

## Survey Paper

Deformation transfer survey<sup>☆</sup>Richard A. Roberts<sup>a,\*</sup>, Rafael Kuffner dos Anjos<sup>a</sup>, Akinobu Maejima<sup>b</sup>, Ken Anjyo<sup>a,b</sup><sup>a</sup> Computational Media Innovation Centre, Victoria University of Wellington, New Zealand<sup>b</sup> OLM Digital, Japan

## ARTICLE INFO

## Article history:

Received 12 June 2020

Revised 25 September 2020

Accepted 13 October 2020

Available online 21 October 2020

## Keywords:

Deformation transfer

Retargeting

Animation

Industry applications

## ABSTRACT

Deformation transfer is a type of retargeting method that operates directly on the mesh and, by doing so, enables reuse of animation without setting up character rigs and a mapping between the source and target geometries. Deformation transfer can potentially reduce the costs of animation and give studios a competitive edge when keeping up with the latest computer animation technology. Unfortunately, deformation transfer has limitations and is yet to become standard practice in the industry. This survey starts by introducing Sumner and Popović's [18] seminal work and highlights key issues for industry settings. We then review related work in sections, organized by these key issues. After surveying related work, we discuss how their advances open the door to several practical applications of deformation transfer. To conclude, we highlight areas of future work.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

Studios need to develop tools that enable artists to move beyond manual keyframing and raw motion editing to keep up with increasing demands for high quality animation.

Retargeting can provide studios with a competitive edge when used effectively. The goal of research in retargeting is to develop techniques that enable us to copy and paste animation between characters. One key advantage of retargeting, for studios, is that they can reuse animation. They might employ an animator to craft animation for a template character, but then copy that work to an entire crowd of orcs in a fantasy film or onto multiple side-characters in a game. Another important advantage is that if the director demands design changes, retargeting techniques can help artists to avoid losing work by transferring animation between design iterations. With these advantages, retargeting is an important tool for both small and large scale productions. When used effectively, it ensures animation work can be completed to a high quality in less time.

Unfortunately, commercially available solutions for retargeting are often not appropriate in many industry settings. As one example, the HUMANIK tool in Autodesk Maya lets an artist specify pairs of joints between the source (animated) and target (unanimated)

characters. The tool then transfers animation by copying changes in rotation between the pairs of joints. This solution is problematic in that (1) the pairings need to be entered for every unique pair of characters being retargeted, which is a repetitive and laborious task; (2) secondary animations are lost during transfer when they cannot be expressed by joints; and, perhaps most importantly, (3) retargeting across vastly different characters is not possible. Studios are left to implement their own solutions when these problems are prohibitive.

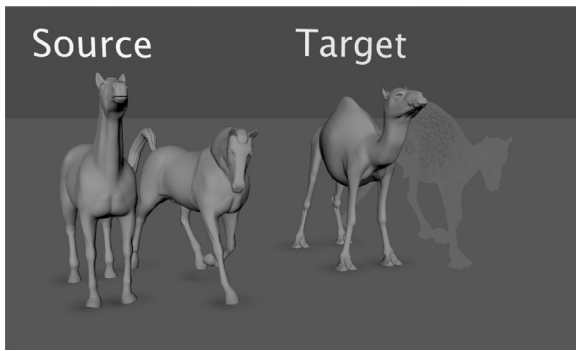
In deformation transfer, an approach to retargeting, the goal is to transfer animation via the mesh directly. Fig. 1 illustrates this goal. While not as simple to understand and implement as alternatives, deformation transfer offers the key advantage that retargeting is possible without the need to first engineer and map between character rigs. Thus, deformation transfer offers retargeting without placing a burden on artist time.

Unfortunately, there are limitations of seminal deformation transfer [18] that hamper a practical application. These limitations are addressed by more recent work, and some of their proposed solutions have already been used successfully in digital productions. We present this survey as an introduction to deformation transfer, with focus on exposing its potential for practical application in industry settings.

<sup>☆</sup> This article was recommended for publication by David Bommes.

\* Corresponding author.

E-mail address: [richard.roberts@vuw.ac.nz](mailto:richard.roberts@vuw.ac.nz) (R.A. Roberts).



**Fig. 1.** The goal of deformation transfer is to deform a target mesh (camel), by directly manipulating the mesh, such that it best recreates a given source pose (horse, with leg forward).

## 1.1. Outline

Section 2 introduces seminal deformation transfer and highlights key issues. We then introduce and discuss the related work in sections based on which of the key issues they solve: Section 3 includes work that changes shape and deformations representations to favour performance, to broaden the range of applicable meshes, or to limit artifacts; Section 4 focuses on techniques that use implicit models to enable partial or full automation over the process; and Section 5 describes how deformation transfer of semantic properties can be realized. As an auxiliary, Section 6 introduces work that highlights how transfer often conflates deformations resulting from shape and from pose. Insight from this work can help to explain why some methods produce unnatural results. To conclude, we summarize and discuss a table listing the surveyed work (Section 7), discuss potential for industry applications (Section 8), and provide ideas for future work (Section 9).

## 2. Seminal deformation transfer

In 2004, Sumner and Popović [18] introduced deformation transfer as the first retargeting solution that operates directly between meshes. Since the source and target rarely match geometrically, the underlying challenge is to develop a principled way to copy a change in the source pose onto the target. They address this challenge through a two-step process that first corresponds the source and target with one another, and then performs an optimization process to transfer the deformation. Throughout this paper, we use the title *seminal deformation transfer* to denote Sumner and Popović's two step approach of building the correspondence map using an optimization method and performing transfer using a linear system that represents deformation via deformation gradients.

As input, the artist should provide the source and target in their reference poses. Conventionally the reference pose has the characters in a natural stance, but any pose for the reference is possible provided that the source and target are both posed in the same way.

The first step of seminal deformation transfer is to build the correspondence map. The correspondence map specifies exactly how the triangles of the source character map onto the target, and vice-versa. To build the map, the artist should first specify a set of corresponding points. Given these points, an optimization algorithm finds the best match between the source and target and, once matched, nearest triangles are considered to be corresponding pairs.

Next, the artist provides a new deformed pose for the source. The transfer step aims to pose the target to match. In this second

step, the change between the source's triangles in their reference and deformed states are modelled by a set of deformation gradients. In practice, a deformation gradient is an affine matrix that encodes how a triangle rotates and scales to transition from its shape in the reference pose to that of the deformed pose. To find a deformed pose for the target, they propose a least-squares solver that chooses target vertex positions such that the target triangles best recreate the deformation gradient of their corresponding source pairs.

In this section, we introduce the correspondence and transfer steps in greater detail (Sections 2.1 and 2.2). We then highlight the key issues that we identified in discussion with our industry partner (Section 2.3).

### 2.1. Correspondence step

While a mapping is obvious when the geometry of both the source and target are similar, it is difficult when this is not the case. For example, consider how it might be difficult to correspond the humps on a camel's back to the spine of a horse (at the level of triangles).

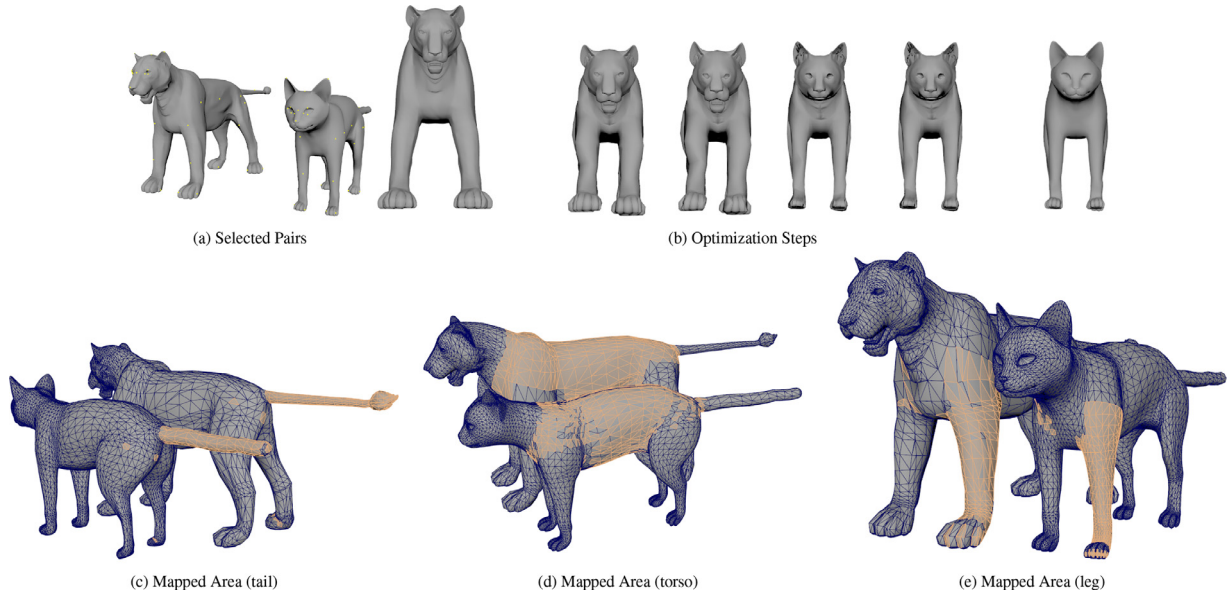
Borrowing from template-fitting algorithms, seminal deformation transfer proposes a method to build the correspondence map. In their method, they employ an optimization scheme that warps the source until it matches the target exactly, or vice-versa. In other words, one mesh is deformed to *become* the other. Once fitted in this way, pairings can be found by identifying the source and target triangles that are closest to one another. To initialize this step, the artist should first select pairs of points that outline how the source and target correspond with one another. Fig. 2(a) provides an example. Next, an optimization method tweaks vertices of the target until a pose that places the correspondence pairs together while also ensuring that the target mesh does not become malformed. The optimization method achieves this by minimizing error function that measures (1) the distance between correspondence pairs, (2) the amount of deformation, and (3) local deformation smoothness. Results created with our testing implementation are shown in Fig. 2(b).

The correspondence map can be built once the target has been fitted. To build the map the algorithm finds all similarly oriented source triangles nearest to a given target and vice-versa, resulting in a many-to-many mapping. Fig. 2(d)–(f) presents examples of mapping regions. With the correspondence map in place, the first issue of encoding the geometric relationship between the source and the target has been solved.

### 2.2. Transfer step

With the correspondence map in place, the second step is to calculate a target pose that best recreates deformations observed for a given source pose. The challenge underlying this step is two-fold: one the one hand, assigning a deformation for each triangle independently results in a target pose where the surface is no longer intact (edges of adjacent triangles would become disconnected); and, on the other, we can observe that there will rarely be a solution that features optimal deformations for each target triangle (there may be multiple source deformation gradients for each target triangle). Thus, we need a method to choose target triangles deformations that both keep the target surface intact while also faithfully recreating the deformed source pose.

As a potential solution, Sumner and Popović suggest using an optimization method. Like the one used for the correspondence step, the method tweaks deformation gradients of the target triangles to minimize an objective function under the constraint that adjacent edges stay connected. While simple to implement



**Fig. 2.** In seminal deformation transfer, the correspondence map is calculated from a sparse set of corresponding points selected by the artist (a). From this input, a series of four optimization steps are employed to warp the target lion into the source cat (b). Once the target has been transformed into the source, a many-to-many mapping is formed between nearest valid triangles. A few examples of corresponding areas are displayed (c)–(e). Transfer becomes computationally tractable given this mapping.

this approach, unfortunately, would be too slow for practical application.

To enable viable performance, Sumner and Popović develop an alternative method for the transfer step. By solving over vertices, instead of triangles, they avoid the complication that the mesh surface can become disconnected. The challenge, in this case, is to determine how to best move the vertices to recreate the appropriate triangles deformations. Impressively, they designed a linear system that models this problem using deformation gradients. Their linear system places the deformation gradients of the source on one side of the linear system. Then, on the other side, a matrix that negates the target's reference pose and a vector of unknown vertex positions; when these are multiplied together, deformation gradients for the target triangles are produced. Solving the linear system finds the deformation gradients that produce a target pose best matching the source. While more difficult to understand, this linear system enables a much faster and practical solution to deformation transfer.

### 2.3. Key issues

**Performance.** Even with the linear system in place, it is large and cannot be solved fast enough for interactive applications [5,7,25].

**Broadness.** The deformation gradients representation is only compatible with triangle meshes [5,9,26]. While triangulation can be used to address this issue in part, doing so exasperates performance overheads. Transfer between meshes featuring multiple components is not possible.

**Artifacts.** Deformation transfer results tend to exhibit artifacts that detract from natural appearance [12,16].

**Artist input.** Specifying pairs of points to initialize the correspondence map can be a laborious task [13,22]. This problem is exasperated in that tweaking the selected pairs does not lead to a proportionate change in the transfer. This disconnect inhibits artist from intuitively refining results.

**Automation.** A lack of automation limits the range of applications to those in which an artist is available [6,22].

**Semantic Transfer.** Finally, while geometric properties are transferred, semantic properties are not.

## 3. Representations to improve deformation transfer

Motivated to improve upon key issues that we outlined earlier in Section 2.3, many related works have proposed new or extended ways to represent and transfer deformation. Some choices of representation help speed up transfer, others enable more robust application, and others limit the occurrence of artifacts.

Here we survey previous work that focuses on improving performance (Section 3.1), enabling broader application (Section 3.3), and limiting the occurrence of artifacts (Section 3.2). We conclude with discussion in Section 3.4.

### 3.1. Improving performance using cages

Previous work has reduced the computational overhead of transfer, primarily by using a cage that offers a lower-dimensional interpretation of the mesh.

Ben-Chen et al. [5] present the first method to use a cage for transfer. They first provide an algorithm to build the cage, which starts from a dense sampling of the original mesh's surface and then interactively removes and repositions vertices. The resulting cage ensures sparsity and a tight-fitting. The relationship between the cage and the underlying mesh is encoded using the variational harmonic functions of [24]. The functions form a basis that, when weighted appropriately, can modify the pose. A least-squares solution is used for transfer, in which an algorithm calculates offsets for target cage vertices such that the deformed target best matches the given source (where closeness is measured in terms of artist-specified points, which operate in place of the correspondence map).

In similar previous work, Chen et al. [7] also propose to enclose the source and target in a cage and perform transfer by optimizing positional changes in a sparse set of artist-defined landmarks. Making their work distinct from that of Chen et al., they use Green coordinate interpolation to propagate changes of the target cage back to the target mesh. They argue that the Green's coordinate interpolation, which is biased to preserve angles between edges over their lengths, improves transfer in detailed areas.

Most recently, Yifan et al. [23] proposed a novel technique that employs a deep learning model – called CageNet – to learn

cage deformation. Where deformation transfer aims to recreate a change between poses, the focus of their work is to deform a given source model to take on the shape of a target model, while preserving local geometric details. In their approach, they first encode both meshes into a latent space and then apply two different decoders: one decoder creates a cage for the source mesh, while the other creates an offset that should be applied to that cage's vertices to best reproduce the target shape.

While not the primary focus of their work, Yifan et al. demonstrate how their approach can be modified to perform deformation transfer of human models. They train a model of cage deformation over a database of exemplary motions, in which they learn to fit the cage to best reproduce poses from that database. Next, they use this trained model to transfer deformation: given a new human mesh with a sparse selection of artist-selected landmarks and a desired pose from the original database, they first align the source to the target reference pose, then employ an optimization step to generate a cage for this source model, and finally query the trained model to obtain an offset for the generated cage that produces the deformed target pose.

Cages are a powerful solution that exploit sparse representation to enable fast performance. With the enhanced performance, these techniques make deformation transfer suitable for interactive use. This is a critical advantage for any applications where artists need to explore and potentially refine transfer results interactively. Unfortunately, the sparser representation has the drawback that fine-scale deformations can be lost. In applications requiring higher fidelity, such as transfer of wrinkling details in faces, typical cage representations are inappropriate.

### 3.2. Limiting artifacts for better transfer

Other previous work has focused on the issue that the resulting target pose often features visually unnatural artifacts. Some of the more common artifacts of seminal deformation transfer are crumpling and self-intersection.

Zhao et al. [25] introduce the dual-mesh representation for deformation transfer.<sup>1</sup> Distinct from seminal deformation transfer, the dual mesh representation encodes and optimizes deformation in terms of surface normals, which helps to eliminate artifacts that arise in areas with fewer vertices or with complex shapes.

Saito [16] extend the linear system used in seminal deformation transfer with new constraints over intersection and smoothness. For intersections, they add *virtual* triangles that encapsulate the holes in the model, such as eye holes and the mouth. The virtual triangles are appended to the linear system used for solving transfer, which helps to preserve the shape of the holes as well as the mesh and, consequently, intersections across these holes are unlikely to occur. Second, they add a Laplacian-based regularization term that leads to smoother deformations. With these two terms added to the linear system, transfer results exhibit fewer artifacts.

In another work Shabayek et al. [10] observe that Euclidean spaces can be ineffective for representing deformation. They adapt the Lie Bodies representation [11] to decompose triangle deformations into a group of transformations expressed using a Lie Algebra. A manifold is formed by combining the elements of the algebra for each of triangle deformations, such that each pose corresponds to a point on that manifold. Using this manifold, in which degenerate cases of deformation are avoided by design, they present impressive results with fewer artifacts. Furthermore, other advantages of this manifold are that interpolation and even composition of deformations become possible, which could be

a significant advantage for some applications, such as simulated animation involving many incremental transfers. Other advanced operations like measuring variance between poses are possible.

Resolving transfer artifacts would become a common task if deformation transfer were to be applied more broadly in industry. In some cases, an artist would find that fixing a few minor artifacts by adjusting vertex positions would be relatively trivial. However, in other cases, the artifacts may be too numerous or occur in complex areas and demand more intensive work to resolve. Furthermore, in other applications, an artist may not be available (perhaps due to limited budget or real-time application), in which case artifacts will detract from quality. The work introduced here helps to combat these issues: Zhao et al.'s dual-mesh can help to limit artifacts without requiring additional artist work, and Saito's virtual triangles help to prevent intersections (particularly important for facial animation). In broader applications, where it may be difficult to create virtual triangles, artifact-free transfer remains an open problem. In these cases, we can take inspiration from Shabayek et al.'s Lie Bodies representation that could provide more natural looking results.

Despite their advantages, these techniques retain higher computational costs and, unlike the cage-based approaches, are not suitable when interactive performance is required.

### 3.3. Broader mesh types

Another focus of previous work has been to enable transfer for more generic mesh types.<sup>2</sup>

Domadiya et al. [9] introduce a vector graph representation, which enables deformation transfer to be applied to meshes with any type of polygons. The vector graph extends the mesh by placing a new vertex at the centre of each face and then adding new edges that span these new vertices. This process effectively triangulates the mesh and, while this would generally slow down the solution, they introduce an optimization that scans through the correspondence map to select a subset of elements (approximately half) to use when solving for transfer. This optimization makes their vector graph amenable for transfer with similar performance to seminal deformation transfer.

Zhou et al. [26] propose a solution to enable transfer between multi-component meshes. Their extension finds spatial relationships between the multiple components of a character and uses these to define a new error term that is minimized when those spatial relations are preserved by transfer. This new error term is non-linear and so they must use the optimization method for transfer (described in Section 2.2; however, they demonstrate that each step in the optimization scheme is linear, which means that performance is still amenable for some applications. Despite this drawback, their solution enables impressive transfer between characters composed of multiple parts and opens the door to advanced transfer applications (perhaps deforming a cloud of particles based on a template animation, or between a template animation and a robot composed of many small parts).

Enabling deformation transfer for a broader range of mesh types, while retaining the ability to be computationally feasible is a difficult problem. In one sense, the previous work on enveloping meshes using a cage-representation could already solve this problem; however, a key drawback of these techniques is that they lose fine-scale details. The methods presented here operate on the

<sup>1</sup> Au et al. [2] developed the dual-mesh representation for editing meshes using Laplacian operations. These types of operations are common in applications that blend two or more images or meshes to appear seamless.

<sup>2</sup> Seminal deformation transfer supports only triangular meshes. Other types of meshes are also used in computer graphics, such as meshes containing quadrilateral polygons and even polygons with higher numbers of vertices. Some meshes combine different types of polygons, which are typically called *hybrid meshes*. Furthermore, some meshes are composed of multiple separate parts, called *multi-component meshes*.



source and target meshes directly, enabling transfer for a broader range of meshes without the drawback of detail loss.

Despite the advantages of these techniques, they also retain higher computational costs and, again, are not suitable when interactive performance is required.

### 3.4. Discussion

In this section, we surveyed related work that tackles key issues of performance, artifacts, and generality.

To improve performance, a sparse representation – typically a cage that envelopes the mesh – have been proposed [5,7,8,14,25]. The cage offers significant gains in performance as the transfer can operate over a much sparser representation. However, this gain in performance tends to come with the cost of detail loss.

To reduce artifacts, the related work has proposed to extend transfer with a representation of the negative space and with a preference for deformation smoothness. Virtualization of negative space can be achieved by adding virtual triangles and deformation smoothness can be encoded either implicitly through an alternative deformation representation (like the dual-mesh in [25]) or explicitly through a regularization term (like the Laplacian in [16]). Finally, most recently, Shabayek et al. [10] have introduced a non-Euclidean deformation representation that avoids degenerate transformations by design.

Finally, extensions have been proposed to generalize the range of meshes to which transfer can be applied. Transfer can be performed for meshes containing any types of polygons with the vector-graph [9], and for multi-component meshes when spatial relationships are found and added as a new term to the transfer method [26].

While an all-encompassing solution for fast, artifact-free, and generalized deformation transfer is yet to be proposed, the advantages provided by the surveyed work can already be used to enable practical applications of deformation transfer in many settings.

## 4. Toward automatic correspondence

Seminal deformation transfer requires the artist to manually specify pairs of points. Allowing manual input from the artist is a desirable feature, especially in production scenarios where an artist can tune the selection of correspondence pairs to affect transfer results (at least through trial and error). However, there are other situations where automation is helpful.

Motivated to provide partial or full automation, related work has proposed novel methods to reduce artist involvement when initializing the correspondence map. This may be done by (1) developing a method to help find correspondence points or by (2) using an implicit correspondence map, which omits the need for the artist to specify correspondence pairs. With either approach, deformation transfer can be applied with less artist involvement. The task of shape correspondence or matching, which can be used for finding an explicit or implicit correspondence map, has applications in several areas other than deformation transfer such as 3D scan alignment, reconstruction, and classification. Outside of deformation transfer, other work has developed approaches to match data of different representations (points, surfaces, skeletons) and dimensions (2D, 3D, temporal or non-temporal), providing different type of correspondences (dense, sparse, full, partial, probabilistic, one-to-many, many-to-many, affine or rigid transformations), and taking different approaches to the problem of correspondence mapping. The survey from van Kaick et al. [20] provides an in-depth review of different approaches and discusses their use in alternative applications.

While the broader field of shape correspondence and matching could be applied to deformation transfer, here we introduce only the correspondence methods proposed in work on deformation transfer.

### 4.1. Finding correspondence pairs

Bian et al. [6] present a fully automatic approach that finds correspondence pairs for transfer between faces. In their approach, they search a mesh projected onto a 2D image to find landmarks around features such as the eyes. They find one landmark in each eye corner, two landmarks in each of the upper lids, and two in each of the lower lids. Once found for both the source and the target, they use the inverse projection to derive which vertices of the mesh match the identified landmarks. Through doing this landmark search for both meshes, they can automatically find points correspondence pairs. While this approach is successful in automating over faces, their search mechanism cannot be extended trivially to other applications. Nevertheless, this principle of using domain-specific knowledge to find similar points between the source and target is novel in that it can automatically suggest candidates to the artist (thus reducing overheads) or, when the found correspondence pairs are already sufficient, be used to automate the algorithm altogether.

Based on the observation that the task of choosing which points to use for correspondence is complex, yet the task of finding a point corresponding to a given point is more simple, Yang et al. [22] explore how to automatically choose ideal correspondence points for the source (leaving the task of pairing them to the artist). To choose the source points, they employ harmonic analysis [24], segmentation, and clustering. Next, they identify a point representing each cluster and provide this set as candidates. The artist then completes the easier task of finding their pairs on the target model. Since the pairing task is easier, significantly less artist time is required to initialize the correspondence map. Interestingly, while the resulting correspondence pairs could be used as input to the seminal method, Yang et al. propose an alternative transfer method where deformation is copied between the source to target pairs directly (with an automatic skinning step used to deform the target mesh to best fit the updated handles). This direct scheme is significantly faster although, much like the cage representation, is prone to detail loss.

### 4.2. Implicit correspondence map

As another approach, deep learning methods excel in terms of enabling implicit mapping of correspondence and a powerful solution for full automation.

Gao et al. [13] combine a variational autoencoder (VAE) with a cycle-consistent generative adversarial network (CycleGAN) to model the correspondence mapping and transfer processes. In their technique, they set up a pair of variational autoencoders (VAEs), one for modelling the source character and the other for modelling the target character. Each VAE is similar in structure to MeshVAE [19], a VAE trained with a rotation-invariant metric that encodes a given pose to a vector in a compact latent space. Different from MeshVAE, which uses fully connected layers, Gao et al. use convolutional layers that improve generalizability. Furthermore, they use an as-consistent-as-possible (ACAP) deformation representation [12] that defines features in terms of vertices. The ACAP deformation representation is beneficial in that convolutional operations are easier to define and larger deformations are better handled.

Using the latent vectors from the VAEs as input, they train a CycleGAN to model a pair of generators: one generator maps a vector from the source latent to another in the target latent space,

while the other generator performs the reverse mapping. The outputs from each generator are then decoded (using the VAEs) to obtain the estimated pose. To train the CycleGAN model, poses from a database of varied human motion are organized into source and target pairs and these poses are then encoded into their latent space vectors through the VAEs. They train the CycleGAN to improve the generators until they can faithfully reproduce the ground-truth poses. Once trained, they generators are effective for transfer between unseen human source and target characters.

Most recently, borrowing a model architecture developed for style-transfer in images, Wang et al. [21] introduced the first solution for deformation transfer without the need for a source reference pose. In their solution, they use an encoder-decoder neural network structure to model the correspondence and transfer process. Given a deformed source pose, they encode it to a latent space using a pose feature extractor that observes localized details of the mesh. Their decoder features several layers, inspired by style-transfer methods, and estimates a target pose that features similar local details to those observed for the source. A unique advantage of this solution is that it can perform transfer without ever observing the deformation directly (there is no reference source pose); however, it may be prone to creating artifacts when the shape of source and target characters differ significantly.

Overall, deep learning solutions offer a way to leverage all the information within meshes to enable powerful and automatic deformation transfer. The use of latent spaces means that no explicit definition of correspondence between is required and, therefore, a distinctive advantage of these techniques is that the artist need not manually define correspondence pairs between the characters. Furthermore, no extra computation time is required for calculating the dense correspondence map. One disadvantage, however, is that a database is required for training. This need has currently limited the application of these techniques to characters where those data are available, such as human scans.

#### 4.3. Discussion

The approaches surveyed in this section enable many practical applications of deformation transfer.

When we can make assumptions about the domain of the problem, we can take inspiration from Bian et al.'s work [6] to develop heuristics to automatically find correspondence pairs. This approach is most easily applicable to transfer between faces but may also prove useful to other applications such as transfer between similar fantasy characters. Employing heuristics based on domain-specific knowledge has the advantage of being fast to compute but also the drawbacks that those heuristics can be difficult to discover and that, once developed, are limited to their specific applications.

Deep-learning approaches are potentially the most powerful. They avoid the need to specify correspondence by encoding it indirectly through latent space. While expensive to train, the resulting models enable transfer that is both fast and automatic.

Unfortunately, extensive data is required for training, which will not be readily available outside of human characters (and perhaps domestic animals where motion capture may be used). Furthermore, without any way for the artist to guide the deformation result, deep learning solutions are only applicable to situations where the result is already suitable for the intended application, which limits their practical use in productions of films and games where significant artistic refinement will be required. Given the restrictions imposed by learning methods, we might consider that Yang et al.'s [22] method of automatically choosing candidates to reduce artist time required for initialization is the most feasible approach to those production situations.

## 5. Semantic transfer

Seminal deformation transfer solves transfer by minimizing deformation gradients that encode geometric differences between reference and deformed poses. However, there are many cases where geometrically corresponding a given source and target is not possible: how should we correspond a flamingo with two legs to a horse with four?

Baran et al. [4] present the first approach for transferring deformation semantically, rather than geometrically. In semantic transfer, the idea is to pose the target to recreate the meaning of the source pose, more so than changes observed in geometry. The key idea behind semantic transfer is to set up two spaces, one for the source and one for the target, that semantically match one another. The spaces enable pairwise interpolation of exemplary poses to produce semantically matching results, and thus transfer can be performed by first projecting into the source space and then interpolating in the target space.

In summary of Baran et al.'s algorithm: two sets of matching poses for the source and target are provided as input. For example, the first pair of poses might feature the source and target standing, the second pair might feature the top of a jump, the pair poses might feature a crouch, and so on. Whatever each pair of poses depict, they must be a semantic match. Given the sets of poses, they convert them into coordinates that span a low-dimensional shape space. Since the coordinates of the source and target space share semantic meaning, the spaces implicitly correspond to each other. Due to their correspondence, transfer can be performed by projection and interpolation. First, their algorithm projects a given source pose into the source space to determine its coordinate. Through the projection, they obtain a set of weights that describe how to combine the basis coordinates (the coordinates of the original source poses) to best recreate the given deformed pose. Second, they interpolate the target space using these weights to obtain a target coordinate that corresponds to the identified source coordinate. Finally, to obtain a target pose, they employ a least-squares solver to choose vertex positions that, when projected, is nearest to the interpolated coordinate.

Semantic deformation transfer is ideal for applications that demand transfer between characters of vastly different shapes. As one example, consider the task of animating a horse from a human mocap. The artist might choose to scan through the mocap and identify a few representative poses, from which they pose the horse manually. After creating a modest library of poses, they can apply semantic deformation transfer to transfer the rest of the motion automatically.

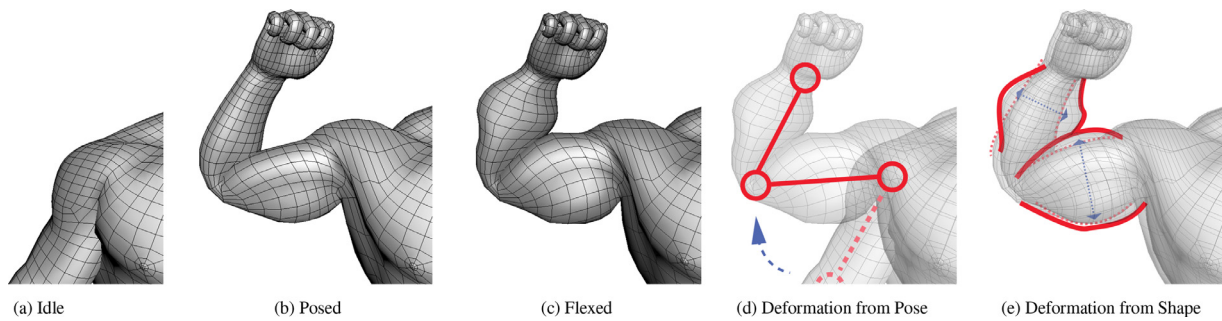
Like the deep learning methods described in Section 4.2, semantic deformation transfer also encodes the correspondence map implicitly and therefore removes the need for the artist time to manually define correspondence pairs.

While Baran et al. propose the only solution for semantic transfer, one drawback is that two sets of poses that sufficiently define the deformation space must be provided. While extensive data is not required, it may be infeasible to produce the poses sets in production settings [9,22]. For example, in a production setting using fantasy characters, it may be too expensive to have the artist make several sculpted poses.

## 6. Decoupling shape and pose deformations

Another idea that must be considered, tangential to the concept of semantic versus geometric deformation transfer, is whether deformation results from a change in pose or is unique to the shape of the source or target.

Pose-based deformation is any deformation that results directly from a change in pose. It is deformation that occurs independently



**Fig. 3.** The example above highlights the difference between deformations relating to pose and shape. The example depicts a close view of the right-arm of a muscly character. There are three key poses: the arm is relaxed at their side (a), the arm strikes the pose but remains relaxed (b), and the arm now flexed (c). Deformation lies between the poses: in (a) and (b) deformation relates to pose (d), and in (b) and (c) deformation is unique to the shape of this muscly character (e). Changing pose directly from (a) to (c) would involve both the pose (d) and shape (e) deformations occurring simultaneously. Consider how a weak character would have smaller muscles and therefore not feature the same shape deformation, in this case seminal deformation transfer will correctly apply pose-based deformations but may also erroneously transfer shape-based deformations. By considering the impact of conflating pose and shape deformations we can see how artifacts occur in seminal deformation transfer – imagine how shape deformations vary between skinny and overweight characters, short and tall characters, and young and old characters. *Model provided by Turbo Squid, under Royalty Free License.*

from the shape of the character. For example, an arm will bend as the elbow joint closes. In contrast, shape-based deformation is any deformation unique to the shape of the object. For example, the bulging unique to a muscly character. Fig. 3 illustrates these differences.

Anguelov et al. [1] developed a method that learns two parametric models that separate shape and pose deformation. Given scans of different poses, they first deform a template mesh to match and, from these deformed poses, compile a pose model. The parameters of this model adjust only the pose of this template character. Next, they construct another model, but this time develop parameters that vary its shape to deform the template into a variety of scans that depict different humans. With the two models in place, Anguelov et al. demonstrate that natural pose-based deformations can be created for a range of characters despite their shape variation.

The separation between pose-based and shape-based deformations, as highlighted in Anguelov et al.'s work, reveals a key problem for deformation transfer. Transfer, when solved geometrically, conflates both the pose and the shape aspects of deformation. This conflation is one reason why artifacts occur in the target pose resulting from transfer. If we can decouple pose-based and shape-based deformations during transfer, we may be able to avoid much of these artifacts.

For now, decoupling deformation in its pose and shape components remains an open problem for deformation transfer. Balan et al. [3] expose one possible direction in related work: they present a technique that tweaks Anguelov et al.'s model parameters to pose a character that matches both the shape and pose observed in images of people. Their results demonstrate that the fitted model accurately depicts the pose and the shape of the human. One could consider an approach that performs transfer over corresponding pose-based and shape-based parameters, although this is yet to be explored.

Interestingly, approaches that model correspondence implicitly tend to avoid the problem of inadvertently transferring shape-based deformation. In semantic transfer (Section 2), conflating pose and shape deformations is avoided as a separate shape space is used for the source and the target models (their transfer cannot inadvertently copy over shape-based deformations because of this separation). The recent work on deep-learning (Section 4.2) also avoids the problem implicitly, as the models are trained against sets of poses unique to each character (thus, again, separation avoids the problem).

## 7. Summary of related work

Table 1 lists the surveyed work that introduces novel methods for deformation transfer. In this section, we summarize the work with a focus on their choices of shape and deformation representation, denoted by columns 2 and 3. We also comment on how these representations underpin the type of correspondence mapping and the method of transfer (columns 4 and 5), along with their key advantages and limitations (columns 6 and 7).

Successful deformation transfer is heavily reliant on an effective representation of both the shape (mesh) and the deformation being transferred. Recalling Section 2, seminal deformation transfer proposes that we imagine forming a tetrahedron over each triangle. The tetrahedron connects each of the vertices along with an additional vertex that sits at the end of the face normal. Since this family of tetrahedrons, together, express the shape of the mesh, we refer to them as the *shape representation*. Using the shape representation, we can easily define a deformation for a given pose: we calculate an affine matrix that transforms each tetrahedron from its shape in the reference pose to that of the given pose. By performing this calculation for each tetrahedron,<sup>3</sup> we can fully express the deformation of the mesh; thus, we refer to the resulting set of affine matrices as the *deformation representation*.

The representations used in seminal deformation transfer are ideal in that they fully capture the deformation of the mesh. Due to this advantage, it remains a common choice when surface deformations are the focus of transfer (used in [6,16–18,25,26]). However, this representation is granular and, consequently, many deformations must be transferred such that a large linear system is required and renders the algorithm too slow for interactive use. Furthermore, the optimal deformations per triangle can conflict with one another and so several artifacts can arise as at least some sub-optimal deformations must be present in the results. Finally, the representation is only suited to triangle meshes, which does not support broad application.

Some works extend the representations of seminal deformation transfer with additional properties. Recalling Section 3.2, Zhao et al. [25] use a dual-mesh shape representation that limits

<sup>3</sup> Note that, in practice, adding the extra vertices to form tetrahedrons would be disadvantageous since doing so would increase the size of the linear system for transfer. As summarized in [17], we can derive deformation gradients that do not require the extra vertex by examining the deformation of edge matrices (hence the notation of edges in column 2 of Table 1).

**Table 1**

A summary of key work surveyed in this report. The first column specifies the citation and title, columns 2 and 3 denote the representations used to model shape and deformation; columns 4 notes the type of correspondence mapping and 5 the method of transfer; and columns 6 and 7 summarize advantages and limitations.

Reference	Shape Rep.	Deformation Rep.	Correspondence	Transfer Method	Key Advantages	Key Limitations
[18] Sumner and Popović	triangles	edges, affine	dense	least squares		triangles only, artifacts, speed
[5] Ben-Chen et al.	cage	landmarks	sparse	least squares	faster, broader application	fine-scale loss
[7] Chen et al.	cage	landmarks	sparse	optimization	faster, broader application	fine-scale loss
[23] Yifan et al.	cage	landmarks	sparse	deformation network	learn cage deformations	fine-scale loss, need data
[26] Zhou et al.	triangles + spatial	edges, affine	dense	optimization	broader application	speed
[9] Domadiya et al.	vector graph	vertices, frames	dense	least squares	broader application	speed
[16] Saito	triangles + virtual	edges, affine	dense	least squares	limit artifacts, enable constraints	triangles only, speed
[25] Zhao et al.	dual mesh	vertices, affine	dense	optimization	limit artifacts	triangles only, speed
[10] Shabayek et al.	triangles	triangles, groups	dense	least squares	limit artifacts	triangles only, speed
[6] Bian et al.	triangles	edges, affine	dense	least squares	automatic correspondence	triangles only, speed, artifacts
[22] Yang et al.	clusters	landmarks	sparse	copy directly	semi-automatic correspondence	fine-scale-loss
[13] Gao et al.	latent	ACAP [12]	implicit	mapping functions	automatic correspondence	need data
[21] Wang et al.	latent	pose feature	implicit	decoder	transfer without source identity	need data
[4] Baran et al.	shape space	coordinates	implicit	project + interpolate	faster, semantic transfer possible	triangles only, artifacts

artifacts when paired with a Laplacian error term, at the cost of requiring a more expensive optimization process for transfer. And, recalling Section 3.3, Zhou et al.'s [26] method appends spatial relations that enable stable transfer for multi-component meshes. These appended elements successfully enable broader transfer but have the drawback that a larger linear system must be solved (although, in practice, this overhead should be relatively minimal).

It is also possible to use alternative surface-based representations. For example, Domadiya et al. [9] employ a vector graph as their shape representation, that effectively converts a hybrid mesh into a triangular mesh. For deformation, they create local coordinate frames describing how each vertex of the graph moves between the reference and deformed poses. As well as enabling broader application, their formulation has another advantage in that each deformation is expressed in a local coordinate system that more easily enables post-processing (they apply a Poisson interpolation post-process to improve temporal properties of transferred sequences). As another example, Shabayek et al. [10] employ a deformation representation that encodes each triangle deformation as a group containing a rotation, in-place deformation, and anisotropic scaling. By projecting these transformation groups into a special type of manifold that features a Riemannian metric, they can produce results that appear more natural than those of seminal deformation transfer. Furthermore, their deformation representation enables new operations such as interpolation and composition. While powerful, this approach is currently limited to triangle meshes.

Other works combat computational complexity, primarily by using a cage as the shape representation, paired with constraints at landmarks for the deformation representation [5,7,22]. Using a cage for the shape representation has the key advantage that the linear system to be solved is smaller and thus can be executed fast enough for interactive use. While fast and broadly applicable, the sparsity of the cage means that finer-scale deformations are lost.

Another powerful approach is to represent shape and deformation in a way that enables implicit correspondence of the source and target. Recalling Sections 4.2 and 5, several methods have been proposed for implicitly representing the correspondence between the source and target. A pair of shape spaces are used in semantic transfer [4], and a latent space in the deep learning methods [13,21]. The implicit representations reduce the burden on artist time (since the artist no longer needs to manually

identify corresponding points to initialize the algorithm) and can potentially offer automation. Furthermore, they are the only techniques that have the potential to transfer between characters of significantly differing shapes. While powerful, these methods require databases to train and this limits their application to data that is readily available.

In summary, the choice of shape and deformation representation is perhaps the most critical consideration when choosing one of the above techniques, and this choice often prescribes what the type of correspondence mapping and transfer method are to be used. The deformation gradients, of seminal deformation transfer, are still reasonable as the default choice; they excel in capturing the full surface deformation and transfer can be applied by pairing a dense correspondence map with a least-squares solver. When speed is an important factor, the cage representation enables a smaller linear system that can be used for interactive processing (an artist can see the results in real-time). Alternatively, if preserving fine-scale details while also reducing artifacts is a focus, then one should consider techniques that add additional factors to the shape representation (such as the dual-mesh representation, the vector graph, or the virtual triangles). When automation is favourable, then one should consider the latent representations that are used by the latest deep-learning techniques to enable implicit correspondence. Finally, for transfer between vastly different characters, one can look to the work on semantic deformation transfer.

## 8. Industry application

Through the discussion of Section 7, we can conclude that deformation transfer is already developed to a point where it can be applied as a practical solution in industry settings.

With a correspondence map set up in advance, cage-based techniques can support transfer in interactive media such as video-games or interactive VR/AR experiences. Surface-based methods run fast enough for artists to supervise results. They offer transfer with a higher level of detail, being best suited to the purpose of reusing animation between similar characters (such as different design iterations of a lead character). These can be used when creating animation for side-characters or crowds. When transfer is required between characters that vary dramatically in shape – such as transfer between a human source and a non-human monster or



another fantasy character – semantic transfer can be applied. And, once trained, deep learning methods push the boundary forward for applications requiring automation.

As one specific example, Saito's [16] deformation transfer has been used successfully in the production of a full feature length film. In this case, seminal deformation transfer was extended with Saito's virtual triangles and smoothness constraints and was used to create facial blendshapes for custom characters by transferring poses from a predefined template model. While blendshapes are a standard in industry, the cost of creating additional shapes tailored to each unique character is considerable, and for productions with lower budgets this cost is infeasible. To address this issue, the studio employed deformation transfer to create blendshapes almost automatically. And, with Saito's extensions reducing the occurrence of artifacts, there was little need to fix issues such as invalid creasing near the corners of the lips and eyes.

Ultimately, by considering the demands of the given application and carefully choosing an appropriate variation, we believe that deformation transfer can provide studios with a competitive edge to keep up with the growing demands of animation production.

## 9. Conclusion

Seminal deformation transfer enables artists to copy and paste animation between two characters without first needing to create and map between customized controls for each of those characters. The advantages provided by more recent work address the key issues of efficiency, robustness, and automation.

To conclude this report, we suggest areas of future work that may help further the practical application of deformation transfer.

*Cages without detail loss.* Cage-based approaches [5,7,23] are critical for realizing interactive performance, but risk losing fine-scale deformations. Future work should consider algorithms to adapt cages to best preserve fine-scale details.

*Resolve intersection artifacts.* Saito [16] highlights that transfer results often feature intersections, which can be resolved in part by adding virtual triangles that represent the space between different parts of the mesh. Future work considering broader solutions to resolving intersections would be valuable.

*Artist guidance.* It is critical that artists be able to refine transfer results. A clear and intuitive mechanism for artists to guide the transfer is yet to be proposed.

*Temporal editing.* Domadiya et al. [9] notes that temporal artifacts need to be addressed in transfer results. Future work might consider deformation representations that model both spatial and temporal properties to ensure that transfer faithfully recreates both the shape and the timing of the source.

*Hybrid techniques.* Deep learning techniques [13,21,23] enable fast, robust, and automatic deformation transfer once trained effectively, but their application is generally limited due to the lack of input data for non-human and fantasy characters. Important future work would be to consider a hybrid approach, where traditional deformation transfer techniques are used to create missing data that can then be used to train deep learning techniques.

*Shape matching for better correspondence.* The range of correspondence methods explored for deformation transfer is relatively small in comparison to the variety surveyed in [20]. Valuable future work would be to apply more advanced shape matching solutions for correspondence mapping in deformation transfer. Recent work by Roufosse et al. [15] provides an exciting starting point.

*Decoupling shape and pose deformations.* Recalling Section 6, future work should develop new representations for shape and deformation that can isolate deformations as being unique to either pose or shape. Doing so would further enable artifact-free transfer.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRediT authorship contribution statement

**Richard A. Roberts:** Conceptualization, Methodology, Software, Investigation, Writing - original draft, Writing - review & editing. **Rafael Kuffner dos Anjos:** Conceptualization, Methodology, Writing - review & editing. **Akinobu Maejima:** Conceptualization, Methodology, Writing - review & editing. **Ken Anjyo:** Conceptualization, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

## Acknowledgments

We would like to call special thanks to Ayumi Kimura and other colleagues from the R&D team at OLM Digital for their continuous support. We also thank Junhong Zhao and Andrew Chalmers, and other peers from CMIC, for their discussion and suggestions. Finally, we thank the anonymous reviewers and editor for their many insightful comments. The help from everyone has been invaluable and greatly improved our paper. This project was supported by IMAGICA Group Inc. in Japan and the Entrepreneurial University Program funded by TEC in New Zealand.

## References

- [1] Angelov D, Srinivasan P, Koller D, Thrun S, Rodgers J, Davis J. Scape: shape completion and animation of people. *ACM Trans Graph* 2005;24(3):408–16. doi:10.1145/1073204.1073207.
- [2] Au OK, Tai CL, Liu L, Fu H. Dual Laplacian editing for meshes. *IEEE Trans Vis Comput Graph* 2006;12(3):386–95.
- [3] Balan AO, Sigal L, Black MJ, Davis JE, Haussecker HW. Detailed human shape and pose from images. In: *Proceedings of the 2007 IEEE conference on computer vision and pattern recognition*; 2007. p. 1–8.
- [4] Baran I, Vlasic D, Grinspun E, Popović J. Semantic deformation transfer. *ACM Trans Graph* 2009;28(3). doi:10.1145/1531326.1531342.
- [5] Ben-Chen M, Weber O, Gotsman C. Spatial deformation transfer. In: *Proceedings of the 2009 ACM SIGGRAPH/Eurographics symposium on computer animation*. New York, NY, USA: Association for Computing Machinery; 2009. p. 67–74. ISBN 9781605586106. doi:10.1145/1599470.1599479.
- [6] Bian S, Zheng A, Gao L, Maguire G, Kokke W, Macey J, et al. Fully automatic facial deformation transfer. *Symmetry* 2019;12(1). doi:10.3390/sym12010027.
- [7] Chen L, Huang J, Sun H, Bao H. Cage-based deformation transfer. *Comput Graph* 2010;34(2):107–18. doi:10.1016/j.cag.2010.01.003.
- [8] Deng Z-J, Luo X-N, Miao X-P. Automatic cage building with quadric error metrics. *J Comput Sci Technol* 2011;26(3):538.
- [9] Domadiya PM, Shah D, Mitra S. Guided deformation transfer. In: *Proceedings of the European conference on visual media production*. New York, NY, USA: Association for Computing Machinery; 2019. https://doi.org/10.1145/3359998.3369408. ISBN 9781450370035.
- [10] El Rahman Shabayek A, Aouada D, Saint A, Ottersten B. Deformation transfer of 3d human shapes and poses on manifolds. In: *Proceedings of the 2017 IEEE international conference on image processing (ICIP)*; 2017. p. 220–4.
- [11] Freifeld O, Black MJ. Lie bodies: a manifold representation of 3d human shape. In: Fitzgibbon A, Lazebnik S, Perona P, Sato Y, Schmid C, editors. *Computer Vision – ECCV 2012*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 1–14.
- [12] Gao L, Lai Y, Yang J, Ling-Xiao Z, Xia S, Kobbelt L. Sparse data driven mesh deformation. *IEEE Trans Vis Comput Graph* 2019. doi:10.1109/TVCG.2019.2941200. 1–1
- [13] Gao L, Yang J, Qiao Y-L, Lai Y-K, Rosin PL, Xu W, et al. Automatic unpaired shape deformation transfer. *ACM Trans Graph* 2018;37(6). doi:10.1145/3272127.3275028.
- [14] Le BH, Deng Z. Interactive cage generation for mesh deformation. In: *Proceedings of the twenty-first ACM SIGGRAPH symposium on interactive 3D graphics and games*. New York, NY, USA: Association for Computing Machinery; 2017. https://doi.org/10.1145/3023368.3023369. ISBN 9781450348867.
- [15] Roufosse J, Sharma A, Ovsjanikov M. Unsupervised deep learning for structured shape matching. In: *Proceedings of the 2019 IEEE/CVF international conference on computer vision (ICCV)*; 2019. p. 1617–27.
- [16] Saito J. Smooth contact-aware facial blendshapes transfer. In: *Proceedings of the symposium on digital production*. New York, NY, USA: Association for Computing Machinery; 2013. p. 7–12. ISBN 9781450321365. doi:10.1145/2491832.2491836.

- [17] Sumner RW. Mesh modification using deformation gradients. Massachusetts Institute of Technology; 2005.
- [18] Sumner RW, Popović J. Deformation transfer for triangle meshes. *ACM Trans Graph* 2004;23(3):399–405. doi:[10.1145/1015706.1015736](https://doi.org/10.1145/1015706.1015736).
- [19] Tan Q, Gao L, Lai Y-K, Xia S. Variational autoencoders for deforming 3d mesh models. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2018. p. 5841–50.
- [20] van Kaick O, Zhang H, Hamarneh G, Cohen-Or D. A survey on shape correspondence. *Comput Graph Forum* 2011;30(6):1681–707. doi:[10.1111/j.1467-8659.2011.01884.x](https://doi.org/10.1111/j.1467-8659.2011.01884.x).
- [21] Wang J., Wen C., Fu Y., Lin H., Zou T., Xue X., et al. Neural pose transfer by spatially adaptive instance normalization. 2020. [2003.07254](https://arxiv.org/abs/2003.07254).
- [22] Yang J, Gao L, Lai Y-K, Rosin P, Xia S. Biharmonic deformation transfer with automatic key point selection. *Graph Models* 2018;98. doi:[10.1016/j.gmod.2018.05.003](https://doi.org/10.1016/j.gmod.2018.05.003).
- [23] Yifan W, Aigerman N, Kim VG, Chaudhuri S, Sorkine-Hornung O. Neural cages for detail-preserving 3d deformations. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*; 2020.
- [24] Zayer R, Rössl C, Karni Z, Seidel H-P. Harmonic guidance for surface deformation. *Comput Graph Forum* 2005;24(3):601–9. doi:[10.1111/j.1467-8659.2005.00885.x](https://doi.org/10.1111/j.1467-8659.2005.00885.x).
- [25] Zhao Y, Pan B, Xiao C, Peng Q. Dual-domain deformation transfer for triangular meshes. *Comput Animat Virtual Worlds* 2012;23(3–4):447–56. doi:[10.1002/cav.1466](https://doi.org/10.1002/cav.1466).
- [26] Zhou K, Xu W, Tong Y, Desbrun M. Deformation transfer to multi-component objects. *Comput Graph Forum* 2010. doi:[10.1111/j.1467-8659.2009.01601.x](https://doi.org/10.1111/j.1467-8659.2009.01601.x).