Fat Pad Cages for Facial Posing

Adèle Colas  
IMT Atlantique, France  
adele.colas@telecom-bretagne.eu

Florent Guiotte  
ESIR, France  
florent.guiotte@uhb.fr

Fabien Danieau  
InterDigital, France  
fabien.danieau@interdigital.com

François Le Clerc  
InterDigital, France  
francois.leclerc@interdigital.com

Quentin Avril  
InterDigital, France  
quentin.avril@interdigital.com

ABSTRACT
We introduce Fat Pad cages for posing facial meshes. It combines cage representation and facial anatomical elements, and enables users with no artistic skill to quickly sketch realistic facial expressions. The model relies on one or several cage(s) that deform(s) the mesh following the human fat pads map. We propose a new function to filter Green Coordinates using geodesic distances preventing global deformation while ensuring smooth deformations at the borders. Lips, nostrils and eyelids are processed slightly differently to allow folding up and opening. Cages are automatically created and fit any new unknown facial mesh. To validate our approach, we present a user study comparing our Fat Pad cages to regular Green Coordinates. Results show that Fat Pad cages bring a significant improvement in reproducing existing facial expressions.

1 INTRODUCTION
Facial expressions are key elements in the realistic characters creation process [18]. Recent movie characters have facial rigs with thousands of blend shapes. Blend shapes represent facial elementary displacements and are often related to muscles [11]. Modeling them requires strong modeling skills and a lot of time (weeks for hero characters). Posing faces can be achieved through linear piece-wise modeling approaches [24] that require a lot of facial performances motion capture data. Physics-based approaches are also possible [8] but a high-end polyhedral facial anatomical representation is necessary to simulate bones, muscles and fat. Simpler solution would use any free-form deformation (FFD) techniques [22]. However, FFD are generic and do not take into account facial semantics. They can lead to severe uncanny artifacts for non-experts as a face is deformed as any other meshes. These non-expert users may be professionals, like stage directors who want to sketch facial expressions on a tablet to give guidelines, or beginners such as digital art school students. The final shape of a facial expression is a complex combination of the skull, the muscles and the fat [28]. As skin is soft and thin, it may rather be considered as a protective and aesthetic layer than a dynamic and active one. The skull gives the macroscopic shape of the face, and except for the jaws, it should not move. Actually, muscles are the cause of the action, and fat is the main visual element of the final shape (wrinkles, dimples, gaps and folds). Fat on the face is divided in several pads. The name “fat pad” comes from medical dictionaries and plastic surgery. Each pad moves almost independently if it is activated by the related muscle, and the skin joins all the pads together. We present the novel concept of fat pads for facial posing, enabling an interactive mesh deformation that respects the constraints of a face and is intuitive and easy to use for beginners. We propose a new automatic way for creating facial cages and a new weighted function based on geodesic distances to localize and limit the influence of cage deformation. Finally, we present a user study with 17 naive users that validates the easiness and intuitiveness of the Fat Pad cages.

2 RELATED WORK
Cages. Cages are polyhedra with a small number of polygons encompassing a mesh. Mesh vertices are linear sums of the cage’s vertices multiplied by weight functions (coordinates). Floater et al. [7] first proposed a way to use cages with a random geometry with the Mean Values Coordinates (MVC). But undesired artifacts appear with negative coordinates. Joshi et al. then proposed the
Harmonic Coordinates (HC) that resolve Floater’s issues but require a long computation time [10]. Lipman et al. put forward the Green Coordinates (GC) that do not have negativity issues [16]. Computation time allows interactivity, surface’s details and mesh volume are correctly preserved. Jacobson et al. propose three rules for cage generation [9]: it must be low resolution to reduce manual interactions, it must fully and tightly bind the enveloped model and it should respect the mesh topology for the manipulation to be intuitive. Yang et al. propose to use a template based cage generation type [27]. The nested cage [19] uses the character mesh to build progressive decimated cage mesh around the model. Xian proposes to generate oriented bounding boxes for each mesh part and then registering them together [26] but it is hard to apply on faces due to the lack of articulation. Chen et al. proposes a skeleton based approach [3] but face meshes do not have any skeleton. Lee et al.’s cage relies on user inputs [13] that cuts the model and the slides set is used to construct an initial cage, but our target is an automatic approach with no user interaction.

Facial Posing. To model faces, there are several approaches such as blend shapes [11, 14, 15, 20], region-based model [24] and physics-based model [5, 8, 21]. Blend shapes remain a predominant technique and are in most digital content creation tools. Tena et al. [24] propose a region-based linear model, that uses PCA. It separates the mesh in several models and the formulation restricts the solutions to have semi-consistent boundaries while enforcing user-given constraints. It does not allow to model other expressions or tissues in-between main fat pads. To avoid a complex modeling and fill the gaps between the pads, we propose to overlap them as shown on Figure 1. It avoids sharp borders and ensure smooth borders’ deformations. A plugin in Maya was developed to paint fat pads on meshes and place handles.

Attenuation Matrix. A weight $w_{v,h}$ is computed for each pad’s vertex $v$ and handle $h$, the closer the vertex to the border, the higher the attenuation. To avoid any border aberration, the weight must decrease slowly to zero at the border. We propose to use the following function that decreases according to the geodesic distance between the vertex position and the border,

$$w_{v,h} = \frac{(d(v, h) - d(i, h))^2}{d(i, h)^2}, \quad (1)$$

with $d$ the geodesic distance between two vertices, and $i$ the vertex located at the intersection of the border of the pad and the line $\overline{oh}$ (see Figure 2 and 3). The use of a quadratic radial basis function kernel appears appropriate as it smoothly goes from 1 (at the handle) to 0 (at the borders), thus restricting the influence on the fat pads. The function is used to compute the weight matrix $W_h$ for each vertex of each fat pad. This kernel has a zero derivative at the boundaries but it is not an issue in our case as we set the boundaries’ vertices fixed and unmovable (weight is 0).

Specific Borders’ Vertices. Lips are a good example where fat pads must be processed differently. Indeed, it has to follow the deformation of the fat pads to allow folding up, opening, and closing. Such borders can be specified in our model as an exception not to set the attenuation at the border to zero as displayed in Figure 3. It is the left upper lip fat pad with its two handles (blue) and its movable border (red). This is applied to the nostrils and the eyelids as well. In our posing process, we set the weight for this specific borders’ vertices to 1, thus it moves freely according to the handle motion as if it was not linked to another adjacent pad.

The final positions $V'$ of vertices after moving a handle $h$ is:

$$V' = V + (V_{gc} - V)W_h \quad (2)$$

with $V$ the current vertices’ positions, $V_{gc}$ the GC positions and $W_h$ the weight matrix associated to the handle $h$.

Pad Shape Processing. As fat pads may have different shapes, the weight matrix $W_h$ must be computed for each new mesh. $d(i, h)$,
A way to remove intersection candidates is to filter according to the distance between the handle and the border passing through vertex \( v \), a key value to compute \( W_h \). It is different for each vertex and each handle. To compute \( d(i, h) \), we set plane \( P \), defined by line \( L \) (between handle and vertex) and handle’s normal \( \hat{n} \). Plane \( P \) intersects the border. A vertex \( v \in A \) with handle \( h \), \( d(i,A, h) \), geodesic dist. between \( h \) and \( i \in A \) intersection between pad’s border and \( h \mathbf{v} \). B, C: non-convex fat pad, many intersections may exist.

Figure 2: A vertex \( v_A \) with handle \( h \), \( d(i_A, h) \), geodesic dist. between \( h \) and \( i_A \) intersection between pad’s border and \( h \mathbf{v} \).

Figure 3: Left: Fat pad movable border and handle. Middle: given a handle and a vertex, plane \( P \) intersects the border. Right: nostril fat pad has a concave shape.

the interaction with the pads. The duplication prevents this issue. To respect (2), cage’s vertices are moved away from the mesh along the normal to ensure no intersection between the mesh and the cage. First, we apply a uniform scaling to the upper part of the face and a double uniform one to the lower part due to smaller and tighter number of handles on the jaws. Then, based on their position on the face, a specific non-uniform adjustment scaling is applied to some vertices. For example, nose handles are only scaled according to z-axis (towards the image plane) to keep them in the middle of the face. To respect (3), two new vertices are positioned at the back of the head of the model. GC are meant to affect all the vertices within the cage. For instance if the upper lip handle is moved up to open the mouth, and the lower lip handle is moved down, vertices of the lower lip are moved relatively to the new position of those of the upper lip. The expected behavior is that they should be independent. To tackle this, two cages are built, one for the upper part of the head and one for the lower one.

4 RESULTS

Figure 4 presents posing results, obtained by a non-expert on three different facial meshes, compared to a professional artist (top row). Only parts with movable handles (not the ones added to close the cage) are displayed. When a handle is grabbed, the associated fat pad is colored in light blue and the deformation center with a red sphere (actual handle’s position on the pad). The artist used generic mesh deformers [22] whereas the non-expert has used our Fat Pad Cage. The three meshes were obtained through a digital double creation pipeline [6]. All meshes have the same mesh topology allowing the transfer of the template fat pads map. We selected two fundamental action units: AU-2 (Outer Brow Raiser) and AU-12 (Lip Corner Puller). They are involved in many facial human expressions.
(e.g. Happiness, Sadness, Fear etc.). Heat maps show differences between the initial mesh and the final modeling. Dark blue is static. We notice how close to the ground-truth the deformations are. It also emphasizes that our new Fat Pad cage prevents undesired mesh deformation due to global deformation (recomputation of the whole mesh when a cage vertex is moved) and ensures smooth deformation at the borders of the pads. With our cage, most of the mesh remains static, deformations only occur in the specific locations the user wanted to model.

Figure 5: Example of user results. From left to right: sadness (GC | ref | Fat Pad) and disgust (GC | ref | Fat Pad).

5 EVALUATION
A user study was conducted with two cage systems: GC and Fat Pad. Two expressions were used as reference to reproduce: "sadness" and "disgust". With each cage participants had to realize the two expressions, leading to four tasks. We have chosen to compare our method to a single GC cage as used in the literature, and because a two-parts GC cage leads to strong artifacts around the mouth. By design, a local GC cage moves the whole mesh. Objective measures were time, number of handles’ manipulation, number of undos, and the root mean square (RMS) of the Hausdorff distances between each participant’s result and the reference model [1]. Subjective measures were the following questionnaire evaluated on a 5-point Likert scale, from 1 (totally disagree) to 5 (totally agree): [Q1] The result is similar to the model, [Q2] The task was easy, [Q3] The manipulation was easy, [Q4] I am happy with my result. Participants had two screens, a front one displaying the interactive UI and a side one showing the reference meshes. It started with an exploration phase of 10 min to get used to the UI. The four tasks were then randomly presented. They stopped whenever they were satisfied with the result. Total duration was about 40 min. 17 participants took part in the study (age $\overline{X} = 33.75 \sigma = 12.920$). They were all used to 3D visualization and navigation ($\overline{X} = 3.65 \sigma = 1.069$ on a Likert scale 1 to 5) but no expertise in 3D modeling ($\overline{X} = 1.75 \sigma = 0.958$). A results sample can be seen on Figure 5. For the sad expression, results produced with the Fat Pad cage are visually similar to the GC cage, both close to the reference. However, for the disgust expression, the Fat Pad cage is closer to the reference because of the possibility to open the mouth. Figure 6 shows the mean of the RMS of the Hausdorff distance between meshes. The lower the better. Right: answers for sadness (up) and disgust (bottom). The higher the better.

cages ($p > 0.05$). For a simple task, the performance with the Fat Pad is not perceived differently from with the GC. This is in line with the visual analysis (Figure 5). However, with the disgust task, all assertions were better rated for the Fat Pad (Q1: $p = 0.001$, Q2: $p = 0.001$, Q3: $p = 0.03$, and Q4: $p = 0.002$). This preference is also visible from the produced meshes, and was confirmed by the post-test interviews.

6 CONCLUSIONS AND PERSPECTIVES
We presented Fat Pad cages, a first combination between cage-based deformation and facial anatomical model. The paper has described three main contributions: the Fat Pads concept enabling an interactive mesh deformation that respects facial anatomical constraints, a new automatic way of creating personalized cages for any facial mesh, and a user study validating the high interest and preference of Fat Pad cages. The new filter function appears to be more suitable to pose faces than Green Coordinates. It prevents global deformation and ensures smooth deformation at the borders of the pads. The generated cages closely fit the shape of the mesh and can be considered as an adaptive extension of the head. The user study confirmed the interest of our approach and provided valuable insight to improve our system. For instance, allowing the symmetrical control of the handles would be necessary to speed up the design of expressions. In a longer perspective, we want to focus our work to include machine learning to improve facial posing. Acquiring and analyzing facial fat motion would allow to better define pad’s shape and regions of impact. We also want to extend the concept of multiple cages to other facial parts while keeping making facial posing easier and more accessible to non experts.

REFERENCES
[1] Nicolas Aspert, Diego Santa-cruz, and Touradj Ebrahimi. 2002. MESH: Measuring Errors between Surfaces Using the Hausdorff Distance. Proceedings of the IEEE International Conference in Multimedia and Expo (ICME) 1 (Nov. 2002), 705 – 708. https://doi.org/10.1109/ICME.2002.1035879
[2] Sofien Bouaziz, Yangang Wang, and Mark Pauly. 2013. Online modeling for realtime facial animation. ACM Trans. Graph 32, 4 (2013), 40:1–40:10.
[3] Xue Chen and Jieqing Feng. 2014. Adaptive Skeleton-driven Cages for Mesh Sequences. Comput. Animat. Virtual Worlds 25, 3-4 (May 2014), 447–455. https://doi.org/10.1002/cav.1577
