Static and Dynamic Human Shape Modeling

Zhiqing Cheng[1] and Kathleen Robinette[2]

[1] Infoscitex Corporation, 4027 Colonel Glenn Highway, Suite 210, Dayton, OH 45431, USA
[2] 711th Human Performance Wing, Air Force Research Laboratory, 2800 Q Street, Wright-Patterson AFB, OH 45433, USA
{zhiqing.cheng,kathleen.robinette}@wpafb.af.mil

Abstract. Recent developments in static human shape modeling based on range scan data and dynamic human shape modeling from video imagery are reviewed. The topics discussed include shape description, surface registration, hole filling, shape characterization, and shape reconstruction for static modeling and pose identification, skeleton modeling, shape deformation, motion tracking, dynamic shape capture and reconstruction, and animation for dynamic modeling. A new method for human shape modeling is introduced.

Keywords: Human body, shape modeling, pose, animation.

1 Introduction

From the perspective of the motion status of the subject to be modeled, human shape modeling can be classified as either static or dynamic. Static shape modeling creates a model to describe human shape at a particular pose and addresses shape description, registration, hole filling, shape variation characterization, and shape reconstruction. Dynamic shape modeling addresses the shape variations due to pose changes (pose identification, skeleton modeling, and shape deformation) or while the subject is in motion (motion tracking, shape capture, shape reconstruction, and animation). Extensive investigations have been performed on human shape modeling [1-10]. Recent developments of human shape modeling, in particular, static shape modeling based on range scan data and dynamic shape modeling from video imagery, are reviewed in this paper. A new method for human shape modeling based on body segmentation and contour lines is introduced.

2 Static Shape Modeling

2.1 Shape Description

Shape description is a fundamental problem for human shape modeling. Traditional anthropometry is based on a set of measurements corresponding to linear distances between anatomical landmarks and circumference values at predefined locations. These measurements provide limited information about the human body shape [11]. With the advances in surface digitization technology, a 3-D surface scan of the whole body can be acquired in a few seconds. While whole body 3-D surface scan provides

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2009, LNCS 5620, pp. 3–12, 2009.

very detailed description of the body shape, the verbose scan data cannot be used directly for shape analysis. Therefore, it is necessary to convert 3-D scans to a form of compact representation. For searching and mining from a large 3-D scan database, Robinette [12] investigated 3-D shape descriptors where the Paquet Shape Descriptor (PSD) developed by Paquet and Rioux [13] was examined in detail. While PSD is able to discriminate or characterize different human shapes, it is not invertible. In other words, it is impossible to reconstruct a human shape from PSD.

An ideal human shape descriptor should be concise, unique, and complete for human shape description, efficient for shape indexing and searching, and invertible for shape reconstruction. Finding such a descriptor still remains a challenge. Alternatively, various graphic elements or graphic representation methods can be used to describe the human shape. For instance, Allen et al [2] and Anguelov et al [7] basically dealt directly with the vertices or polygons of a scanned surface for shape description. Allen et al [1] used subdivision surface in their pose modeling. Ben Azouz et al [6] utilized volumetric representation to convert vertices to voxels in their human shape modeling. While these methods guarantee reconstruction, they are not quite efficient for shape identification, discrimination, and searching.

2.2 Surface Registration

Surface registration or point-to-point correspondence among the scan data of different subjects is essential to many problems such as the study of human shape variability [2, 14] and pose modeling and animation [1, 7] where multiple subjects or multiple poses are involved. One methodology for establishing point-to-point correspondence among different scan data sets or models is usually called non-rigid registration. Given a set of markers between two meshes, non-rigid registration brings the meshes into close alignment while simultaneously aligning the markers.

Allen et al [1, 2] solved the correspondence problem between subjects by deforming a template model which is a hole-free, artist-generated mesh to fit individual scans. The resulting individually fitted scans or individual "models" all have the same number of triangles and point-to-point correspondences. The fitting process relies on a set of anthropometric landmarks provided in the CAESAR (Civilian American and European Surface Anthropometry Resource) database [15]. Anguelov et al [16] developed an unsupervised algorithm for registering 3-D surface scans of an object among different poses undergoing significant deformations. The algorithm called Correlated Correspondence (CC) does not use markers, nor does it assume prior knowledge about object shape, the dynamics of its deformation, or scan alignment. The algorithm registers two meshes with significant deformations by optimizing a joint probabilistic model over all point-to-point correspondences between them. This model enforces preservation of local mesh geometry as well as global constraints that capture the preservation of geodesic distance between corresponding point pairs. To obtain the markers for the non-rigid registration, Anguelov et al [7] then used the CC algorithm to compute the consistent embedding of each instance mesh (the mesh of a particular pose) into the template mesh (the mesh of a reference pose). Ben Azouz et al [14] used a volumetric representation of human 3-D surface to establish the correspondences between the scan data of different subjects. By converting their polygonal mesh descriptions to a volumetric representation, the 3D scans of different subjects are aligned inside a volume of fixed dimensions, which is sampled to a set of voxels. A human 3-D shape is then characterized by an array of signed distances between the voxels and their nearest point on the body surface. Correspondence is achieved by comparing for each voxel the signed distances attributed to different subjects without using anatomical landmarks.

2.3 Hole Filling

Surfaces acquired with scanners are typically incomplete and contain holes. Filling a hole is a challenging problem in its own right, as discussed by Davis et al. [18]. A common way to complete a hole is to fill it with a smooth surface patch that meets the boundary conditions of the hole. While these methods fill holes in a smooth manner, which is reasonable in some areas such as the top of the head and possibly in the underarm, other areas should not be filled smoothly. Therefore, Allen et al [2] developed a method that maps a surface from a template model to the hole area. Alternatively, hole-filling can be based on the contour lines of a scan surface [14].

2.4 Shape Variation Characterization

The human body comes in all shapes and sizes. Characterizing human shape variation is traditionally the subject of anthropometry—the study of human body measurement. The sparse measurements of traditional anthropometric shape characterization curtail its ability to capture the detailed shape variations needed for realism. While characterizing human shape variation based on a 3-D range scan could capture the details of shape variation, the method relies on three conditions: noise elimination, hole-filling and surface completion, and point-to-point correspondence. Also, whole body scanners generate verbose data that cannot be used directly for shape variation analysis. Therefore, it is necessary to convert 3-D scans to a compact representation that retains information of the body shape. Principal components analysis (PCA) is a potential solution to the problem. Allen et al [2] captured the variability of human shape by performing PCA over the displacements of the points from the template surface to an instance surface. Anguelov et al [7] also used PCA to characterize the shape deformation and then used the principal components for shape completion. Ben Azouz et al [14] applied PCA to the volumetric models where the vector is formed by the signed distance from a voxel to the surface of the model.

In order to explore the variations of the human body with intuitive control parameters (e.g., height, weight, age, and sex), Allen et al [2] showed how to relate several variables simultaneously by learning a linear mapping between the control parameters and the PCA weights. Ben Azouz et al [6,14,21] attempted to link the principal modes to some intuitive body shape variations by visualizing the first five modes of variation and gave interpretations of these modes. While PCA is shown to be effective in characterizing global shape variations, it may smear local variations for which other methods (e.g., wavelets) may be more effective.

2.5 Shape Reconstruction

Given a number of scan data sets of different subjects, a novel human shape can be created that has resemblance to the samples but is not the exact copy of any existing ones. This can be realized in four ways.

- *Interpolation or morphing.* One shape can be gradually morphed to another by interpolating between their vertices or other graphic entities [2]. In order to create a faithful intermediate shape between two individuals, it is critical that all features are well-aligned; otherwise, features will cross-fade instead of moving.
- *Reconstruction from eigen-space.* After PCA analysis, the features of sample shapes are characterized by eigen-vectors or eigen-persons which form an eigne-space. Any new shape model can be generated from this space by combining a number of eigen-models with appropriate weighting factors [14].
- *Feature-based synthesis.* Once the relationship between human anthropometric features and eigen-vectors is established, a new shape model can be constructed from the eigen-space with desired features by editing multiple correlated attributes, such as height and weight [2] or fat percentage and hip-to-waist ratio [4].
- *Marker-only matching.* Marker-only matching can be considered as a way of reconstruction with provided markers [2]. This is important for many applications such as deriving a model from video imagery, since marker data can be obtained using less expensive equipment than a laser range scanner.

3 Pose Change Modeling

During pose changing or body movement, muscles, bones, and other anatomical structures continuously shift and change the shape of the body. For pose modeling, scanning the subject in every pose is impractical; instead, body shape can be scanned in a set of key poses, and then the body shapes corresponding to intermediate poses are determined by smoothly interpolating among these poses. The issues involved in pose modeling include pose definition and identification, skeleton model derivation, shape deformation (skinning), and pose mapping.

3.1 Pose Definition and Identification

The human body can assume various poses. In order to have a common basis for pose modeling, a distinct, unique description of difference poses is required. Since it is impossible to collect the data or create template models for all possible poses, it is necessary to define a set of standard, typical poses. This is pose definition. A convention for pose definition is yet to be established. One approach is to use joint angle changes as the measures to characterize human pose changing and gross motion. This means that poses can be defined by joint angles. By defining poses and motion in such a way, the body shape variations caused by pose changing and motion will consist of both rigid and non-rigid deformation. Rigid deformation is associated with the orientation and position of segments that connect joints. Non-rigid deformation is related to the changes in shape of soft tissues associated with segments in motion, which, however, excludes local deformation caused by muscle action alone. One method for measuring and defining joint angles is using a skeleton model. In the model, the human body is divided into multiple segments according to major joints of the body, each segment is represented by a rigid linkage, and an appropriate joint is placed between the two corresponding linkages. Given a set of scan data, imagery, or photos, the determination or identification of the corresponding pose can be done by fitting a skeleton model to the data set. The skeleton model derivation will be discussed in the following section. Alternatively, there are several methods for pose identification that are not based on skeleton models. Mittal et al [22] studied human body pose estimation using silhouette shape analysis. Cohen and Li [23] proposed an approach for inferring the body posture using a 3-D visual-hull constructed from a set of silhouettes.

3.2 Skeleton Model

Allen et al [1] constructed a kinematic skeleton model to identify the pose of a scan data set using markers captured during range scanning. Anguelov et al [24] developed an algorithm that automatically recovers from 3-D range data a decomposition of the object into approximately rigid parts, the location of the parts in the different poses, and the articulated object skeleton linking the parts. Robertson and Trucco [25] developed an evolutionary approach to estimating upper-body posture from multi-view markerless sequences. Sundaresan et al [26] proposed a general approach that uses Laplacian eigen-maps and a graphical model of the human body to segment 3-D voxel data of humans into different articulated chains.

3.3 Body Deformation Modeling

Body deformation modeling is also referred to as skinning in animation. Two main approaches for modeling body deformations are anatomical modeling and example-based modeling. The anatomical modeling is based on an accurate representation of the major bones, muscles, and other interior structures of the body [27]. These structures are deformed as necessary when the body moves, and a skin model is wrapped around the underlying anatomy to obtain the final geometry of the body shape. The finite element method is the primary modeling technique used for anatomical modeling. In the example-based approach, a model of some body part in several different poses with the same underlying mesh structure can be generated by an artist. These poses are correlated to various degrees of freedom, such as joint angles. An animator can then supply values for the degrees of freedom of a new pose and the body shape for that new pose is interpolated appropriately. Lewis et al [28] and Sloan et al [29] developed similar techniques for applying example-based approaches to meshes. Instead of using artist-generated models, recent work on the example-based modeling uses range-scan data.

Allen et al [1] presented an example-based method for calculating skeleton-driven body deformations. Their example data consists of range scans of a human body in a variety of poses. Using markers captured during range scanning, a kinematic skeleton is constructed first to identify the pose of each scan. Then a mutually consistent parameterization of all the scans is constructed using a posable subdivision surface template. Anguelov et al [7] developed a method that incorporates both articulated and non-rigid deformations. A pose deformation model was constructed from training scan data that derives the non-rigid surface deformation as a function of the pose of the articulated skeleton. A separate model of shape variation was derived from the training data also. The two models were combined to produce a 3D surface model with realistic muscle deformation for different people in different poses. The method (model) is referred to as the SCAPE (Shape Completion and Animation for People).

For pose modeling, it is impossible to acquire the pose deformation for each person at each pose. Instead, pose deformation can be transferred from one person to another for a given pose. Anguelov et al [7] addressed this issue by integrating a pose model with a shape model reconstructed from eigen-space. As such, they were able to generate a mesh for any body shape in their PCA space in any pose.

4 Shape Modeling of Human in Motion

4.1 Motion Tracking

Human motion tracking or capturing is an area that has attracted a lot of study and investigation. Today's performance of off-the-shelf computer hardware enables marker-free, non-intrusive optical tracking of the human body. For example, Theobalt et al [30] developed a system to capture human motion at interactive frame rates without the use of markers or scene-introducing devices. The algorithms for 2-D computer vision and 3-D volumetric scene reconstruction were applied directly to the image data. A person was recorded by multiple synchronized cameras, and a multi-layer hierarchical kinematic skeleton was fitted into each frame in a two-stage process.

4.2 Dynamic Shape Capture

During dynamic activities, the surface of the human body moves in many subtle but visually significant ways: bending, bulging, jiggling, and stretching. Park and Hodgins [8] developed a technique for capturing and animating those motions using a commercial motion capture system with approximately 350 markers. Supplemented with a detailed, actor specific surface model, the motion of the skin was then computed by segmenting the markers into the motion of a set of rigid parts and a residual deformation. Sand et al [5] developed a method (a needle model) for the acquisition of deformable human geometry from silhouettes. Their technique uses a commercial tracking system to determine the motion of the skeleton and then estimates geometry for each bone using constraints provided by the silhouettes from one or more cameras.

4.3 Shape Reconstruction from Imagery Data

- *From Photos.* Seo et al [31] presented a data-driven shape model for reconstructing human body models from one or more 2-D photos. A data-driven, parameterized deformable model acquired from a collection of range scans of a real human body is used to complement the image-based reconstruction by leveraging the quality, shape, and statistical information accumulated from multiple shapes of range-scanned people.

- *From Video Sequences.* One recent work was done by Balan et al [10] that proposed a method for recovering human shape models directly from images. Specifically, the human body shape is represented by the SCAPE [7] and the parameters of the model are directly estimated from image data. A cost function between image observations and a hypothesized mesh is defined and the problem is formulated as an optimization.

4.4 Animation

The animation of the subject can be realized by displaying a series of human shape models for a prescribed sequence of poses. Hilton et al [3] built a framework for construction of animated models from the captured surface shape of real objects. Seo et al [4] developed a synthesizer where for any synthesized model, the underlying bone and skin structure is properly adjusted, so that the model remains completely animatable using the underlying skeleton. Aguiar et al [9] developed a novel versatile, fast and simple framework to generate high quality animations of scanned human characters from input motion data. The method is purely mesh-based and can easily transfer motions between human subjects of completely different shape and proportions.

5 A New Method

In the static human shape modeling based on 3-D laser scan data, polygons/vertices are usually used as the basic graphic entities for the representation of a human body shape. Usually approximately 20,000 ~ 500,000 vertices are required to describe a full body shape, depending upon surface resolution. This way of surface representation incurs a large computational cost and cannot ensure point-to-point correspondence among the scans of different subjects. Thus we developed a new method that uses contour lines as the basic entities for the shape modeling. The entire procedure of the method is as follows. (1) *Joint center calculation* The human body is treated as a multi-segment system where segments are connected to each other by joints, which in turn, are defined by respective landmarks. (2) *Skeleton model building* A skeleton model is formed by connecting respective joint centers to represent the articulated structure and segments of the human body, as shown in Fig. 1. (3) *Segmentation* The entire body scan is divided into segments according to the skeleton model with some special treatment in certain body areas. (4) *Slicing* The scan of each segment is sliced along the main axis of each segment at fixed intervals, which produces the contour lines of the segment. Figure 2 displays the segmentation and slicing of a whole body scan. (5) *Discretizing* Each contour line is discretized with respect to a polar angle. As such, the two-dimensional contour curve is represented by a vector. (6) *Hole-filling* The hole-filling is performed on contour lines for each segment. Figure 3 shows the original surface and filled surface of the abdomen segment. (7) *Parameterization* The vector of each discretized contour line is represented by a set of wavelet coefficients. (8) *Registration* The point-to-point correspondence between the scans of two bodies is established with respect to the contour lines of each segment. (9) *Shape description and PCA* The assembly of the wavelet coefficients of all segments is used as the shape description vector. Principal component analysis (PCA) is performed on a selection of subjects from the CAESAR database. (10) *Shape reconstruction* A 3D human shape model is reconstructed in the following way: (a) From the shape description vector to wavelet coefficients; (b) From wavelet coefficients to contour lines; (c) From contour lines to 3D surface model; and (d) Part blending as needed.

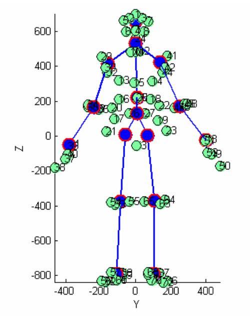

Fig. 1. Landmarks, joint centers, and the skeleton model

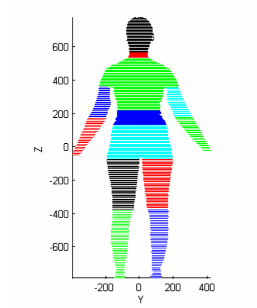

Fig. 2. Segmentation and slicing

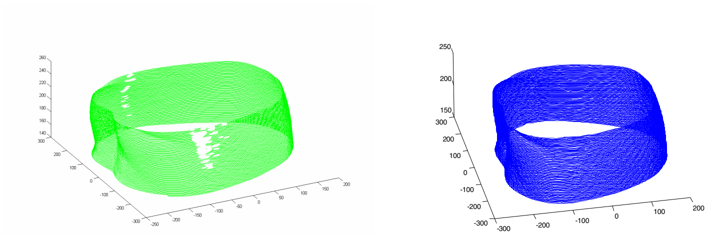

Fig. 3. Hole filling based on contour lines

6 Concluding Remarks

Human shape modeling spans various research areas from anthropometry, computer graphics and computer vision to machine intelligence and optimization. It simply would not be possible to present a full survey of the related work. Instead, this paper just intended to provide an indication of the current state of the art. In addition to traditional uses, human modeling is finding many new applications with great challenges, such as virtual environment, human identification, and human-borne threat detection.

References

1. Allen, B., Curless, B., Popovic, Z.: Articulated Body Deformation from Range Scan Data. In: ACM SIGGRAPH 2002, San Antonio, TX, USA, pp. 21–26 (2002)
2. Allen, B., Curless, B., Popvic, Z.: The space of human body shapes: reconstruction and parameterization from range scans. In: ACM SIGGRAPH 2003, San Diego, CA, USA, 27-31 July (2003)

3. Hilton, A., Starck, J., Collins, G.: From 3D Shape Capture to Animated Models. In: Proceedings of First International Symposium on 3D Data Processing Visualization and Transmission, pp. 246–255 (2002)
4. Seo, H., Cordier, F., Thalmann, N.M.: Synthesizing Animatable Body Models with Parameterized Shape Modifications. In: Eurographics/SIGGRAPH Symposium on Computer Animation (2003)
5. Sand, P., McMillan, L., Popovic, J.: Continuous Capture of Skin Deformation. ACM Transactions on Graphics 22(3), 578–586 (2003)
6. Ben Azouz, Z., Rioux, M., Shu, C., Lepage, R.: Analysis of Human Shape Variation using Volumetric Techniques. In: Proc. of 17th Annual Conference on Computer Animation and Social Agents Geneva, Switzerland (2004)
7. Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J.: SCAPE: Shape Completion and Animation of People. ACM Transactions on Graphics 24(3) (2005)
8. Park, S.I., Hodgins, J.K.: Capturing and Animating Skin Deformation in Human Motion. In: ACM Transaction on Graphics (SIGGRAPH 2006), vol. 25(3), pp. 881–889 (2006)
9. Aguiar, E., Zayer, R., Theobalt, C., Magnor, M., Seidel, H.P.: A Framework for Natural Animation of Digitized Models. MPI-I-2006-4-003 (2006)
10. Balan, A., Sigal, L., Black, M., Davis, J., and Haussecker, H.: Detailed Human Shape and Pose from Images. In: IEEE Conf. on Comp. Vision and Pattern Recognition (2007).
11. Robinette, K.M., Vannier, M.W., Rioux, M., Jones, P.: 3-D surface anthropometry: Review of technologies. In: North Atlantic Treaty Organization Advisory Group for Aerospace Rearch & Development, Aerospace Medical Panel (1997)
12. Robinette, K.M.: An Investigation of 3-D Anthropometric Shape Descriptors for Database Mining, Ph.D. Thesis, University of Cincinnati (2003)
13. Paquet, E., Rioux, M.: Content-based access of VRML libraries. In: Ip, H.H.-S., Smeulders, A.W.M. (eds.) MINAR 1998. LNCS, vol. 1464. Springer, Heidelberg (1998)
14. Ben Azouz, Z.B., Shu, C., Lepage, R., Rioux, M.: Extracting Main Modes of Human Body Shape Variation from 3-D Anthropometric Data. In: Proceedings of the Fifth International Conference on 3-D Digital Imaging and Modeling (2005)
15. Robinette, K., Daanen, H., Paquet, E.: The Caesar Project: A 3-D Surface Anthropometry Survey. In: Second International Conference on 3-D Digital Imaging and Modeling (3DIM 1999), Ottawa, Canada, pp. 380–386 (1999)
16. Anguelov, D., Srinivasan, P., Pang, H.C., Koller, D., Thrun, S., Davis, J.: The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. Advances in Neural Information Processing Systems 17, 33–40 (2005)
17. Ben Azou, Z.B., Shu, C., Mantel, A.: Automatic Locating of Anthropometrics Landmarks on 3D Human models. In: Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (2006)
18. Davis, J., Marschner, S., Garr, M., Levoy, M.: Filling holes in complex surfaces using volumetric diffusion. In: Proceedings of the First International Symposium on 3D Data Processing, Visualization and Transmission. Padua, Italy (2002)
19. Curless, B., Levoy, M.: A volumetric method of building complex models from range images. In: Proceedings of SIGGRAPH 1996, 303–312 (1996)
20. Liepa, P.: Filling holes in meshes. In: Proc. of the Eurographics/ACM SIGGRAPH symposium on Geometry processing, pp. 200–205 (2003)
21. Ben Azouz, B.Z., Rioux, M., Shu, C., Lepage, R.: Characterizing Human Shape Variation Using 3-D Anthropometric Data. International Journal of Computer Graphics 22(5), 302–314 (2005)

22. Mittal, A., Zhao, L., Davis, L.S.: Human Body Pose Estimation Using Silhouette Shape Analysis. In: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (2003)
23. Cohen, I., Li, H.X.: Inference of Human Postures by Classification of 3D Human Body Shape. In: Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (2003)
24. Anguelov, D., Koller, D., Pang, H.C., Srinivasan, P., Thrun, S.: Recovering Articulated Object Models from 3D Range Data. In: Proceedings of the 20th conference on Uncertainty in artificial intelligence, pp. 18–26 (2004)
25. Robertson, C., Trucco, E.: Human body posture via hierarchical evolutionary optimization. In: BMVC 2006 (2006)
26. Sundaresan, A., Chellappa, R.: Model driven segmentation of articulating humans in Laplacian Eigenspace. IEEE Transaction: Patter Analysis and Machine Intelligence (2007)
27. Aubel, A., Thalmann, D.: Interactive modeling of the human musculature. In: Proc. of Computer Animation (2001)
28. Lewis, J.P., Cordner, M., Fong, N.: Pose space deformations: A unified approach to shape interpolation and skeleton-driven deformation. In: ACM SIGGRAPH, pp. 165–172 (2000)
29. Sloan, P.P., Rose, C., Cohen, M.F.: Shape by example. In: Proceedings of 2001 Symposium on Interactive 3D Graphics (2001)
30. Theobalt, C., Magnor, M., Schüler, P., Seidel, H.P.: Combining 2D Feature Tracking and Volume Reconstructions for Online Video-Based Human Motion Capture. International Journal of Image and Graphics 4(4), 563–583 (2004)
31. Seo, H., Yeo, Y.I., Wohn, K.: 3D Body Reconstruction from Photos Based on Range Scan. In: Pan, Z., Aylett, R.S., Diener, H., Jin, X., Göbel, S., Li, L. (eds.) Edutainment 2006. LNCS, vol. 3942, pp. 849–860. Springer, Heidelberg (2006)