tele-operated Systems, Multi-modal Adaptive Interface for virtual worlds, Semantic Virtual Environments and new services for graphic mobile devices. He is author of several papers in journals and international conferences in the fields of Human Computer Interaction, Robotics and Computer Graphics. He is a member of several conference program committees (AAMAS, SMI, SVE and IWVR) and an expert for different institutions and companies. He received his Ph.D. in Computer Science in 2000 from University of Reims and Bachelor's and Master's degrees from University of METZ. He also has contributed to various European projects.

LAURENT MOCCOZET is a senior researcher at MIRALab, University of Geneva. His main research field is geometric modeling and geometric deformations for articulated characters such as virtual humans. He is also investigating semantics-based representation for high-level models of virtual humans. He has been actively involved in several European and Swiss Projects focused on human shape modeling for virtual environments.

NADIA MAGNENAT-THALMANN has pioneered research in virtual humans over the last 20 years. She studied at University of Geneva and obtained several degrees in Psychology, Biology, Chemistry, and a Ph.D. in Quantum Mechanics from University of Geneva in 1977. From 1977 to 1989, she was a professor at University of Montreal, Canada. In 1989, she founded MIRALab, an interdisciplinary creative research laboratory at University of Geneva. She has published more than 200 papers, and is editor-in-chief of Visual Computer and the Journal of Visualization and Computer Animation journals. She has received several awards, and recently she has been nominated at the Swiss Academy of Technical Sciences.

MICHELA MORTARA graduated in Computer Science from the University of Genoa in 1999. She also received a Ph.D. in Robotics in 2004. Since 2000 she has been a researcher at the Institute for Applied Mathematics and Information Technologies of CNR in Genoa, where she has been involved in some National and International Projects. She is mainly involved in shape modeling, in particular in shape similarity, skeletonization, shape analysis and decomposition.

MICHELA SPAGNUOLO is senior researcher at IMATI–CNR Genova. Her research interests are related to computational topology for shape understanding, classification and retrieval, and shape-based approaches to modeling and processing digital shapes. She is currently involved in the AIM@SHAPE project as leader of the research program related to "Analysis and Structuring" of digital shapes. She is a member of the SMI (Shape Modeling International) steering committee and member of the editorial board of IJSM.