COMPUTER GRAPHICS I
COMP.5460

LITERATURE REVIEW III

PRANITHA KEERTHI
01719717

SCAPE: SHAPE COMPLETION AND ANIMATION OF PEOPLE

This paper introduces the SCAPE method (Shape Completion and Animation for People) — a data-driven method for building a unified model of human shape. Graphics applications often require a complete surface model for rendering and animation. Obtaining a complete model of a particular person is often difficult or impossible. When the task is to obtain a 3D sequence of the person in motion, the situation can be even more difficult.

This method learns separate models of body deformation — one accounting for changes in pose and one accounting for differences in body shape between humans. The models provide a level of detail sufficient to produce dense full body meshes, and capture details such as muscle deformations of the body in different poses. A key aspect of this pose model is that it decouples deformation into a rigid and a non-rigid component. The pose deformation component of this model is acquired from a set of dense 3D scans of a single person in multiple poses. The pose deformation is relatively low dimensional, allowing the shape deformation to be learned automatically, from limited training data.

This model component can be acquired from a set of 3D scans of different people in different poses. The model of shape variation does not get confounded by deformations due to pose, as those are accounted for separately. The two parts of the model form a single unified framework for shape variability of people. The framework can be used to generate a complete surface mesh given only a succinct specification of the desired shape — the angles of the human skeleton and the eigen-coefficients describing the body shape.

The body-shape variation is modeled independently of the pose variation, by introducing a new set of linear transformation matrices S i k , one for each instance i and each triangle k. This method was used to learn a SCAPE body shape deformation model using the 45 instances in the body shape data set, and taking as a starting point the pose deformation model , the mean shape and the first four principal components in our PCA decomposition of the shape space. These components represent very reasonable variations in weight and height, gender, abdominal fat and chest muscles, and bulkiness of the chest versus the hips.

There are two important graphic tasks that are highlighted in this work. The first is partial view completion. Most scanned surface models of humans 408 have significant missing regions. Given a partial mesh of a person for whom there is no previous data, this method finds the shape that best fits the observed partial data in the space of human shapes. The model can then be used to predict a full 3D mesh. The second task is producing a full 3D animation of a moving person from marker motion capture data. This problem is approached as a shape completion task.

The author then describes the basic pipeline for data acquisition and pre-processing of the data meshes. The specific design of the pipeline is inessential for the main contribution of this paper;

they describe it first to introduce the type of data used for learning this model. The authors acquired their surface data using a Cyberware WBX whole-body scanner. The scanner captures range scans from four directions simultaneously and the models contain about 200K points. The next step in the data acquisition pipeline brings the template mesh into correspondence with each of the other mesh instances. Given a set of markers between two meshes, the task of non-rigid registration is well understood and a variety of algorithms exist. The task is to bring the meshes into close alignment, while simultaneously aligning the markers. A skeleton is then constructed for the template mesh automatically, using only the meshes in their data set.

Next, how to use the SCAPE model to address the task of shape completion is discussed, which is the main focus of this work. A solution to this optimization problem is a completed mesh Y[Z] that both fits the observed marker locations and is consistent with the predictions of the learned SCAPE model. An obvious application of the shape completion method is to the task of partial view completion. Here, a partial scan of a human body is given; the task is to produce a full 3D mesh which is consistent with the observed partial scan, and provides a realistic completion for the unseen parts.

The shape completion framework can also be applied to produce animations from marker motion capture sequences. In this case, there exists a sequence of frames, each specifying the 3D positions for some set of markers. Currently, the framework includes no prior over poses. Thus, when encountering occlusions, this work cannot use the observed position of some body parts to constrain the likely location of others. This model can easily be extended to encompass such a prior, in a modular way.

Finally, it can be concluded that this approach is purely data driven, generating the entire model from a set of data scans. Human intervention is required only for placing a small set of markers on the scans, as a starting point for registration. Thus, the model can easily be applied to other data sets, allowing the authors to generate models specific to certain types of body shapes or certain poses.

EMBODIED HANDS: MODELING AND CAPTURING HAANDS AND BODIES TOGETHER

Humans move their hands and bodies together to communicate and solve tasks. Capturing and replicating such coordinated activity is critical for virtual characters that behave realistically. Here the authors formulate a model of hands and bodies interacting together and fit it to full-body 4D sequences. Several factors have led to the separation of hands from bodies. Full body models such as SCAPE [Anguelov et al. 2005] are learned from subjects making a tight fist, while more recent models like SMPL [Loper et al. 2015] assume an open, rigid, hand. Neither looks realistic when animated. Such 3D body models are created from 3D scans of the whole body.

This model makes several contributions, which can be roughly divided into two categories: learning a new model of the hand, and tracking hands and bodies together. First a new database of detailed hand scans of 31 subjects in up to 51 poses is collected. Both left and right hands of men and women are captured, using a wide range of poses, and capture hands interacting with objects. Second, this data is used to build a statistical hand model similar to the SMPL body model [Loper et al. 2015]. Like SMPL, the model factors geometric changes into those inherent to the identity of the subject and those caused by pose. Third, the MANO is combined with the SMPL body model to give a new combined model of hands and bodies interacting together. Fourth, the problem of capturing full bodies and hands in motion is addressed.

Given the difficulty of capturing hands with bodies, the model advocate for a two-stage approach in the creation of a dexterous full body model. First, the model collects a large number of scans of hands in isolation. These scans are obtained with a scanner configured specifically to capture hands with a fixed wrist position. This allows to capture the nuances of hand deformation. Then a hand model is trained using an iterative process of aligning a template to the scans using the model and learning a model from the registered scans. In the second stage, this hand model is integrated with the full body to obtain a single, dexterous and fully articulated body model.

The next step to building a hand model is to register, or align, a template to all the hand scans, bringing them into correspondence. The process of registering noisy hand scan data is challenging. Mesh registration is a challenging problem in general, but hands are especially difficult due to a high degree of self-similar structures (fingers) and a high level of self and object occlusion. While some previous work uses landmarks to simplify registration, this work do not do so because it is impractical when registering a large number of scans. Instead, the creation of the model is bootstrapped by manually curating the registrations. Specifically, this model starts with a crude hand model, use it to register all our scans, manually curate good ones, learn an improved model, and repeat. This gradually improves the alignments and the model.

This approach is relevant for industrial applications in several different ways. First, while this capture environment is complex, the ability to capture fully body and hand performance without markers may make this appropriate for high end applications. Second, the body model with hand articulation is directly usable by animators as it is designed to be compatible with existing animation systems. It is believed that this is the first learned model of bodies and hands for which this is true. Third, this provides a practical step towards the simultaneous capture of hands and bodies, which will enable animators to learn how body pose and hand pose are correlated; this is useful both as reference material and for training machine learning models relating the two.

Finally it is concluded that the authors propose MANO, a new model of the human hand that is learned from examples, is low-dimensional, is easy to pose and fit to data, and is compatible with graphic engines because it is built on linear blend skinning with blend shapes.

SIMILARITY BETWEEN THE TWO PAPERS

Both the papers deals with the graphical representation of humans. Humans move their hands and bodies together to communicate and solve tasks. Capturing and replicating such coordinated activity is critical for virtual characters that behave realistically. The primary paper mainly concentrates on the movement of hands and bodies and how they coordinate and the secondary paper concentrates on animation for the shape of human. Full body models such as the secondary paper is learned from subjects making a tight fist whereas the primary paper roughly divides the representation into two categories: learning a new model of the hand, and tracking hands and bodies together. The primary paper proposes a new model of hand shape and pose that is learned from data, compatible with existing graphics systems, lowdimensional, realistic and compatible with the SMPL body model. The secondary paper introduces a data-driven method for building a human shape model that spans variation in both subject shape and pose.

REFERENCES

PRIMARY PAPER:

EMBODIED HANDS: MODELING AND CAPTURING HAANDS AND BODIES TOGETHER

ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH Asia 2017

Volume 36 Issue 6, November 2017
Article No. 245

SECONDARY PAPER:

SCAPE: SHAPE COMPLETION AND ANIMATION OF PEOPLE

ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2005

Volume 24 Issue 3, July 2005