

Nadia Magnenat-Thalmann  
Daniel Thalmann

## Virtual humans: thirty years of research, what next?

Published online: 24 November 2005  
© Springer-Verlag 2005

N. Magnenat-Thalmann  
MIRALab, University of Geneva,  
Switzerland

D. Thalmann  
Vrlab, EPFL, Switzerland

**Abstract** In this paper, we present research results and future challenges in creating realistic and believable Virtual Humans. To realize these modeling goals, real-time realistic representation is essential, but we also need interactive and perceptive Virtual Humans to populate the Virtual Worlds. Three levels of modeling should be considered to create these believable Virtual Humans: 1) realistic appearance modeling, 2) realistic, smooth and flexible motion modeling, and 3) realistic high-level behaviors modeling. At first, the issues of creating virtual humans with better skeleton and realistic deformable bodies are illustrated. To give a level of believable behavior, challenges are laid on generating on

the fly flexible motion and complex behaviours of Virtual Humans inside their environments using a realistic perception of the environment. Interactivity and group behaviours are also important parameters to create believable Virtual Humans which have challenges in creating believable relationship between real and virtual humans based on emotion and personality, and simulating realistic and believable behaviors of groups and crowds. Finally, issues in generating realistic virtual clothed and haired people are presented.

**Keywords** Virtual human · Motion modeling · Crowd modeling · Cloth and hair simulation · Believable behavior

### 1 Introduction

In the beginning of eighties, we could see short movies and demos involving high quality animation sequences involving very realistic virtual humans. We may cite Adam Powers, the Juggler produced by Triple I or “sexy robot” from Robert Abel. Twenty years later, can we say that there was an incredible progress in virtual humans? This is a very key question and we can answer it in different ways.

The rendering of the juggler and sexy robot were very nice at that time. Although some aspects including textures are better today, it is very hard to tell that the animation at current stage is any better than that concerning the motion. Both the juggler and sexy robot are realistically animated using typical motion captures, which are

still well used for current movies like Final Fantasy or even to build databases of motion for video games.

So, if rendering and motion are rather similar to twenty years ago, does this mean that the research in the area is dead? Surely not, because, there have been, in fact, considerable advances.

- 1) In appearance, recent virtual humans have real clothes and hair, which was not possible in the eighties
- 2) The juggler, for example, was nice, but the body was rigid. Today’s virtual humans have deformable bodies: it is possible to generate automatic bodies with various measurements, a very important aspect for virtual fashion, for example.
- 3) Video games, VR systems, and real-time systems require adaptive and spontaneous motion, which is hard

to accomplish with motion capture. It needs to have sophisticated motion control methods and models.

- 4) Creating interesting games and VR systems with virtual characters requires a high-level behavior of these characters.

So, the research in cloth and hair modeling, body deformation, motion control, and behavioural animation has made considerable progress in the last decade. However, we are still in the infancy age and a lot of further developments are needed.

Let us try to consider a virtual human as it should be modeled in a complete and exhaustive way. We may consider three levels of modeling.

- 1) *Appearance modeling*. This means that visually the virtual humans should look like real people, this includes face and body shapes of real humans, skin textures, nice hairstyle, and realistic clothes.
- 2) *Realistic, smooth, and flexible motion in any situation*. This is a challenge, as flexibility is obtained by creating parameter-based models. However, no model can provide smooth and realistic motion as obtained by motion capture.
- 3) *Realistic high-level behaviour*. This is the main challenge, can we expect to model the manner in which humans behave. This is what A.I. and agent technology have been trying for many years. Unfortunately, we cannot expect better results in animation.

We have also to consider that these levels are not independent and can even also be in conflict each other. For example, clothes and hair are an important aspect in the visual appearance of the virtual humans, but they are also most of the time in motion. Their animation has to be clearly considered. More fundamentally, body motion is not only the result of a skeleton motion, but the deformation of the body itself is an important part and makes the difference with a typical robot. We need to represent these body deformations using sophisticated deformation algorithms and many approaches were proposed in the very early days. We can observe today even a trend to base such algorithms on an anatomy-based approaches derived from medical research.

Let us consider a very simple example. A virtual human is supposed to enter into a kitchen, look for a glass, find one in a cupboard, grasp it and bring it out of the kitchen. How can one make such a sequence?

First approach: we just ask a real person to do the job and record it using an optical or magnetic motion capture system; this is the easiest way, very similar to the 80's juggler, and still the way most of today's computer-based movies are done. Now imagine that we decide to replace the glass by a plate. Then the sequence of grasping the object has to be changed, because we do not grasp a glass and a plate in the same way. And what if we replace the human by a child who cannot reach the cupboard? And what if the door of the kitchen is closed? And what if there is

a table in front of the cupboard? Each new condition imposes an adaptation of the animation that would probably require the entire process to be restarted.

Second approach: we use a series of motion controllers including a walking engine, a grasping engine, inverse kinematics. This should solve some problems like changing the glass into a plate, or changing the location of the object, but not the problem of the table obstacle.

Third approach: we use a task planner with an obstacle avoidance algorithm; this will allow us to solve our problem of the table in front of the cupboard.

However, will the sequence produced by the most sophisticated approach be better than the original motion capture based sequence? Unfortunately, the answer is no, and even the results will be clearly worse. This is exactly the point. We can make more intelligent virtual humans but they look poorly animated, which gives the wrong impression to people who just watch sequences that after 30 years computer animation has made no progress and it has even regressed.

What is the future of virtual humans? We may think about virtual worlds inhabited by a virtual human society, where virtual humans will co-operate, negotiate, make friends, communicate, group and ungroup, depending on their likes, moods, emotions, goals, fears, etc. However, the computational models of such interaction and corresponding groups are very hard to develop or even impossible. Behaviour should emerge as a result of a multi-agent system sharing a common environment, in our case, sharing a virtual environment. For example in a panic situation, we should not model the group behaviour as an entity, because each human reacts differently depending, for example, on its level of fear. If we model the individual entity, there will be groups of different behaviours (not programmed explicitly) as result of the interaction of simple individual behaviours. Virtual humans should also have some natural needs like hunger, tiredness, etc., which guide selection of their behaviours. In order to behave in a believable way, these agents also have to act in accordance with their surrounding environment, be able to react to its changes, to the other agents and also to the actions of real humans interacting with the virtual world.

One important point in the simulation is the believability of the individual virtual humans. They should behave as real humans and have capabilities such as perception, language understanding and generation, emotions, goal-driven behaviour, reactivity to the environment including with other virtual humans, memory, inference, appearance of thought and personalities, interpersonal interactions, social skills and possibly others.

## 2 Further great challenges in virtual humans

Real-time realistic virtual humans will be essential in the future, and we will also need interactive perceptive virtual



They assumed a simple skeleton (see Fig. 1), but the human skeleton is very different from this structure. As a result, dynamics-based motions were more appropriate for robot-like structures than for human characters. One question arose: should we continue to develop very complex algorithms for controlling the motion of primitive structures that do not correspond to the human skeleton?

The skeleton determines the general shape of the body. It is made up of bones connected by joints that allow relative motion between different parts of the skeleton. Different classes of human joints can be distinguished according to the type of motion they allow. Hinge and pivot joints allow rotations about a single transverse or longitudinal axis.

The ball-and-socket joint allows rotational movement in all directions; a bone with a rounded end (ball) fits in a cuplike cavity (socket) of another bone. Ellipsoidal and saddle joints are similar to the ball-and-socket, except that no axial rotation is possible due to the different structure of the articulating ends.

The revolute joint allows a rotation about an arbitrarily fixed axis. The natural parameterization is the rotation angle with respect to a reference configuration. For example, Aubel and Thalmann [2] model both hinge and pivot joints with the revolute type. The knee joint performs two successive rotations about two orthogonal axes. It is not equivalent to two independent revolute joints located at the same point, because the twist limits are functionally coupled with the extension. In practice, the range of axial motion becomes more and more restricted as the leg approaches full extension, as in reality.

The swing joint allows all rotational motions except axial rotation. Both saddle and ellipsoidal joints are therefore modeled by swing joints in our system. In terms of parameterization and limits, a swing joint is strictly equivalent to a ball-and-socket joint whose axial motion is restricted altogether.

### 3.2 Body deformations

Modeling and deformation of 3-D characters, and especially human bodies, during the animation process is an important but difficult problem. Researchers have devoted significant efforts to the representation and deformation of the human body's shape. Broadly, we can classify their models into two categories: the surface model and the multilayered model. The surface model [3] is conceptually simple, containing a skeleton and an outer skin layer. The envelope is composed of planar or curved patches. One problem with this model is that it requires the tedious input of the significant points or vertices that define the surface. Another problem is that it is hard to control the realistic evolution of the surface across joints. Surface singularities or anomalies can easily be produced.

The multilayered model [4] contains a skeleton layer, intermediate layers that simulate the physical behavior of

muscle, bone, fat tissue, etc., and a skin layer. Since the overall appearance of a human body is very much influenced by its internal muscle structures, the multilayered model is the most promising for realistic human animation. The key advantage of the layered methodology is that once the layered character is constructed, only the underlying skeleton needs be scripted for an animation; consistent yet expressive shape deformations are generated automatically. Henne [5] represents the skin by a mesh of bicubic surface patches, whose control points deform at joints in accordance with several constraints such as elasticity, area preservation, and implicit repulsive force fields that mimic bones and muscles. Deformations that occur away from joints are ignored by this approach.

The question is still remains as to how to model or capture the multilayered model. Yoshomito [6] showed that implicit formulations like metaballs provide efficient ways of creating beautiful virtual humans at a reduced storage cost. Thalmann et al. [7] extended this approach by combining implicit surfaces and B-spline patches. Ellipsoids and ellipsoidal metaballs represent the gross shape of bone, muscle, and fat tissue. The motion/deformation of each primitive with respect to underlying joints is specified via a graphical interface. A skin of fixed topology is extracted by casting rays from the skeleton segments in a star-shaped manner and using the intersection points as control points of B-spline patches.

More recent work aims at mimicking more closely the actual anatomy of humans or animals. Wilhelms [8] developed an interactive tool for designing and animating monkeys and cats. In her system, ellipsoids or triangular meshes represent bones and muscle. Each muscle is a generalized cylinder made up of a certain number of cross-sections that consist, in turn, of a certain number of points. The muscles show a relative incompressibility when deformed. In their work on anatomically modeling the human musculature, Scheepers et al. [9] stressed the role of underlying components (muscles, tendons, etc.) on the form. They use three volume-preserving geometric primitives for three different types of muscles. Isometric contraction is handled by introducing scaling factors and tension parameters. The skin is obtained by fitting bicubic patches to an implicit surface created from the geometric primitives. Porcher-Nedel and Thalmann [10] introduced the idea of abstracting muscles by an action line (a polyline in practice), representing the force produced by the muscle on the bones, and a surface mesh deformed by an equivalent mass-spring mesh. In order to smooth out mesh discontinuities, they employ special springs, termed angular springs, which tend to restore the initial curvature of the surface at each vertex. However, angular springs cannot deal with local inversions of the curvature. Aubel and Thalmann [11] also use an action line and a muscle mesh. The action line, represented by a polyline with any number of vertices, is moved for each posture using a predefined behavior and a simple, physically based simulation (see





**Fig. 2.** Action lines of the pectoral muscle during shoulder abduction. Right: the use of two attractive force fields (solid and wireframe ellipsoids)

Fig. 2). It is then used as a skeleton for the surface mesh, and the deformations are produced in the usual way. Seo et al. [12] propose a very fast method of deformation for MPEG-4-based applications.

## 4 Generating on-the-fly flexible motion control

### 4.1 A need for motion control methods

Human motion has always been difficult to simulate with an acceptable level of physical realism and/or believability. This is especially the case when strong update rate requirements are imposed as in interaction context. Results from neurophysiology suggest that many basic human skills could be seen as motion patterns managed at such a low level as the spinal cord while the central brain is only in charge of modulating it according to the current context [13]. Such findings are very stimulating for real-time human animation, as relatively easy methods for motion capture are now available. Numerous approaches aiming at re-using and generalizing this raw motion material at low cost have been proposed [14]. Among them, statistical and probabilistic methods aim at encapsulating the structure of the motion implicitly, with only minimal normalization of the data [15]. Additional classification stages can map the motion database onto some high-level qualitative dimension. At the exploitation stage, specifying a value along that dimension leads to the low cost production of a new motion with the desired level of that qualitative property. The major limitation of this family of approaches comes from its black box nature and the unavoidable coupling between the various high-level qualities desired. Adding new samples to the motion database may improve some aspects while perturbing other dimensions. An alternate computer animation technique is procedural modeling that strives to establish explicit analytic formulation linking the high-level qualitative dimension to the animation parameters (i.e. the skeleton joint values).

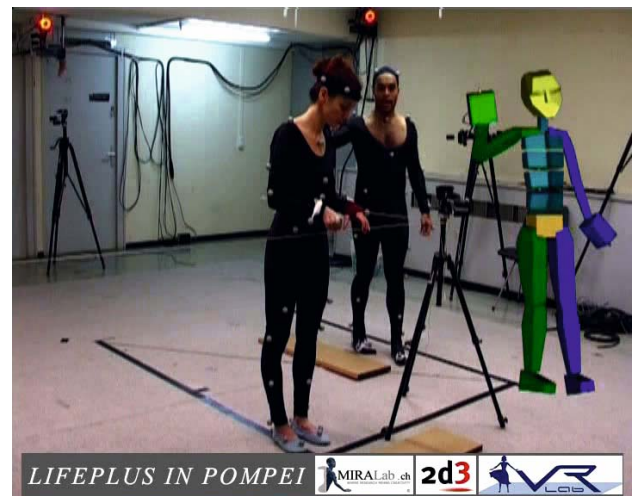
In any case, there is a need for a motion control method behind. It specifies how a virtual human is animated and may be characterized according to the type of information it privileged in animating the character. For example, in a keyframe system for an articulated body, the privileged information to be manipulated is the angle. In a forward dynamics-based system, the privileged information is a set

of forces and torques; of course, in solving the dynamic equations, joint angles are also obtained in this system, but we consider these as derived information. In fact, any motion control method will eventually have to deal with geometric information (typically joint angles), but only geometric motion control methods explicitly privilege this information at the level of animation control. Typically, motion is defined in terms of coordinates, angles and other shape characteristics. A physical motion control method uses physical characteristics and laws as a basis for calculating motion. Privileged information for these motion control methods include physical characteristics such as mass, moments of inertia, and stiffness. The physical laws involved are mainly those of mechanics, and more particularly dynamics.

### 4.2 Retargeted motion versus modeled motion

To create this flexible virtual humans with individualities, there are mainly two approaches.

The first approach consists in recording the motion (see Fig. 3) using motion capture systems (magnetic or optical), then to try to alterate such a motion to create this individuality. This process is tedious and there is no reliable method at this stage. Even if it is fairly easy to correct



**Fig. 3.** Motion capture session

one posture by modifying its angular parameters (with an inverse kinematics engine, for instance), it becomes a difficult task to perform this over the whole motion sequence while ensuring that some spatial constraints are respected over a certain time range, and that no discontinuities arise. When one tries to adapt a captured motion to a different character, the constraints are usually violated, leading to problems such as the feet going into the ground or a hand unable to reach an object that the character should grab. The problem of adaptation and adjustment is usually referred to as the motion retargeting problem.

Witkin and Popovic [16] proposed a technique for editing motions, by modifying the motion curves through warping functions and produced some of the first interesting results. In a more recent paper [17], they have extended their method to handle physical elements, such as mass and gravity, and also described how to animate characters with different numbers of degrees of freedom. Their algorithm is based on the reduction of the character to an abstract character, which is much simpler and only contains the degrees of freedom that are useful for a particular animation. The edition and modification are then computed on this simplified character and mapped again onto the end user skeleton. Bruderlin and Williams [18] have described some basic facilities to change the animation, by modifying the motion parameter curves. The user can define a particular posture at time  $t$ , and the system is then responsible for smoothly blending the motion around  $t$ . They also introduced the notion of the motion displacement map, which is an offset added to each motion curve. The motion retargeting problem term was brought up by Michael Gleicher [19]. He designed a space-time constraints solver, into which every constraint is added, leading to a big optimization problem. He mainly focused on optimizing his solver, to avoid enormous computation time, and achieved very good results. Bindiganavale and Badler [20] also addressed the motion retargeting problem, introducing new elements: using the zero-crossing of the second derivative to detect significant changes in the motion, visual attention tracking (and the way to handle the gaze direction) and applying inverse kinematics to enforce constraints, by defining six sub-chains (the two arms and legs, the spine and the neck). Finally, Lee and Shin [21] used in their system a coarse-to-fine hierarchy of B-splines to interpolate the solutions computed by their inverse kinematics solver. They also reduced the complexity of the IK problem by analytically handling the degrees of freedom for the four human limbs.

### 4.3 Creating computational models

The second approach consists in creating computational models that are controlled by a few parameters. One of the major problems is to find such models and to compose them to create complex motion. Such models can be created for walking or grasping objects, but also for groups

and crowds. Let us take the example of walking, where models have been proposed by many researchers. Walking has global and specific characteristics. From a global point of view, every human walk has comparable joint angle variations. However, at a close-up, we notice that individual walk characteristics are overlaid to the global walking gait.

For example, Boulic et al. proposed the walking engine described in [22] which has been extended in the context of a European project on virtual human modeling [23]. Their contribution consists in integrating the walking engine as a specialized action in the animation framework. Walking is defined as a motion where the center of gravity alternatively balances from the right to the left side. It has the following characteristics.

- At any time, at least one foot is in contact with the floor, the 'single support' duration (ds).
- There exists a short instant during the walk cycle, where both feet are in contact with the floor, the 'double support' duration (dds).
- It is a periodic motion that has to be normalized in order to adapt to different anatomies.

The joint angle variations are synthesized by a set of periodic motions, which we briefly mention here.

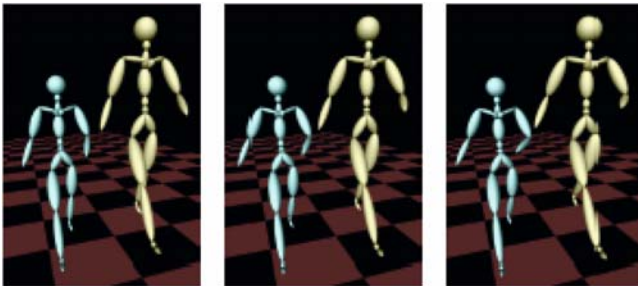
- Sinus functions with varying amplitudes and frequencies for the humanoid's global translations (vertical, lateral and frontal) and the humanoid's pelvic motions (forward/backward, left/right and torsion).
- Periodic functions based on control points and interpolating Hermite splines. They are applied to the hip flexion, knee flexion, ankle flexion, chest torsion, shoulder flexion and elbow flexion.

### 4.4 Generating modeled motion engines

One promising way is the development of a methodology to generate generic real-time motion. This is the approach developed by Glardon et al. [24] for the simulation of human locomotion. Any motion engine needs input data on which it will create new motion by interpolation and extrapolation. The first step consists of getting through motion capture a database of walking and running motions at various speed and height jumps (without initial speed).

In this approach, the principal component analysis (PCA) method is used to represent the motion capture data in a new, smaller space. As the first PCs (principal components) contain the most variance of the data, an original methodology is used to extract essential parameters of a motion. This method decomposes the PCA in a hierarchical structure of sub-PCA spaces. At each level of the hierarchy, an important parameter (personification, type of motion, speed) of a motion is extracted and a related function is elaborated, allowing not only motion interpolation but also extrapolation. Moreover, effort achieved

concerning the possibility not only to qualify a parameter but also to quantify it, improves this method in comparison to other similar works. Generic animation, applicable to any kind of human size, is another important aspect of such research. The method cannot only normalize a generated motion, it may animate humans with a large range of different size. In order to have a complete locomotion engine, transitions between different motions should be handled. The PCA structure offers the capability of performing transition from a specific movement to another one in a very smooth and efficient way. Indeed, as coefficient vectors represent a motion, interpolation between these vectors is sufficient to generate intermediate motion, without computing into the joint angle space. Glatton et al. applied these new concepts on a home-made motion database composed of walking and running motion on a treadmill, at different specific speeds. They developed an engine proposing real-time variation of high-level parameters (style, speed, type of motion, human size), generating a smooth animation (see Fig. 4).



**Fig. 4.** Postures at identical moment in the walk cycle, with two different skeleton sizes

#### 4.5 Adding “idle” motion

Although virtual models of humans continue to improve both in real-time and non-real-time applications, controlling and animating them realistically still remains a difficult task. In nature there exists no motionless character, while in computer animation we often encounter cases where no planned actions, such as waiting for another actor finishing his/her part, is implemented as a stop/frozen animation. We identify many situations where a flexible idle motion generator can help: from synchronization of speech/body animation duration, to dynamic creation of stand still variations in between two active plays. Additionally, as every person has a unique way of moving and standing, an individualized motor control system is required to transfer this property to virtual human motions. From now on, we will call the motions that occur in waiting/idle states between animation clips idle motions. These motions include changing balance because of fatigue, small variations in body posture caused by small muscle contractions or eye blinking. Apart from generat-

ing random movements for separate joints, another possibility is to take captured animation files as a basis for generating idle motions. This results in idle motions that affect all joints, but unfortunately, they are very repetitive and inflexible. Secondly, this approach generates transition problems between the animation clips. Figure 5 shows an example.



**Fig. 5.** Virtual humans waiting at a tram stop in Geneva

Egges et al. [25] propose a novel animation approach based on principal component analysis that allows generating two layers of subtle motions: small posture variations and personalized change of balance. Such a motion generator is needed in many cases when one attempts to create an animation sequence out of a set of existing clips. In nature there exists no motionless character, while in computer animation we often encounter cases where no planned actions, such as waiting for another actor finishing his/her part, is implemented as a stop/frozen animation. They identify many situations where a flexible idle motion generator can help: from synchronisation of speech/body animation duration, to dynamic creation of stand still variations in between two active plays. Their approach overcomes the limitations of using a small set of existing clips as a basis for synthesizing idle motions, such as unnatural repetition of movements and difficulties to insert idle motions into an animation without breaking its continuity. A realistic animation is obtained by blending small posture variations with personalized balance shifting animations.

## 5 Generating complex behaviours of virtual humans inside their environments

### 5.1 Behavioural models

By behavioral animation we refer to the techniques applied to the synthesis of autonomous animation of virtual

characters. Autonomy is one of the most priced goals in the field of character animation. We generally consider autonomy as the quality or state of being self-governing. Perception of the elements in the environment is essential, as it gives the agent the awareness of what is changing around it. It is indeed the most important element that one should simulate before going further. Most common perceptions include (but are not limited to) simulated visual and auditory feedback. Adaptation and intelligence then define how the agent is capable of reasoning about what it perceives, especially when unpredictable events happen. On the other hand, when predictable elements are showing up again, it is necessary to have a memory capability, so that similar behaviour can be selected again. Lastly, emotion instantaneously adds realism by defining affective relationships between agents.

The subtle and spontaneous gestures are part of low-level behaviors, and they are essential to communicate any message and convey emotions. These ideas are just beginning to be explored by the scientific community, some studies focus on non-human characters, like dogs or other virtual animals [26], while others put special attention on human-like facial gestures.

Continuous animation (low level behavior) is the result of proprioception (this includes motivations, internal “mental” or emotional states, etc.), and reactions to external stimuli.

The low-level behavioral animation is an open research avenue. Creating virtual characters able to select the action to perform in an automatic way is not new. State of the art autonomous characters and virtual humans are able to perform a variety of tasks, and have some level of autonomy in their decision making, but still they don’t look like real people. One of the problems is that the virtual humans do not display the subtle gestures and mannerisms that characterize a real human being: facial expressions, body language, etc.

Several studies and implementations have been done with the objective of giving a virtual character the ability to perceive its virtual environment and react to it by means of executing adequate tasks [27, 28], creating the illusion that the synthetic character is alive: it moves by itself and achieves relatively complex tasks. The autonomous virtual human explores an unknown environment constructed on mental models as well as a “cognitive map” based on this exploration. Navigation is carried out in two ways: globally (pre-learned model of the virtual environment, few changes and the search for performance with a path planning algorithm) and locally (direct acquisition of the virtual environment.) A 3D geometrical model in the form of a grid is implemented with the help of an octree combined with the approach proposed by Noser [29] and Kuffner [30]. Nevertheless, this approach requires a filtering of irrelevant information from the variable positions of the virtual sensors, Elfes [31]. A virtual human is situated in a virtual environment (VE) equipped with sensors

for vision, audition and touch, informing it of the external VE and its internal state. A virtual human possesses effectors, which allow it to exert an influence on the VE and a control architecture, which coordinates its perceptions and actions. The behaviour of a virtual human can be qualified as being adaptive as long as the control architecture allows it to maintain its variables in their validity zone. Learning methodologies can modify the organization of the control architecture. We will then integrate them in the further pursuit of our research. Conde and Thalmann propose an artificial life environment framework equips a virtual human with the main virtual sensors, based on an original approach inspired by neuroscience in the form of a small nervous system with a simplified control architecture to optimise the management of its virtual sensors as well as the virtual perception part. The processes of filtering, selection and simplification are carried out after obtaining the sensorial information; this approach allows us to obtain some persistence in the form of a “cognitive map” and also serves as a pre-learning framework to activate learning methods of low and high level concerning the behaviour of a virtual human.

The objective pursued is to allow the virtual human to explore virtual environments until then unknown and to construct mental structures and models, cognitive maps or plans from this exploration. Once its representation has been constructed, this knowledge can be passed on to other virtual humans. The autonomous virtual human perceives objects and other virtual humans with the help of its virtual environment, which provides the information concerning the nature and position of the latter. The behavioural model to decide which actions the Virtual human should take such as walking, handling an object, etc., and then uses this information.

## 5.2 Perception and sensors

The Conde and Thalmann [32] approach is based on the multi-sensory integration of the standard theory of neuroscience, where signals of a single object coming from distinct sensory systems are combined. The acquisition steps of signals, filtering, selection and simplification intervening before proprioception, active and predictive perception are integrated into virtual sensors and a virtual environment. We will focus on two aspects: 1) the assignment problem: determining which sensory stimuli belong to the same virtual object, and 2) the sensory recoding problem: recoding signals in a common format before combining them. We have developed three novel methodologies to map the information coming from the virtual sensors of vision, audition and touch, as well as that of the virtual environment in the form of a “cognitive map”.

An autonomous virtual agent (AVA) situated in a virtual environment (VE) is equipped with sensors for vision, audition and touch that inform it of the external VE and its internal state. An AVA possesses effectors to let it exert



an influence on the VE and control architecture to coordinate its perceptions and actions. The behaviour of an AVA is adaptive as long as the control architecture allows it to maintain its variables in their validity zone.

Artificial life environment (ALifeE) based on an original approach inspired by neuroscience equips an AVA with the main virtual sensors in the form of a small nervous system [1]. The control architecture is kept simple to optimise the management of the AVA's virtual sensors and perception. The processes of filtering, selection and simplification are carried out after obtaining the sensorial information. This approach allows us to achieve some persistence in the form of a "cognitive map".

The "mental processes" of an AVA can be simulated. Behavioural animation includes the techniques applied to make an AVA intelligent and autonomous, to react to its VE and to make decisions based on its perceptive system, its short-term memory and long-term reasoning. Intelligence is the ability to plan and carry out the tasks based on the model of the current state of the VE.

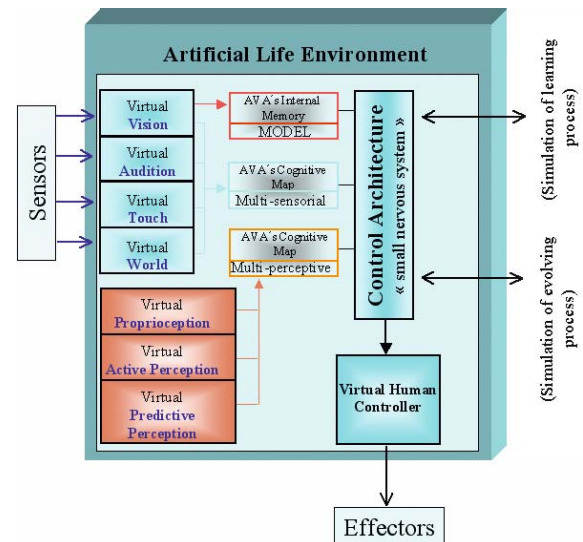
Our objective is to permit the AVA to explore unknown VEs and to construct mental structures and models, cognitive maps or plans from this exploration. Once its representation has been created, the knowledge can be communicated to other AVAs. Each AVA perceives objects and other AVAs with the help of its VE, which provides information concerning their nature and positions. The behavioural model decides which action the AVA should take (such as walking or handling an object) and then uses the knowledge.

In a VE, an AVA requires a combination of perception and action to behave in an autonomous way. The perception system provides a uniform interface to various techniques in the field of virtual perception, including synthetic vision, audition and touch. In usual approaches, different behaviours implement their own perception mechanisms, which lead to computation duplication when multiple behaviours are involved. Basically, an AVA maintains a set of perception pipelines, each corresponding to a particular type of virtual sensor. A pipeline is composed of filters that coordinate themselves in order to extract relevant information from the data sent by the associated virtual sensor.

Figure 6 shows a schematic representation of such an artificial life organization.

### 5.3 A case study: interacting with objects

One of the challenge in simulating human behaviour is to correctly model interactions between an object and a virtual human agent (hereafter just referred to as an agent), it appears in most applications of computer animation and simulation. Such applications encompass several domains, for example, virtual autonomous agents living and working in virtual environments, human factors analysis, training, education, virtual prototyping, and simulation-based



**Fig. 6.** A schematic representation of our ALifeE. Virtual vision discovers the VE, constructs the different types of perception and updates the AVA's cognitive map to obtain a multi-perceptive mapping. Then the control architecture uses both the "cognitive maps" and the "memory model" to interact with the learning, development, and control processes of the AVA (virtual human controller)

design. A good overview of such areas is presented by Badler [33]. An example of an application using agent-object interactions is presented by Johnson et al. [34], whose purpose is to train equipment usage in a populated virtual environment.

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.

To overcome such difficulties, a natural way is to include within the object description, more useful information than only intrinsic object properties. Some proposed systems already use this kind of approach. In particular, the object specific reasoning [35] creates a relational table to inform object purpose and, for each object graspable site, the appropriate hand shape and grasp approach direction. This set of information may be sufficient to perform a grasping task, but more information is needed to perform different types of interactions.

Another interesting way is 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 [36] 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 [37] is used to include all desired information in objects. A graphical interface program permits the user to interactively specify different features in the object, and save them as a script file.

## 6 Creating realistic groups and crowds of virtual people

Another challenge in virtual humans is the simulation of larger numbers of virtual human agents for interactive virtual environments such as virtual reality training systems [38]. In recent years, virtual crowds have become a more and more common element of our cinematic experience. Whether it was a photo-realistic crowd of digital passengers in the Titanic, a legion of animated ants in AntZ, or a rendered army of droids in Star Wars, computer graphics (CG) generated crowds of characters added significantly to the impact of the movies employing them, allowing us to visualize scenes not possible only some years earlier. Crowds of CG human and non-human characters help to overcome prohibiting costs and complexities of working with large numbers of extras, stand as virtual stuntmen in dangerous shots and integrate well with the virtual scenes.

Several works in different fields have been exploring issues connected to the domain of crowd simulations. In his pioneer work, Reynolds [39] described a distributed behavioral model for simulating aggregate motion of a flock of birds. Bouvier and Guilloteau [40] used a combination of particle systems and transition networks to model human crowds in visualization of urban spaces. Brogan and Hodgins [41] simulated group behaviors for systems with significant dynamics. Aubel and Thalmann [42] introduced dynamically generated impostors to render virtual humans. Tecchia et al. [43] proposed an image-based method for real-time rendering of animated crowds in virtual cities. McPhail et al. [44] studied individual and collective actions in temporary gatherings. Still [45] used mobile cellular automata for simulation and analysis of crowd evacuations. However, only a few works tried to explore more general crowd models integrating several sub-components, such as collision avoidance, path-planning, higher-level behaviors, interaction or rendering.

With current consumer-grade personal computers it is possible to display 3D virtual scenes with thousands of animated individual entities at interactive frame rates using different techniques as animated impostors [46]. When increasing the number of involved individuals it is becoming more difficult to create unique and varied content of scenarios with large numbers of entities. If we want to create or modify features of every single individual one by one, it will soon become too laborious. If, on the other hand, we apply a set of features (either uniform, or patterned) to many individuals at once, it could create unwanted ar-

tifacts on a larger scale, resulting in an “army-like” appearance with too uniform, or periodic distributions of individuals or characteristics.

The use of random distributions can alleviate such problems; however, it can be very difficult to capture the desired constraints into a set of mathematical equations, especially considering integration into common art production pipelines.

The challenge is how to create complex scenes resembling a variety-rich look of the real world. Bottom-up approaches, such as local rule based flocking [39] can create such complexity, however they are difficult to control if we want to achieve particular end configurations (how to set local rules to get a global result). In recent work, Anderson et al. [47] achieved interesting results for a particular case of flocking animation. Nevertheless, the algorithm can get very costly when increasing the number of entities and simulation time.

Major 3D content creation packages used by the media industry now offer tools to improve working with a large number of virtual characters. The production of massively populated scenes is still in the majority of cases a lengthy and manual-intervention intensive process, operating mostly in a non-real-time mode.

In previous works, crowds have been considered as already formed units with more or less uniform behavior placed in particular environments corresponding only to the limited purpose of the simulation, e.g. pedestrians just fleeing from burning building [48] or a marching crowd during a demonstration, the crowd assigning people to particular groups and determining their behaviors. Collective behavior emerges from interaction of individuals, crowds are dynamically assembled and disassembled and over the time they change their behavior. Musse and Thalmann take inspiration from the field of sociology. McPhail [49] argues that members of a gathering do not act as a whole, but rather that subsections of the number assembled react in similar way, which, when observed externally appears as a collective behavior. In our system, a crowd is modeled as collection of individuals, which react to the environment, other agents and real human participants of the simulation and can have very different behaviors both for one agent in different situations and for many agents in the same situation.

Ulicny and Thalmann [50] implemented our emergent crowd system to reproduce simple simulation consists of group of autonomous virtual human agents existing in dynamic virtual 3D environment. In order to behave in believable way, these agents have to act in accordance with their surrounding environment, be able to react to its changes, to the other agents and also to the actions of real humans interacting with the virtual world. Agents contain a set of internal attributes corresponding to various psychological or physiological states (e.g. level of curiosity, tiredness, etc.), a set of higher-level complex behaviors and a set of rules determining selection of these behav-

iors. Events provide a way of agents' interaction with their environment, other agents or human participants of the simulation. Each agent is able to receive events from the environment objects, other agents or user interface. Combinations of different received events and different levels of agent's attributes produce both changes of its internal attributes and change of the overt behavior.

The behavior model is based on a combination of rules [51, 52] and finite state machines [53] for controlling the agent's behavior using a layered approach. The first layer deals with the selection of higher-level complex behavior appropriate to the agent's situation, the second layer implements these behaviors using low-level actions provided by the virtual human [54]. At the higher level, rules select complex behaviors (such as fleeing) according to the agent's state (constituted by attributes) and the state of the virtual environment (conveyed by events). In rules we specify for who (e.g. which particular agent, or agents in a particular group) and when the rule is applicable (e.g. at defined time, after receiving event or when some attribute reached specified value), and what is the consequence of rule firing (e.g. change of agent's high-level behavior or attribute). An example of such rule is:

```
FOR ALL
WHEN EVENT = in_danger_area AND ATTRIBUTE fear > 50%
THEN BEHAVIOR FLEE
```

The variety of the reactions to the same situation is achieved by different agents having different values of the attributes (at the beginning through different initializations, later because of their different histories), which consequently leads to different rules being triggered.

At the middle level, high-level behaviors are implemented using hierarchical finite state machines. Each behavior is realized by one FSM that drives selection of the low-level actions for virtual human (like movement of location, playing a short animation sequence), manages connections with the environment (like path queries, or event sending) and also can call other FSMs to delegate subtasks such as path following. There are two types of high-level behaviors. First, we can specify scripted behavior which is more precise, but less autonomous and with less environment coupling by using explicit sequences of low-level actions. Or second, we can let agents perform autonomously complex behaviors with feedback from the environment. Examples of such autonomous behaviors are wandering, fleeing, neutralizing the threat, or requesting and providing help. Both types can be mixed as needed. Figure 7 shows an example.

At the lowest level, motion control is also organized into layers. As the result of higher-level behavior, agent's behavior FSM decides (or is told directly by the rule) that the agent wants to move to particular location. This request is forwarded to the path-finding layer, which constructs sequence of waypoints that need to be passed to get to the location. Finally, it is the responsibility of the



Fig. 7. Real-time crowd at a dance party

collision-avoidance layer to move agent between waypoints correcting trajectory in order to avoid collisions.

Ulicny et al. [55] propose an approach to give full creative power to designers using metaphors of artistic tools, operating on a two-dimensional canvas familiar from image manipulation programs working in WYSIWYG (what you see is what you get) mode, with a real-time view of the authored scene. The advantages of immediate feedback, intuitive interface and familiarity allows us to better express the artist's vision and can also lead to an increase in productivity. The goal is to create a system that would allow authoring of freely navigable real-time 3D scenes, composed of a large number of varied animated individuals in a virtual environment. The authoring should be simple and intuitive, usable by non-programmers. They take inspiration from the domain of image and word processing, where most of the applications use WYSIWYG approaches with immediate feedback of the resulting manipulations. The idea is simple: the designer manipulates virtual tools, working in a two-dimensional screen space, with a mouse and a keyboard. These tools then affect the corresponding objects in a three-dimensional world space (see Fig. 1). Different tools have different visualizations and perform different effects on the scene including creation and deletion of crowd members, changing of their appearances, triggering of various animations, setting of higher-level behavioral parameters, setting waypoints for displacement of the crowd, or sending of events to the behavior subsystem.

## 7 Believable interaction between real and virtual humans

### 7.1 Facial communication

In 1971, Parke produced a representation of the head and face at the University of Utah, and three years later ad-

vances in good enough parametric models to produce a much more realistic face [56].

Nowadays, a lot of interest from both industry and research exists for virtual environments (VEs) and embodied conversational agents (ECAs). A lot of new techniques are being developed to improve the simulation in general, to add more visual detail and make the interaction between human and VE/ECA more natural. Believability is a measure to help to determine how well these different techniques are working. Believability represents the 'outsider' point of view. It is thus a very powerful evaluation tool, since we can use it to make evaluations of different techniques/methods while the evaluations are independent from the underlying techniques. This allows us to compare different approaches and give a meaningful indication of their quality on the level of believability. In this section, we will focus especially on believability and ECAs. Since ECAs are generally modeled after humans (even cartoon characters), one important aspect of their believability is how well an ECA succeeds in acting like a human.

We believe that the key to believable ECAs is the definition of their personality and their emotions. Although quite some research has been done to describe the influence of emotion and personality on ECAs, the results until now are not very convincing. We see several reasons for this.

- Psychological models of emotion/personality are not finalized. The exact structure of our emotions/personality is not certain, nor is the way in which emotions and personality interact with our perception, behaviour and expression.
- When one wants to simulate emotions/personality computationally, one tends to take the model the most suitable for a computational simulation. However, this model is not necessarily the best representation for emotions/personality.
- Even if there exists a perfect emotion/personality model, it is very difficult to distinguish the resulting behaviour from the emotion that is behind it.

Also, other issues interfere with our impression of how emotional an ECA really is, such as its appearance, the surroundings, its capabilities, and so on.

## 7.2 Emotion, personality and perception

In emotion simulation research so far, appraisal is popularly done by a system based on the OCC model [57]. This model specifies how events, agents and objects from the universe are used to elicit an emotional response depending on a set of parameters: the goals, standards and attitudes of the subject. Since the emotional response is generated from a cognitive point of view, this type of appraisal is called cognitive appraisal and it corresponds closely with Lazarus' emotion theory (not taking into account the physiological response). When one wants to develop a computational model of appraisal, not all of the

above mentioned scenarios are suitable to take as a basis, especially those scenarios where arousal plays a crucial role in the determination of the emotional response (ECAs do not yet have a physiology). This raises the question if it is possible to develop a computational model of appraisal that has a high believability. On the level of personality, one could consider the goals, standards and attitudes of the OCC model as a domain dependent 'personality'. However, personality can also be modeled in a more abstract, domain-independent way [58, 59]. Egges et al. [60] discuss how a link between multidimensional personality models and the OCC appraisal model can be established.

The effect of personality and emotion on agent behavior has been researched quite extensively, whether it concerns a general influence on behavior [61], or a more traditional planning based method [62]. Also, rule-based models [63], probabilistic models [64, 65] and fuzzy logic systems [66] have been developed. In the case of real human beings, there are still many questions regarding how emotion influences our behavior, but in the field of Neuroscience, work has been done that partly describes the relationship between emotions and the brain [67, 68].

## 7.3 Emotion, personality and expression

The expression of emotions (see Fig. 8) has been widely researched, and the most well known research is the work done by Ekman [69]. Not only will personality and emotion have an effect on expressions by the face or body; also physiological changes can be measured according to different emotions. Furthermore, emotions and personality have an important effect on speech [70]. In the following two sections, we will concentrate on the relationship between emotions, personality and face/body animation. Also, we will give some examples on how to improve the believability of an ECA using emotions and personality.

When communicating with an ECA, the dialogue itself is only a small part of the interaction that is actually going on. In order to simulate human behavior, all the non-verbal elements of interaction should be taken into account. An ECA can be defined by the following parts.

- Appearance (face model, age, race, etc.).
- Behaviour (choice of non-verbal behaviour, accompanying speech).
- Expressiveness of movement (amplitude, tempo, etc.).

Other information can also be important, like the cultural background or the context. Facial animation synchronized with speech can be improved by different factors such as nonverbal actions, speech intonation, facial expression consistent with the speech and the context, and also facial expressions between speech sequences. All this information helps to increase the believability of facial animation for ECAs. The main problem is to determine when and how this kind of non-verbal behavior should be expressed.





Fig. 8. Various expressions for the virtual human

Finally, one of the most important points for increasing believability of facial and body animation is the synchronization between verbal and non-verbal expressions [71, 72].

## 8 Generating realistic virtual clothed and haired people

Even when virtual humans have nice bodies and faces, and behave realistically, they still need a realistic appearance in terms of hair, skin, and clothes.

### 8.1 Cloth animation

Dressing a virtual body involves designing complex garment shapes, as well as an advanced simulation system able to detect and to handle multiple collisions generated between the cloth and the body. Very accurate methods are used when simulating complex garments for haute couture for instance, whereas less accurate but interactive ones are used for design and manipulation.

Along with the evolution of cloth simulation techniques, focus was primarily aimed to address realism through the accurate reproduction of the mechanical features of fabric materials. The early models, developed a decade ago, had to accommodate very limited computational power and display device, and therefore were geometrical models that were only meant to reproduce the geometrical features of deforming cloth [73]. Then, real mechanical simulation took over, with accurate cloth models simulating the main mechanical properties of fabric. While some models, mostly intended for computer

graphics, aimed at simulating the complex garments used for dressing virtual characters, [74], [75], other studies focused on the accurate reproduction of mechanical behavior, using particle systems [76] or finite elements [77]. Despite all these developments, all these techniques remain dependent on high computational requirements, limiting their application along the new trends toward highly interactive and real-time applications brought by the highly spreading multimedia technologies. While highly accurate methods such as finite elements are not suitable for such applications, developments are now focusing toward approximate models, that can render, using minimal computation, approximate, but realistic results in a very robust way.

The most popular way for modeling the mechanics of clothes are spring-mass particle systems. They are composed of 3D punctual masses, connected to their neighbors by springs. The mechanical behaviors of clothes are simulated by computing forces applied on each mass given their position and velocity. There are different ways to complete the numerical integration.

Explicit integration computes the next state of clothes at time  $t + h$  by evaluating the forces at time  $t$  such as the Runge–Kutta family methods. Explicit integrations are fast methods to compute (for first-order ones), but they show instability for large time-steps. Another family is implicit integration or backward integration: it uses the predictive derivative for the next timestep and then deduces the state of clothes for the current one. This method allows a larger timestep, while conserving numerical stability [78]. Obviously, despite their stability, implicit methods are not necessarily accurate, particularly for large timesteps. Actually, many dynamical effects may disappear in these approximations such as animated wrinkles. In the case of explicit methods, accurate computation of the motion of every particle requires a lot of computational resources, but the result will be very accurate through the use of high-order methods.

There are hybrid solutions, which take advantage of both integration methods: stiff elements are handled with the time-consuming implicit method, while the others are integrated explicitly [79]. The fashionizer platform approach [80] combines the flexibility obtained in [81] with simulation speeds aimed in [82], which are restricted to regular meshes. Furthermore, a suited implicit integration method has been associated to this model to maximize simulation timesteps and computation speeds without trading away mechanical stability.

Another family of non-physically based methods tries to give the look of real clothes without the expensive cost of physical ones. These methods are well-suited for real-time issues. Kang et al. [83] improved the visual quality of very low detailed (in terms of polygons) garments by tessellating the polygons with cubic spline curves and succeeded in simulating wrinkles (see Fig. 9). Unfortunately, it exhibits undesirable results with a high number of poly-



**Fig. 9.** Simulation using spline curves [83] and real-time simulation of fabrics by [87]

gons that is essential for simulating complex garments. Another interesting approach tries to take advantage of a bump or displacement map technique coupled to a particle system for generating global movement to simulate fast the cloth. However, it requires a lot of designer interaction [84].

Pure geometric methods aim at providing faster results, using geometric primitives such as catenary's curves or surfaces (Fig. 5), but they are not appropriate for cloth movement reproduction [85]. Statistical-based approaches have been introduced very recently and show promising results in a wide range of uses. The main idea is to construct a simulator that will learn cloth deformations through a set of examples, to increase the speed of the computation [86, 87]. A very interesting method [88, 89] achieves real-time simulation. It is based on a segmentation of clothes into three areas and a hybrid simulation using both data-driven and physically-based simulation and was used in the LIFEPLUS project [90] (see Fig. 10).

## 8.2 Hair animation

One of the many challenges in simulating believable virtual humans and animals has been to produce realistic looking hair. Hair simulation can be thought of as having three different subtasks: shape modeling, dynamics and rendering. We can even classify the hair modeling attempts based on their underlying models, namely explicit hair models, cluster hair models and volumetric textures. Hair rendering and shape modeling of fur-like short hair is becoming increasingly available to animators. However, shape modeling and dynamics of long hair has been difficult. The difficulties stem from the sheer number of hair strands, their geometric intricacies and associated complex physical interactions.

Explicit models are tedious for shape modeling and numerically intensive for dynamics, however they are intuitive and close to reality. They are especially suitable for the dynamics of long hair. Daldegan et al. [91] and Rosenblum et al. [92] used a mass-spring-hinge model to



**Fig. 10.** Real-time simulation of a wearing cloth by [3, 12]

control the position and orientation of hair strand. As their dynamic models were for straight hair, they could not be applied in animating hairstyles. Anjyo et al. [93] modeled hair with a simplified cantilever beam and used a one-dimensional projective differential equation of angular momentum to animate hair. Although the algorithm dealt with both hairstyling and hair animation, it had several limitations; the original hairstyle could not be recovered after subsequent animation and the movement of the head was not accounted for. Further, the approximation in

collision detection could generate unnatural results when the algorithm was used to animate long hair. To reduce computations, Daldegan et al. [91] used a wisp model. Recently, Lee et al. [94] developed on Anjyo's work to add some details to model hairstyles. Appreciably, none of the previous attempts considered hair-hair and hair-air interactions. Even individual hair dynamics was grossly approximated to suit available computational power. However, in recent years the computing power has grown considerably. Supercomputing power of the past is becoming increasingly available to animator's workstations.

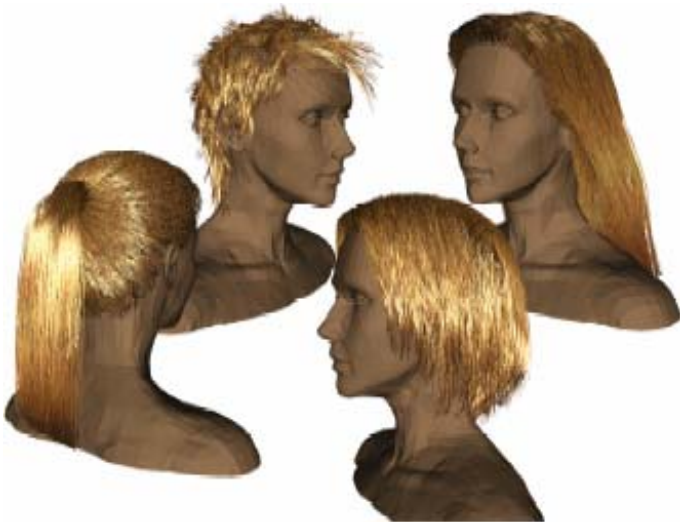
Hair-hair interaction is probably the most difficult problem in achieving visually pleasing hair dynamics. So far, no good models have been developed in this regard. Though there are many advances in collision detection and response [95], they are simply unsuitable for the problem at hand, because of sheer number complexity of hair. Hadap et al. [96] take a radical approach by considering hair as a continuum. The continuum assumption states that the physical properties of a medium such as pressure, density and temperature are defined at each and every point in the specified region. Fluid dynamics regards liquids and gasses as a continuum and even elastic theory regards solids as such, ignoring the fact that they are still composed of individual molecules. Indeed, the assumption is quite realistic at a certain length scale of the observation, but at smaller length scales the assumption may not be reasonable. One might argue that hair-hair spacing is not at all comparable to inter molecular distances to consider hair as a continuum. However, individual hair-hair interaction is of no interest to us apart from its end effect. Hence, we treat the size of individual hair and hair-hair distance as being much smaller than the overall volume of hair, justifying the continuum assumption. As we develop the model further, it will be apparent that the above assumption is not just about approximat-

ing the complex hair-hair interaction. An individual hair is surrounded by air, as it moves it generates a boundary layer of air. This influences many other hair strands in motion. This aerodynamic form of friction is comparable to hair-hair contact friction. In addition, there are electrostatic forces to take part in the dynamics. It is not feasible to model these complex multiple forms of interactions. This inspires us to consider the dynamics of single hair interacting with other surrounding hairs in a global manner through the continuum assumption. That way, we hope to have a sound model for an otherwise very complex phenomenon. Figure 11 shows examples.

As we start considering hair as a continuum, we look at the properties of such a medium, namely the hair medium. There are two possibilities: hair medium could be considered as a solid or a liquid, depending on how it behaves under shearing forces. Under shearing stresses, solids deform till they generate counter stresses. If the shearing stresses are removed, the solids exhibit the ability of retaining their original shape. The liquids are not able to withstand any shearing stresses. Under the influence of the shearing stresses they continue to deform indefinitely and they do not have any shape memory. In the case of hair, if we apply a lateral shearing motion it acts like a liquid. At the same time, length wise, it acts as a solid. Thus there is a duality in the behaviour of hair as a continuum.

However, from an animation point of view, we cannot treat hair solely as a continuum, unless the viewpoint is far enough and individual hair movement is not perceived. Thus, we have to retain the individual character of hair as well, while considering hair as a continuum. We split hair dynamics into two parts.

- Hair-hair, hair-body and hair-air interactions, which are modeled using continuum dynamics, and more precisely fluid dynamics.



**Fig. 11.** Various hairstyles animated by our real-time simulation system





**Fig. 12.** Crowd in virtual park: a) before emergency, b) after gas leak

- Individual hair geometry and stiffness, which are modeled using the dynamics of an elastic fiber. Interestingly, this approach even addresses the solid liquid duality effectively. The model can be visualized as a bunch of hair strands immersed in a fluid. The hair strands are kinematically linked to fluid particles in their vicinity. The individual hair has its own stiffness dynamics and it interacts with the environment through the kinematical link with the fluid. Density, pressure and temperature are the basic constituents of fluid dynamics. The density of the hair medium is not precisely the density of individual hair. It is rather associated with the number density of hair in an elemental volume.

## 9 Conclusion

Should virtual humans become complex like real humans? This is an important question. In fact, this is context-dependent. As a guide to find information on the web: we just want an efficient agent. However, we are not interested in dealing with a virtual human that has problem with his

children or complains about his/her headaches. However, in a simulation of behavior in the case of emergency (see Fig. 12), we want believable humans: panic, stress, and leadership are the key issue. It is essential to take into account psychological aspects.

Finally, the goal of the research effort in virtual humans is to develop believable autonomous virtual humans in virtual environments. In this context, believability means that the behavior of a virtual human has to be indistinguishable from that of a real human. For instance, an actor inhabiting a typical virtual environment composed of objects and other actors, only concerned with performing a given task without reacting to objects or actors it encounters or to what these actors are doing would not look believable. The behavior of these actors would be more that of a robot than that of a human. To be believable, an actor has to be affected by what takes place around it and needs to engage in social behaviors with other actors. Therefore, the behavioral model of an actor needs to be versatile enough to allow a change of behavior, the emotional state of the actor must be reflected and must affect its behavior, the interpersonal relationships with the other actors must be taken into account and possibly bring the actor to engage in social interactions.

## References

1. Magnenat-Thalmann, N., Thalmann, D. (eds): *Handbook of Virtual Humans*. New York, John Wiley (2004)
2. Aubel, A., Thalmann, D.: MuscleBuilder: a modeling tool for human anatomy. *J Comput Sci Technol* **19**(5), 585–595 (2004)
3. Magnenat-Thalmann, N., Thalmann, D.: The Direction of Synthetic Actors in the film *Rendez-vous à Montréal*, *IEEE Computer Graphics and Applications* **7**(12), pp. 9–19 (1987)
4. Chadwick, J., Haumann, D., Parent, R.: Layered construction for deformable animated characters. *Computer Graphics (SIGGRAPH '89 Proceedings)*, pp. 243–252 (1989)
5. Henne, M.: A constraint-based skin model for human figure animation. Master's Thesis, University of California, Santa Cruz, June (1990)



6. Yoshimoto, S.: Ballerinas generated by a personal computer. *J Visualizat Comput Animation* **3**, 85–90 (1992)
7. Thalmann, D., Shen, J., Chauvineau, E.: Fast realistic human body deformations for animation and vr applications. *Computer Graphics International'96*, Pohang, Korea (1996)
8. Wilhelms, J.: Animals with anatomy. *IEEE Comput Graph Appl* **17**(3), pp. 22–30 (1997)
9. Scheepers, F., Parent, R.E., Carlson, W.E., May, S.F.: Anatomy-based modeling of the human musculature. *Computer Graphics (SIGGRAPH'97 Pro-ceedings)*, pp. 163–172 (1997)
10. Porcher Nedel, L., Thalmann, D.: Anatomic modeling of deformable human bodies. *Visual Computer* **16**(6), 306–321 (2000)
11. Aubel, A., Thalmann, D.: Interactive modeling of the human musculature. *Proc. Computer Animation*, Seoul, Korea (2001)
12. Seo, H., Cordier, F., Philippon, L., Magnenat-Thalmann, N.: Interactive modeling of mpeg-4 deformable human body models. In: *Proc. Deform'2000, Workshop on Virtual Humans by IFIP Working Group 5.10 (Computer Graphics and Virtual Worlds)*, pp. 120–131. Kluwer, Dordrecht (2000)
13. Berthoz, A.: *The Brain's Sense of Movement*. Harvard University Press, ISBN 067 400 9800 (2002)
14. Rose, C., Cohen, M.F., Bodenheimer, B.: Verbs and adverbs: multidimensional motion interpolation. *IEEE Comput Graph Appl* **18**(3), 32–40 (1998)
15. Brand, M., Hertzmann, A.: Style machines. *Proc. of SIGGRAPH'00*, New Orleans, and Technical Report TR-2000-14 MERL (2000)
16. Witkin, A., Popovic, Z.: Motion warping. *Proceedings of SIGGRAPH 95*, Los Angeles, August 1995, pp. 105–108. ISBN 0-201-84776-0. Held in Los Angeles, California (1995)
17. Popovic, Z., Witkin, A.: Physically based motion transformation. *Proceedings of SIGGRAPH 99*, Los Angeles, August 2000, pp. 11–20. ISBN 0-20148-560-5
18. Bruderlin, A., Williams, L.: Motion signal processing. In: Cook, R. (ed.) *SIGGRAPH 95 Conference Proceedings*, Annual Conference Series, held in Los Angeles, August, pp. 97–104. ACM SIGGRAPH, Addison Wesley, Reading (1995)
19. Gleicher, M.: Retargeting motion to new characters. In: Cohen, M. (ed.) *SIGGRAPH 98 Conference Proceedings*, Annual Conference Series, pp. 33–42. ACM SIGGRAPH, Addison Wesley, Reading, July 1998. ISBN 0-89791-999-8 (1998)
20. Bindiganavale, R., Badler, N.I.: Motion abstraction and mapping with spatial constraints. In: Magnenat-Thalmann, N., Thalmann, D. (eds.) *Modeling and Motion Capture Techniques for Virtual Environments*. Lecture Notes in Artificial Intelligence, pp. 70–82. Springer, Berlin Heidelberg New York (1998)
21. Jehee, L., Shin, S.Y.: A hierarchical approach. In: *Proceedings of SIGGRAPH 99*, held in Los Angeles, pp. 39–48. ISBN 0-20148-560-5
22. Boulic, R., Magnenat-Thalmann, N., Thalmann, D.: A global human walking model with real-time kinematics personification. *Visual Computer* **6**(6), 344–358
23. Boulic, R., Capin, T., Huang, Z., Mocozet, L., Molet, T., Kalra, P., Lintermann, B., Magnenat-Thalmann, N., Pandzic, I., Saar, K., Schmitt, A., Shen, J., Thalmann, D.: The HUMANOID environment for interactive animation of multiple deformable human characters. *Proc. Eurographics '95*, Maastricht, August 1995, pp. 337–348 (1995)
24. Glardon, P., Boulic, R., Thalmann, D.: PCA-based walking engine using motion capture data. In: *Proceedings of Computer Graphics International*, pp. 292–298, June 2004
25. Egges, A., Visser, R., Magnenat-Thalmann, N.: Example-based idle motion synthesis in a real-time application. *CAPTECH Workshop*, Zermatt, Switzerland, pp. 13–19, December (2004)
26. Isla, D., Blumberg, B.: Object persistence for synthetic creatures. In: *Proceedings of International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, Bologna, Italy, July (2002)
27. Blumberg, B., Galyean, T.: Multi-level direction of autonomous creatures for real-time virtual environments. In: *Proceedings of SIGGRAPH 95* (1995)
28. Tu, X., Terzopoulos, D.: Artificial fishes: physics, locomotion, perception, behavior. *ACM Computer Graphics, Proceedings of SIGGRAPH'94*, July (1994)
29. Noser, H., Renault, O., Thalmann, D., Magnenat Thalmann, N.: Navigation for digital actors based on synthetic vision, memory and learning. *Computers and Graphics* **19**(1), 7–19 (1995)
30. Kuffner, J.-J., Latombre, J.-C.: Fast synthetic vision, memory, and learning models for virtual humans. *Proc. Computer Animation*, IEEE, pp. 118–127 (1999)
31. Elfes, A.: Occupancy grid: a stochastic spatial representation for active robot perception. *6th Conf. on Uncertainty in AI* (1990)
32. Conde, T., Thalmann, D.: An artificial life environment for autonomous virtual agents with multi-sensorial and multi-perceptive features. *Comput Anim Virtual Worlds* **15**(3–4), 311–318 (2004)
33. Badler, N.N.: Virtual humans for animation, ergonomics, and simulation. *IEEE Workshop on Non-Rigid and Articulated Motion*, Puerto Rico, June (1997)
34. Johnson, W.L., Rickel, J.: Steve: an animated pedagogical agent for procedural training in virtual environments. *Sigart Bull* **8**(1–4), 16–21 (1997)
35. Levison, L.: Connecting planning and acting via object-specific reasoning. Ph.D. thesis, Dept. of Computer & Information Science, University of Pennsylvania (1996)
36. Kallmann, M., Thalmann, D.: A behavioral interface to simulate agent-object interactions in real-time. *Proc. Computer Animation 99*, IEEE Computer Society Press (1991)
37. Shah, J.J., Mäntylä, M.: *Parametric and feature-based CAD/CAM*. John Wiley, New York. ISBN 0-471-00214-3 (1995)
38. Ulicny, B., Thalmann, D.: Crowd simulation for interactive virtual environments and VR training systems. *Proc. Eurographics Workshop on Animation and Simulation'01*. Springer, Berlin Heidelberg New York (2001)
39. Reynolds, C.: Flocks, herds and schools: a distributed behavioral model. *Proc. SIGGRAPH'87*. *Comput Graph* **21**(4), 25–34 (1987)
40. Bouvier, E., Guilloteau, P.: Crowd simulation in immersive space management. In: *Proc. Eurographics Workshop on Virtual Environments and Scientific Visualization '96*, pp. 104–110. Springer, Berlin Heidelberg New York (1996)
41. Brogan, D., Hodgins, J.: Group behaviors for systems with significant dynamics. *Auton Robots* **4**, 137–153 (1997)
42. Aubel, A., Thalmann, D.: MuscleBuilder: a modeling tool for human anatomy. *J Comput Sci Technol* **19**(5), 585–595 (2004)
43. Tecchia, F., Chrysanthou, Y.: Real-time rendering of densely populated urban environments. In: *Proc. Eurographics Rendering Workshop*, pp. 83–88 (2000)
44. McPhail, C., Powers, W.T., Tucker, C.W.: Simulating individual and collective actions in temporary gatherings. *Social Sci Comput Rev* **10**(1), 1–28 (1992)
45. Still, G.K.: *Crowd Dynamics*. PhD thesis, Warwick University, 2000
46. Aubel, A., Boulic, R., Thalmann, D.: Real-time display of virtual humans: level of details and impostors. *IEEE Trans Circuits Syst Video Technol, Special Issue on 3D Video Technology* **10**(2), 207–227 (2000)
47. Anderson, M., McDaniel, E., Chenney, S.: Constrained animation of flocks. *Proc. ACM Symposium on Computer Animation* (2003)
48. Helbing, D., Farkas, I., Vicsek, T.: Simulating dynamical features of escape panic. *Nature* **407**, 487–490 (2000)
49. McPhail, C.: *The Myth of Maddening Crowd*. Aldine De Gruyter, New York (1991)
50. Ulicny, B., Thalmann, D.: Towards interactive real-time crowd behavior simulation. *Comput Graph Forum* **21**(4), 767–775 (2002)
51. Kalra, D., Barr, A.H.: Modeling with time and events in computer animation. *Proc.*

- Eurographics'92, pp. 45–58. Blackwell Publishers, Oxford (1992)
52. Rosenbloom, P.S., Laird, J.E., Newell, A.: *The Soar papers: research on artificial intelligence*. MIT Press, Cambridge (1993)
  53. Cremer, J., Kearney, J., Papelis, Y.: HCSM: framework for behavior and scenario control in virtual environments. *ACM Trans Modeling Comput Simul* **5**(3), 42–267 (1995)
  54. Boulic, R., Becheiraz, P., Emering, L., Thalmann, D.: Integration of motion control techniques for virtual human and avatar real-time animation. *Proc. VRST '97*, pp. 111–118. ACM Press, New York (1997)
  55. Ulicny, B., de Heras, P., Thalmann, D.: (2004) Crowdbrush: interactive authoring of real-time crowd scenes. *Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation'04*, pp. 243–252 (2004)
  56. Parke, F.I.: A parametric model for human faces. Ph.D. Thesis, University of Utah, UTEC-CSc-75-04 (1974)
  57. Ortony, A., Clore, G.L., Collins, A.: *The Cognitive Structure of Emotions*. Cambridge University Press (1988)
  58. Eysenck, H.J.: Biological dimensions of personality. In: Pervin, L.A. (ed.) *Handbook of Personality: Theory and Research*, pp. 244–276. Guilford Press, New York (1990)
  59. Costa, P.T., McCrae, R.R.: Normal personality assessment in clinical practice: the NEO personality inventory. *Psychol Assess* **(4)**, 5–13 (1992)
  60. Egges, A., Kshirsagar, S., Magnenat-Thalmann, N.: Generic personality and emotion simulation for conversational agents. *Comput Anim Virtual Worlds* **15**(1), 1–13 (2004)
  61. Marsella, S., Gratch, J.: A step towards irrationality: Using emotion to change belief. In: *Proceedings of the 1st International Joint Conference on Autonomous Agents and Multi-Agent Systems*, Bologna, Italy, July (2002)
  62. Johns, M., Silverman, B.G.: How emotions and personality effect the utility of alternative decisions: a terrorist target selection case study. In: *Tenth Conference On Computer Generated Forces and Behavioral Representation*, May (2001)
  63. André, E., Klesen, M., Gebhard, P., Allen, S., Rist, T.: Integrating models of personality and emotions into lifelike characters. In: *Proceedings International Workshop on Affect in Interactions. Towards a New Generation of Interfaces* (1999)
  64. Ball, G., Breese, J.: Emotion and personality in a conversational character. In: *Proceedings of the Workshop on Embodied Conversational Characters*, pp. 83–84, 119–121, October (1998)
  65. Chittaro, L., Serra, M.: Behavioural programming of autonomous characters based on probabilistic automata and personality. *Comput Anim Virtual Worlds* **15**(3–4), 319–326 (2004)
  66. El-Nasr, M., Ioerger, Z., Yen, J.: A pet with evolving emotional intelligence. In: *Proceedings of Autonomous Agents'99* (1999)
  67. LeDoux, J.E., Hirst, W.: *Mind and Brain: Dialogues in Cognitive Neuroscience*. Cambridge University Press (1986)
  68. Davidson, R.J., Ekman, P., Saron, C.D., Senulis, J.A., Friesen, W.V.: Approach withdrawal and cerebral asymmetry: emotional expression and brain physiology. *J Personal Soc Psychol* **58**, 330–341 (1990)
  69. Ekman, P.: Expression and the nature of emotion. In: Ekman, P., Scherer, K.R. (eds.) *Approaches to emotion*, pp. 319–343. Lawrence Erlbaum, Hillsdale, NJ (1984)
  70. Scherer, K.R.: Emotion expression in speech and music. In: Sundberg, J., Nord, L., Carlson, R. (eds.) *Music, Language, Speech, and Brain*, pp. 146–156. MacMillan, London (1991)
  71. Pelachaud, C., Bilvi, M.: Computational model of believable conversational agents. In: Huget, M.-P. (ed.) *Communication in MAS: Background, Current Trends and Future*, pp. 300–317. Springer, Berlin Heidelberg New York (2003)
  72. Smid, C., Pandzic, I., Radman, V.: Autonomous speaker agent. In *CASA 2004, 17th International Conference on Computer Animation and Agent*, pp. 259–266 (2004)
  73. Weil, J.: The synthesis of cloth objects. *Proc. SIGGRAPH'86, Computer Graphics*, vol. 24, pp. 243–252 (1986)
  74. Carignan, M., Yang, Y., Magnenat-Thalmann, N., Thalmann, D.: Dressing animated synthetic actors with complex deformable clothes. *Computer Graphics (SIGGRAPH'92 Proceedings)*, vol. 26(2), pp. 99–104. Addison-Wesley, Reading (1992)
  75. Volino, P., Courchesne, M., Magnenat-Thalmann, N.: Versatile and efficient techniques for simulating cloth and other deformable objects. *Computer Graphics (SIGGRAPH'95 Proceedings)*, pp. 137–144. Addison-Wesley, Reading (1995)
  76. Breen, D.E., House, D.H., Wozny, M.J.: Predicting the drape of woven cloth using interacting particles. *Computer Graphics (SIGGRAPH'94 Proceedings)*, pp. 365–372. Addison-Wesley, Reading (1994)
  77. Eischen, J.W., Deng, S., Clapp, T.G.: Finite-element modeling and control of flexible fabric parts, computer graphics in textiles and apparel. *IEEE Computer Graphics and Applications*, pp. 71–80 (1996)
  78. Barff, D., Witkin, A.: Large steps in cloth simulation. *ACM Trans Graph. Proceedings of ACM SIGGRAPH 1998, ACM SIGGRAPH*, pp. 43–54, Addison Wesley, Reading (1998)
  79. Eberhardt, B., Etmuss, O., Hauth, M.: Implicit-explicit schemes for fast animation with particles systems. *Proceedings of the Eurographics Workshop on Computer Animation and Simulation*, pp. 137–151. Springer, Berlin Heidelberg New York (2000)
  80. Volino, P., Magnenat-Thalmann, N.: Developing simulation techniques for an interactive clothing system. In: *Virtual Systems and Multimedia (VSMM'97 Proceedings)*, pp. 109–118. IEEE Press, Geneva, Switzerland (1997)
  81. Volino, P., Courchesne, M., Magnenat-Thalmann, N.: Versatile and efficient techniques for simulating cloth and other deformable objects. *Computer Graphics (SIGGRAPH'95 Proceedings)*, pp. 137–144. Addison-Wesley, Reading (1995)
  82. Eberhardt, B., Weber, A., Strasser, W.: A fast, flexible, particle-system model for cloth draping. *Computer Graphics in Textiles and Apparel (IEEE Computer Graphics and Applications)*, pp. 52–59, September (1996)
  83. Kang, Y.-M., Cho, H.-G.: Bilayered approximate integration for rapid and plausible animation of virtual cloth with realistic wrinkles. *Proceedings of Computer Animation 2002*, pp. 203. Geneva, Switzerland, IEEE Press, June (2002)
  84. Hadap, S., Bangarter, E., Volino, P., Magnenat-Thalmann, N.: Animating wrinkles on clothes. *IEEE Visualization '99*, pp. 175–182. IEEE Press, San Francisco (1999)
  85. Ng, H., Grimsdale, R.L.: GEOFF – a geometrical editor for fold formation. *Lecture Notes in Computer Science: Image Analysis Applications and Computer Graphics*, vol. 1024, pp. 124–131. Springer, Berlin Heidelberg New York (1995)
  86. Grzeszczuk, R., Terzopoulos, D., Hinton, G.: Neuroanimator: fast neural network emulation and control of physics-based models. *SIGGRAPH 98 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH*, pp. 9–20. Addison Wesley, Reading (1998)
  87. James, D.L., Fatahalian, K.: Precomputing interactive dynamic deformable scenes. *ACM Transactions on Graphics (TOG)*, *ACM SIGGRAPH*, vol. 22(3), pp. 165–172. Addison Wesley, Reading (2003)
  88. Cordier, F., Magnenat-Thalmann, N.: Real-time animation of dressed virtual human. *Computer Graphics Forum*, vol. 21(3), Blackwell Publishers, Oxford (2002)
  89. Magnenat-Thalmann, N., Seo, H., Cordier, F.: Automatic modeling of virtual humans and body clothing. *J Comput Sci Technol* **19**(6) (in press), Chinese Academy of Sciences, Beijing, China (publisher), December (2004)
  90. Papagiannakis, G., Schertenleib, S., O'Kennedy, B., Poizat, M., Magnenat-Thalmann, N., Stoddart, A., Thalmann, D.: Mixing virtual and real scenes in the site of ancient Pompeii. *J CAVW* **16**(1), 11–24 (2005)

91. Daldegan, A., Magnenat-Thalmann, N., Kurihara, T., Thalmann, D.: An integrated system for modeling, animating and rendering hair. *Computer Graphics Forum* (Eurographics '93), vol. 12(3), pp. 211–221 (1993)
92. Rosenblum, R., Carlson, W., Tripp, E.: Simulating the structure and dynamics of human hair: modeling, rendering and animation. *J Visual Comput Animat* **2**, 141–148 (1991)
93. Anjyo, K., Usami, Y., Kurihara, T.: A simple method for extracting the natural beauty of hair. *Computer Graphics* (Proceedings of SIGGRAPH 92), vol. 26(2), pp. 111–120, July (1992)
94. Lee, D.-W., Ko, H.S.: Natural hairstyle modeling and animation. In: *Proceedings of International Workshop on Human Modeling and Animation*, pp. 11–21, Seoul, Korea, June 2000. Korea Computer Graphics Society (2000)
95. Lin, M., Gottschalk, S.: Collision detection between geometric models: a survey. In: *Proceedings IMA Conf. on Mathematics of Surfaces*, pp.33–52 (1998)
96. Hadap, S., Magnenat-Thalmann, N.: Modeling dynamic hair as a continuum. *Computer Graphics Forum*, vol. 20(3), Eurographics Proceedings, Manchester, United Kingdom, September (2001)



NADIA MAGNENAT-THALMANN has pioneered research into virtual humans over the last 25 years. She obtained several Bachelor's and Master's degrees in various disciplines (Psychology, Biology and Chemistry) and a PhD in Quantum Physics from the University of Geneva. From 1977 to 1989, she was a Professor at the University of Montreal and led the research lab MIRALab in Canada. She moved to the University of Geneva in 1989, where she founded the Swiss MIRALab, an internationally interdisciplinary lab composed of about 30 researchers. She is author and coauthor of a very high number of research papers and books in the field of modeling virtual humans, interacting with them and in augmented life. She has received several scientific and artistic awards for her work and recently, in 1997, she has been elected to the Swiss Academy of Technical Sciences, and has been nominated as a Swiss personality who has contributed to the advance of science in the 150 years history produced by the Swiss Confederation Parliament.



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 Multimedia. He is coeditor-in-chief of the *Journal of Computer Animation and Virtual Worlds* and member of the editorial board of the *Visual Computer* and 4 other journals. Daniel Thalmann was member of numerous Program Committees, Co-chair, and Program Co-chair of several conferences including IEEE VR 2000. He has also organized 5 courses at SIGGRAPH on human animation and crowd simulation. Daniel Thalmann has published numerous papers in Graphics, Animation, and Virtual Reality. He is coeditor of 30 books included the recent "Handbook of Virtual Humans", published by John Wiley and Sons and coauthor of several books. He received his PhD 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.