Stretchable Cartoon Editing for Skeletal Captured Animations

Yann Savoye*
INRIA Bordeaux University, France.

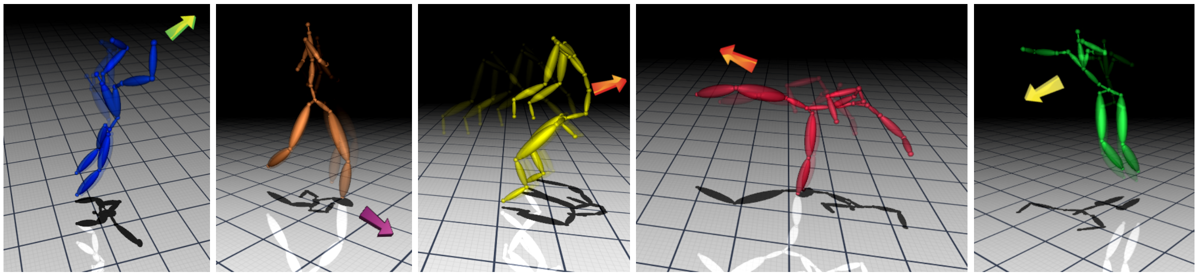

Figure 1: ***Laplacian Cartoon Skeleton Editing**: Optimizing traditional skeletal animated structure with differential-aware bone stretching effects allows animators to efficiently re-use previously captured MoCap data in the context of cartoon productions. We increase Squash-and-Stretch effect on existing realistic* BASKETBALL DUNK, DANCING, HORSE RUADE, KARATE KICK-OFF *and* JUMPING *motion clips (from left to right hand side). Original and intermediate edited poses are displayed in transparency and resulting edited poses in superposition.*

1 Introduction

In this work, we describe a new and simple approach to re-use skeleton of animation with joint-based Laplacian-type regularization, in the context of exaggerated skeleton-based character animation. Despite decades of research, interactive character animations still have a lack of flexibility and editability in order to re-use real vertebral motion. In further details, generation of expressive cartoon animation from real data, is a challenging key task for non-photorealistic animation. Hence, a major problem for artists in production is to enhance the expressiveness of classical motion clips by direct manipulation of the underlying skeletal structure. Relatively small number of researchers present their approach for processing cartoon effects on motion data in [Kwon and Lee 2007; Davis and Kannappan 2002; Bregler et al. 2002]. However existing techniques often avoid dealing with the potential of skeletal-based optimization, while preserving the joint coherence and connectivity. Besides, the majority of characters in cartoons have the flexibility to stretch to extreme positions and squash to astounding shapes. It can also be noticed that squash-and-stretch is easier to realize in traditional animation rather than mocap-based computer generated animation. For this reason, *Ratatouille* a Pixar® movie, did not use a rigid skeleton, while abandoning motion capture to reach such essential non-ultra realistic appeal.

2 Skeletal Graph Laplacian

The idea of the motion capture is to use sensor placed on the subject, and to collect data that describes their motion while they are performing motions. The pose of an articulated figure can be specified by its joint configuration in addition to the position and root orientation segment. Skeleton of animation $\mathcal{S} = (\mathcal{J}, \mathcal{B}, \mathcal{M})$ is composed of a hierarchy of joints enriched with motion data, organized in frames. $\mathcal{S}$ is made of a set $\mathcal{J}$ of n joints, a set $\mathcal{B}$ of bones connecting joints, and a set $\mathcal{M}$ of k motion frames. Two joints i and j are connected into a unique bone only if $(i, j) \in \mathcal{B}$. Motion data consist of bundle of signal defined as a continuous function $\mathbf{f}(t)$. Compiled global rigid transformation matrix from captured motion data associated to the i^{th} joint at the frame t is denoted by $\mathbf{M}_i^t \in \mathbb{R}^{4\times4}$. The global location $\mathbf{p}_i^t$ of the i^{th} joint at the frame t is then the homogeneous zero transformed by the sequence of transformation and can be written as: $\mathbf{p}_i^t = \mathbf{M}_i^t \cdot \left[\vec{0}\,|1\right]^T$. We define the *Differential Joint Coordinates* δ_i^t of joint i at frame t as follows:

$$\delta_i^t = \mathbf{p}_i^t - \sum_{j\in N(i)} \frac{1}{d_i}\left(\left(\mathbf{M}_i^t - \mathbf{M}_j^t\right)\cdot\left[\vec{0}\,|1\right]^T\right)$$

The degree of the joint i denoted by d_i is equal to the number of joints linked to i. The set of immediate adjacent joints to i is denoted by $N(i) = \{j|\,(i, j) \in \mathcal{B}\}$. In addition, the skeleton topology is represented as an open directed acyclic graph. So, we introduce the *Skeletal Graph Laplacian* L_S by $L_S = D - A$ where D is the diagonal matrix of joint degrees and A is the adjacency joint matrix. We denote by $\mathcal{L}_S(.)$ the per-joint Uniform Laplacian Operator applied on the skeleton graph structure. The entries of the corresponding square symmetric $n \times n$ matrix L_S are setup as follows:

$$L_S[i,j] = \begin{cases} 1 & \text{if } i = j \\ -1/d_i & \text{if } (i,j) \in \mathcal{B} \\ 0 & \text{otherwise} \end{cases}$$

We focus on adding non-rigid effects to an existing captured skeletal structure, while preserving its consistency and original connectivity. In our approach, we are referring to the well-known Laplacian shrinking effects (i.e. shearing and stretching distortion) in order to apply non-rigid warps over the rigid skeleton topology.

3 Cartoon Optimization Algorithm

Our algorithm takes an arbitrary input articulated motion signal and perturbates its Euler and Euclidean representations in such a way that the output motion looks more cartoon-like. In order to hack the rigidity, we prefer to deal with the skeletal structure as euclidean joint coordinates. Our technique has two key components: spatio-temporal motion filtering and global joint location optimization. At the beginning of pre-optimization, we apply the cartoon animation filter as suggested in the [Wang et al. 2006] in order to add: the follow-through, exaggeration and anticipation effects on the motion signal. The filtered motion signal $\tilde{\mathbf{f}}(t)$ has interesting properties, especially on kinematic chains that are not explicitly edited automatically or hand-drawn specified. This filter involving a Laplacian of Gaussian (*LoG*), is defined as follows:

$$\tilde{\mathbf{f}}(t) = \mathbf{f}(t) - \mathbf{f}(t) \otimes LoG$$

To continue, we reformulate the *Squash-and-Stretch* problem as a skeletal adaptation optimization. Given the fact it is nearly impossible to get a plausible squash-and-stretch by working exclusively in Euler space with inverse kinematics, we prefer to optimize the whole skeletal structure in term of global joint location of a pre-animated pose satisfying stretching constraints. The core algorithm of our technique relies on *Skeletal Graph Laplacian*, with the aim to ensure spatial relationship of joints under sparse differential-aware stretching features over the whole joint hierarchy. As a result, the driving idea is to employ fast and accurate skeleton fitting process, guided by kinematic-free constraints in each frame.

*ysavoye@siggraph.org

At each frame, the initial solution is the Euclidean parameters of the animated pose, generated by forward kinematics. Thus, pose estimation recovers the pose by minimizing an objective function that is a combination of penalty and data terms. Smoothness term is required for making the warp field regular. To control a cartoon desired pose, the user inputs the targeted global joint positions $\mathbf{q}_l^t$ for a collection $\mathcal{C}$ of edited joints. We conserve the Laplacian coordinates δ_i^t for each joint i in the skeleton hierarchy. The reconstructed positions $\widehat{\mathbf{p}}$ of the skeleton joints in world space coordinates are obtained by solving the following quadratic minimization problem separately from each euclidean dimension:

$$\underset{\widehat{\mathbf{p}}^t}{\operatorname{argmin}} \left(\sum_{l \in \mathcal{C}} \left\| \mathbf{q}_l^t - \widehat{\mathbf{p}}_l^t \right\|^2 + \sum_{i=1}^{n} \left\| \mathcal{L}_{\mathcal{S}} \left(\widehat{\mathbf{p}}_i^t \right) - \delta_i^t \right\|^2 \right)$$

Global joint location is estimated by minimizing the sum of squared difference between the data-driven pre-animated pose and input features cues. In order to avoid purely translation effect provoked by a single dragged-and-dropped joint, we fix by default the root joint, acting as skeleton gravity center. The root joint is generally located in a place where many bones come together. In our framework, bone elongation at frame t can also be automatically established along the bone direction for selected joint l having the k^{th} joint as parent with a time-dependent scaling factor α_t:

$$\mathbf{q}_l^t = \mathbf{p}_l^t \pm \alpha_t \cdot \frac{\left(\mathbf{p}_l^t - \mathbf{p}_k^t\right)}{\left\|\mathbf{p}_l^t - \mathbf{p}_k^t\right\|}$$

Energy terms can be equilibrated by a weighting system. Using such refined constraint formulation united with Laplacian on the bone structure is motivated by the fact that *Squash-and-Stretch* can be accomplished by differential scaling in Euclidean coordinates system. Hence, this minimization problem can be solved efficiently in least-squares sense based on the succeeding expression:

$$\mathbf{A} \cdot \mathbf{X'} = \mathbf{B}$$

Moreover, the global location of joints can be therefore found by solving in real-time the very small sparse linear system using this closed form expression:

$$\mathbf{X'} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{B}$$

This sparse editing technique performs spatio-temporal numerical optimization on MoCap data. As post-optimization step, we enforce temporal smoothness by penalizing inter-keyframe deviation. Our edited motion signal is consequently simply filtered using iterative smoothing convolution. A better solution is to diffuse edited exaggeration across time over the joint trajectories using temporal Laplacian solver with Dirichlet boundary conditions.

4 Results and Discussion

To demonstrate the robustness and usefulness of our techniques, we have implemented this advanced reproducible pipeline using only C++ and OPENGL. As shown on the accompanying video (produced real-time and performed by a non-professional artist), just a single editing feature is sufficient to adapt the whole skeletal pose in coherent manner, as shown in Figure 2. Moreover, results have shown that the proposed technique is robust enough to probably be used in low-budget productions, especially for processing sporty motion clips, as illustrated in Figure 1. We have tested the robustness of our method on a corpus composed of more than forty motion clips with success. In addition, we developed a simple interaction technique coupled with depth first search to handle the closest joint.

Even if human perception system tends to focus on kinematic parameters rather than on structural clues, real-life creatures actually elongate and stretch due to the elasticity of tendons and muscles. Consequently, breaking the rigidity of the underlying armature adds a pleasant realism to cartoon rubber-like effects. The resulting animated structure looks more pleasant while remains rigid in motion. Our technique can be widely used in more complete 2D or 3D animation systems to easily achieve non-rigid, rigged-and-skinned animation, cartoon shape control and elastic skin deformation from MoCap data. Our method offers the potential ability to control stretchable effect while allowing smooth behavior of skeletally-guided skin deformation, as far as the rigging function is smooth.

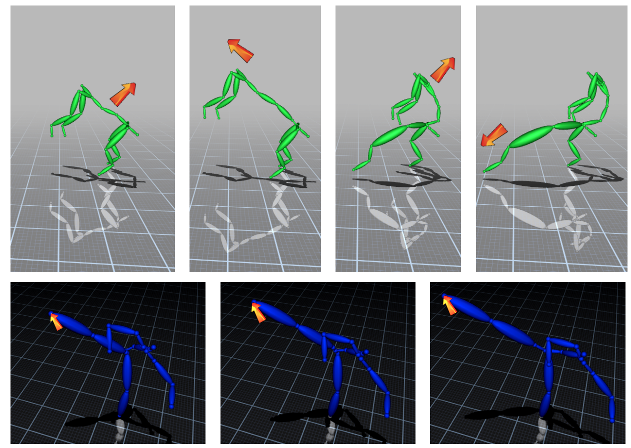

Figure 2: *A single editing feature is sufficient to produce pleasant cartoon gait style over the whole skeletal pose in coherent manner for a horse ruade (**top row**) and a karate kick-off (**bottom row**).*

5 Conclusion and Future Work

We have introduced a novel, accurate, robust, easy to use, cutting-edge skeletal editing and interaction techniques that allow animators to manipulate arbitrary animation with real-time control. Violating physics of rigid motion allows us to obtain cartoon-like effect, independently of the skin layer. Spatial coherence of joints is preserved by differential-aware scaling represented by a quadratic energy function leading to a new *Skeletal Poisson Solver*. The usefulness and flexibility of our Laplacian approach is fully demonstrated to emphasize effective high-quality skeletal-based animation pose. In the context of filmmaking, our pipeline reduces the amount of time needed to animate superbly well-articulated expressive character. More importantly, this contribution makes complex interactions with animation more accessible to non-professionals, and offers real-time motion processing for video games as well. Notwithstanding, there is a large range of innovations to make interactive cartoon characters move naturally under exaggerated cues. For instance, selecting automatically good joint candidates for cartoon editing is an opened and unsolved problem for researchers. An interesting alternative idea is to incorporate *as-kinematic-as possible* attenuation to enforce the original kinematic curvature.

References

BREGLER, C., LOEB, L., CHUANG, E., AND DESHPANDE, H. 2002. Turning to the masters: motion capturing cartoons. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, ACM, New York, NY, USA, SIGGRAPH '02, 399–407.

DAVIS, J. W., AND KANNAPPAN, V. S. 2002. Expressive features for movement exaggeration. In *ACM SIGGRAPH 2002 conference abstracts and applications*, ACM, New York, NY, USA, SIGGRAPH '02, 182–182.

KWON, J.-Y., AND LEE, I.-K. 2007. Rubber-like exaggeration for character animation. In *Proceedings of the 15th Pacific Conference on Computer Graphics and Applications*, IEEE Computer Society, Washington, DC, USA, 18–26.

WANG, J., DRUCKER, S., AGRAWALA, M., AND COHEN, M. F. 2006. The cartoon animation filter. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006) 23*, 3 (July), 1169–1173.