Iterative Cage-based Registration from Multi-view Silhouettes

Yann Savoye*
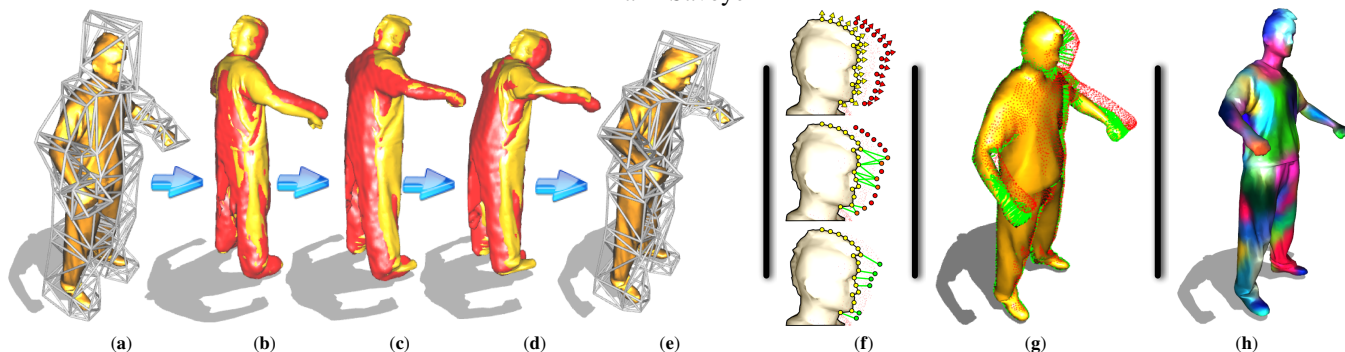

Figure 1: *Dynamic Registration of Life-Like Organic Shapes.* A humanoid-type cage is roughly bound to a laser-scanned template mesh, to perform scalable registration of an observed whole-body motion (**a**). We aim at deforming the source template (in yellow color) toward a noisy target point cloud (in red color) (**b**), reconstructed from sparse captured silhouettes. Our new skin-detached registration procedure pulls cage-based geometry toward estimated target locations (**c**). After multi-iterations, the next geometric target is reconstructed (**d**), and the iterative data-driven deformation is repeated (**e**). Normal-guided pairwise correspondences (in green color) are automatically established (**f**), and pruned to satisfy a normals divergence criteria (**g**). Consequently, handle-aware overlapping biharmonic-rigidities (**h**) offer more controllably to register the complex surface variations (**e**).

Abstract

Consistent dynamic shape capture from casual videos is a fundamental task at the cross-fertilization of Computer Vision and Computer Graphics. Notwithstanding, recent advances in low-cost dynamic scanning turn the cross-parametrization of non-rigid animatable surface into an ill-posed vision-oriented problem. In this paper, we propose a cage-based technique to register non-rigid observed shapes using a meaningful animator-friendly non-rigid reduced embedding. This subspace offers natural silhouette-aware complexity encapsulation and possible reuse of time-varying estimated parameters associated the underlying flexible structure. In particular, we leverage the problem of highly non-rigid spacetime registration by employing an elasto-plastic coarse cage. Thus, we perform scalable handle-aware harmonic shape registration, relying on the higher-level of shape abstraction offered by the space-based paradigm. To the best of our knowledge, our technique is the first to investigate handle-aware elastic overlapping-rigidities for registering life-like dynamic shapes in full-body clothing. Finally, we successfully tested the effectiveness of our proposed solution to a collection of challenging real-world datasets.

CR Categories: I.4.8 [Computer Vision]: Surface fitting—Time-varying imagery I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Physically based modeling;

Keywords: template-based registration, non-rigid ICP, biharmonic rigidity fields, cage-based shape, handle-aware deformation.

1 Introduction

Recent advances in low-cost dynamic scanning turn the cross-parametrization of non-rigid animatable surface into an ill-posed vision-oriented problem. Non-rigid structures in motion are ubiquitous in the world we live in, and their complex behavior respond to the nature's law. Thereupon, acquiring dynamic whole-body live-action via shape registration is a challenging research task that plays a decisive role in various vision-based animation applications, at the crossroad of 4D Computer Vision and Computer Graphics fields.

Cage-Driven Motivation. Numerical geometry of non-rigid shapes has been extensively studied over the years. Nevertheless, registering highly non-rigid dynamic captured surface is a hard and fundamental problem of crucial importance that attracts an ever-increasing interest in recent years. A potential application of non-rigid registration technology featured in the video-clip G-Force of Flo Rida where non-rigid facial multi-view scans leads to an ear-to-ear head alignment. Separately, cage-based methods have already shown their efficiency to drive expressive non-rigid animation of feature-length animated film. In this paper, we propose a new understanding of life-like dynamic captured surface of people in daily-life clothing. We cast the animation reconstruction of laser-scanner mesh template from multi-view images footage as the flexible cage-driven registration of this template across reconstructed geometry information from the input multi-view silhouettes.

Ultimate Goals. Little attention has been paid to the possibility of registering via meaningful reduced set of parameters, abstracting the observed deformations rather than the shape evolution itself. Thus, we aspire to accomplish dynamic shape capture from sparse multi-view videos, by the mean of markerless registration of a static cage-based template via a visual hull prior. Our goal is to estimate optimal time-varying target cage location expressing the unknown deformation of the template under unsupervised correspondences established between the source enclosed model at its current state, and the target visual hull reconstructed from multi-view calibrated silhouettes. Injecting global cross-connectivity in temporally inconsistent is the ultimate goal of our investigation. In our work, we try to estimate optimal cage parameters expressing a plausible deformation with respect to the unknown physical motion observed in the multi-view video stream. To highlight key characteristics of the proposed solution, we catch the reader's attention concerning the fact that we do not employ any user-specified marker or landmark on the template to force point-to-point correspondence. We assume any collection of target cage, any user-specified correspondence or

*ysavoye@siggraph.org

an example-based learning space of deformation as input. User intervention is only required to roughly adjust the geometry of the generic cage connectivity for enveloping the static template mesh, once at the binding step.

Our Contributions. We address the particularly hard and ill-posed inverse problem of non-rigid space-time registration in presence of highly deforming surface wrinkles. Consequently, we propose an incremental cage-based animation reconstruction strategy. We non-rigidly register a given highly-detailed template mesh on time-varying unaligned dense point clouds, encapsulating high resolution details with unknown correspondences. First, we cast the animation-reconstruction of static laser-scanner mesh template from temporally-inconsistent point clouds as unsupervised cage-based registration. Thus, we perform scalable handle-aware biharmonic shape registration, relying on the higher-level of shape abstraction offered by the space-based paradigm. In order to recover an output life-like space-time surface, we propose an elasto-plastic deformation model allowing free-form shape registration iteratively using non-rigid skin-detached subspace, and preserving the local surface variation like cloth wrinkles encapsulated in the captured point clouds. In particular, we leverage the problem of highly non-rigid spacetime registration by employing an elasto-plastic coarse cage as deformable reduced geometric model. The elastic-rigidities are expressed through high-level Laplacian-of-Biharmonics control, defined in term of rigged cage-handles. At the heart of the system, a data-driven re-weighted least-squares formulation allows the production of aligned and highly-detailed boneless meshes.

Algorithm Novelties. Our pipeline enables scalable curve-based surface registration of video-driven whole-body performance in low-dimensional animation-space, loosely decoupled for the shape resolution. More importantly, the output skin-detached curves express highly-detailed template from inconsistent temporal scans with a low-dimensional consistent temporal subspace. Consequently, non-rigid surfaces can be temporally registered via a linear combination of biharmonic basis functions, decoupling the state size from the controlled geometry and dynamic complexity of the mesh. To the best of our knowledge, our technique is the first to investigate handle-aware elastic overlapping-rigidities for registering life-like dynamic shapes in full-body clothing. We demonstrate a first step toward realistic deformation using our algorithm on real-world data of anatomical shape in clothing.

2 Background and Related Works

There is a terrific deal of recent research that is of interest toward cage-based performance capture. In this section, we briefly review few outstanding works that are relevant to the problem of dynamic shape registration for captured videos, covering the following main categories: cage-based embeddings, template-based dynamic capture, animation reconstruction, and space-time non-rigid ICP.

Cage-based Shape Embeddings. Recent years have seen an increasing interest for cage-based deformation: an emerging class of purely geometric space-based techniques, widely used for controlling meshes enclosed in a flexible coarse bounding polytope. The automatic coarse cage-generation for behavior-specific deformation is a difficult problem that has relatively attracted few attention, but having promising solutions using oriented bounding box [Xian et al. 2012] and voxelization [Xian et al. 2009]. By relying on a single generic full-body cage connectivity as single-component cage element for all Experiments, our cage building technique is similar to Ju *et al.* [2008] for bringing more animator-friendly abstraction. Then, a rigging function expressed as generalized barycentric coordinates is associated with the cage-model paradigm to describe realistic and plausible control behavior of the deformation. Various cage-vertex coordinate systems have been developed for space-based manipulations [Ju et al. 2005; Lipman et al. 2007]. Green Coordinates [Lipman et al. 2008] ensures quasi-conformal mapping by integrating cage-face orientation.

Harmonicity for highly-detailed cage-based encoding is introduced with Harmonic Coordinates in [Joshi et al. 2007] defined as the solution of Laplace's equation subject to appropriately chosen boundary condition. Variational harmonic maps of Ben-Chen *et al.* [2009b] are insensitive to cage geometry. Multi-harmonics have recently gained considerable attention and demonstrated better shape-awareness and smoothness [Ofir Weber and Gotsman 2012]. Biharmonic weights described in Jacobson *et al.* [2011] bring sparsity, similarly to the Poisson-based weight reduction proposed in Landerneau *et al.* [2010]. To the best of our knowledge, we are the first to aim at transferring the inherent deformation observed in multi-view video stream into life-like cage-based encoding, following the principle of transferring spatial deformation using cage-based paradigm [Chen et al. 2010b; Ben-Chen et al. 2009a]. Thus, cage-based control lies at the heart of our video-based registration method to pilot dynamic life-like template-based capture.
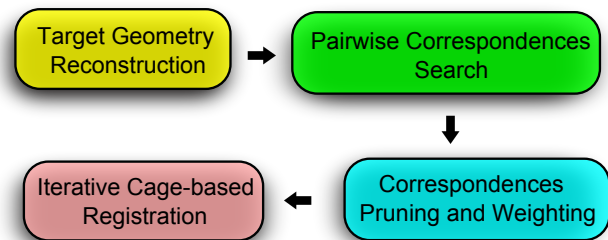

Figure 2: *Overview. Our unsupervised registration technique estimates a collection of meaningful correspondence between the source template and the target point clouds. Resulting data-driven matching drives the cage-based geometry fitting. Our technique allows the registration of sparse-and-detached cage-handle curves, expressing highly-detailed deformation induced by target point clouds.*

Video-based Reconstruction. Since the pioneering work of Laurentini *et al.* [1994] in passive reconstruction from wide-baseline images fusion, various sensor-based techniques have emerged for acquiring silhouette-based point clouds. Firstly, silhouette foreground segmentation as provided by Kim *et al.* [2006] is required for contour-driven reconstruction. Probabilistic approaches [Grauman et al. 2003; Kolev et al. 2012] try to reduce defective segmentation. Secondly, even if irregularly sampled isosurfaces can be obtained by polyhedral algorithm [Matusik et al. 2001; Chen et al. 2010a] without uniform grid carving, we prefer voxel-based polygonization [Montenegro et al. 2006; Sormann et al. 2007] to extract semi-regular geometry without heavy remeshing. Contrary to [Volodine et al. 2007; Guennebaud and Gross 2007] where the ultimate aim is only the surface reconstruction from an input point cloud, we rely on an easy-to-implement silhouette-based watertight inference to derive a series of oriented point clouds. This temporary meshing step is similar to [Hornung and Kobbelt 2006] but relies on sparse set of multi-view of silhouette as input. This visual-hull prior furnishing geometric targets for our registration stage, can be refined by multi-view volumetric stereo [Liu et al. 2010], graph-cut [Vogiatzis et al. 2005] and GPU-accelerated computation [Yous et al. 2007].

Template-based Dynamic Shape Capture. Laplacian-based template processing for point cloud fitting is proposed by Stoll *et al.* [2006] guided by user-specified correspondences, and extended Yeh *et al.* [2010] with an iterative fitting scheme. Since a variety of sensors roughly perform non-intrusive 4D acquisition of life-like surfaces, template-based processing is a key ingredient of De Aguiar *et al.* [2008]. Our goal is also to achieve purely geometric silhouette-consistent performance capture. The pioneering work on surface capture by Starck *et al.* [2007] is extended by Huang *et*

al. [2011] to obtain temporally consistent non-rigid surfaces. However, only few techniques succeed in fitting simultaneously articulated skinned template to sparse multi-view silhouettes with additional silhouette rim vertices correction, as seen in [Vlasic et al. 2008; Ballan and Cortelazzo 2008; Gall et al. 2009] for a single character. Liu *et al.* [2011] exhibit a segmentation-based approach to estimate the pose-and-surface of two-people in the same video.

Space-Time Non-rigid ICP. Various classes of registration approaches for dynamic surfaces have largely benefited from two well-established works in the modern age of Computer Graphics: *Iterative Closest Point* (ICP) algorithm introduced by Besl *et al.* [1992] and *Space-Time Optimization* pioneered by Witkin *et al.* [1988]. The articulated generalization of the ICP for non-rigid spacetime and unsupervised registration has been relatively well-studied as seen in [Corazza et al. 2010; Anguelov et al. 2005; Weilan Luo and Aizawa 2010; Chang and Zwicker 2008] to recover articulated model using the skeleton-guided fitting to point cloud. Even if Zheng *et al.* [2010] only used a curve-skeleton to register captured point cloud, we prefer to abandon the use of piecewise-rigid stick figure for acquiring the 3D surface motion. Contrary to Chang *et al.* [2011], we do not restrict surface deformation to quasi-articulated model. Elastic variants of non-rigid ICP for registering dense deformable model are proposed with coarse to fine strategy [Sagawa et al. 2009], incremental free-form deformation [Abdelmunim and Farag 2011] and elastic orientation-preserving matching [Windheuser et al. 2011]. Space-time surface reconstruction is performed using registration flow in [Sharf et al. 2008; Eckstein et al. 2007]. A formalization for establishing pair-correspondences is attempted in [Connor and Kumar 2010; Low 2004] using the k-nearest neighbor assignment, and fuzzy point correspondences in [Münch et al. 2010].

Animation Reconstruction. Non-rigid ICP-like registration also addresses the problem of space-time reconstruction as seen in Popa *et al.* [2010]. In our work, we seek globally consistent space-time geometry and share the same goal with Wand *et al.* [2007] toward reconstructing a consistent frame sequence with the same connectivity, but not limited to near-isometric deformations. Amberg *et al.* [2007] proposes ICP-like algorithm for surface registration using different regularization but suffering of landmark prior knowledge. Our approach treats non-rigid registration as an optimization problem like Huang *et al.* [2008] but the video-based nature of our method invited us not to assume approximately isometric deformations. We bypass the estimation of affine transformation used in the non-linear registering optimization from a dynamic point cloud sequence of Li *et al.* [2008], Animation Cartography [Tevs et al. 2012] and Animation Transplantation [Süssmuth et al. 2010]. Li *et al.* [2009] also proposes a video-based and template-based surface registration approach using non-rigid ICP. Similarly to temporally coherent completing of dynamic shapes of Li *et al.* [2012], our system relies on the visual hull prior and surface fairing, we prefer to employ a better abstracted underlying subspace than the deformation graph using in [Tong et al. 2012]. In contrast with Li *et al.* [2012], we propose a novel detail-preserving registration approach with resolution-independent control. Finally, we avoid coarse-to-fine surface segmentation at the center of patch-based approaches [Budd et al. 2011; Huang et al. 2011].

The remainder of this paper is organized as follows. Our handle-aware detached registration from multi-view videos is explained in Section 3, with a novel comprehensive formulation. Then, experimental results are detailed in Section 4. Next, the overall procedure is discussed in Section 5. Finally, this paper is concluded, and an outlook to future work is given in Section 6.

3 Handle-Aware Detached Registration

In this section, we describe our skin-detached registration technique build upon silhouette-consistent reconstruction of unrelated series of target point clouds. As the key insight, an iterative elasto-plastic optimizer solves the cage-handles registration by alternating between normal-guided pairwise correspondences computation from target to template and template-based warping. Our non-rigid ICP-like technique is piloted by weight-control to balance between elasticity and plasticity energies. Finally, the registered dynamic surface is resituated according to temporally-registered cage-handle curves and associated space-based deformation.

3.1 Non-Rigid Registration Setup

In our approach, the registration setup is composed of the multi-view full-body recording, and the Laplacian cage-based template describing the underlying reduced deformable model.

Multi-View Full Body Recording. The full-body performance of real-actor is filmed by a synchronized network of only eight fixed and pre-calibrated high-res pinhole cameras, regularly spaced in a chroma-key green room. We assume the continuous observed dynamic scene is represented by a regular sampled sequence $\mathcal{S}_l = \left\{ \mathcal{I}_l^1, \cdots, \mathcal{I}_l^b \right\}$ of foreground-segmented binary silhouette images corresponding to each camera l during b frames. We denote by $\mathcal{I}_l^t$ the binary silhouette image associated to the lth camera at time t. Built upon the estimated 4×3 projection matrix describing the configuration of the lth camera, the projection operator $\Pi_l(\cdot) : \mathbb{R}^3 \mapsto \mathbb{N}^2$ mapping a silhouette pixel coordinates to a given 3D world point included in the camera frustum. Assuming static prior knowledge of a specific shape of interest, we propose to evolve the fixed single-component connectivity $\mathcal{F}$ offered by the given dense, laser-scanned, static template mesh $\mathbb{M} = (\mathcal{V}, \mathcal{F}, \mathcal{G})$ of the corresponding actor to fit the deformation observed in the extracted multi-view silhouette images sequences. The geometry of the closed genus-zero template mesh, composed of n model vertices, is written by $\mathcal{V} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ where $\mathbf{v}_i \in \mathbb{R}^3$ is the location of the ith model vertex. The discrete normal vector field to the template surface is denoted by $\mathcal{G} = \{\tilde{\mathbf{n}}_1, \cdots, \tilde{\mathbf{n}}_n\}$ where $\tilde{\mathbf{n}}_i \in \mathbb{R}^3$ is the normalized surface gradient of the ith model vertex.
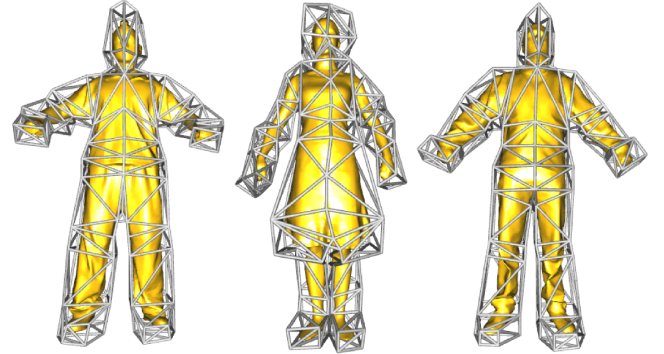

Figure 3: *Reusable Skin-Detached Embedding. A generic humanoid-type cage connectivity performs as-conformal-as possible encoding of whole-body models for scalable editing. This abstraction of skin behavior can be adjusted to a large configuration of the organic topology such as girls in skirt or men in trouser. The same user-specified coarse polytope can bound the space of different laser-scanned body template using a compact set of controllable handles distributed in a shape-aware manner.*

Laplacian Cage-based Shape. In the following, we will refer to $\Omega \subset \mathbb{R}^3$ as the bounded polyhedral domain included by m control cage-handles enveloping the motion-less template mesh. The set of cage-handles is denoted by $C = \{\mathbf{c}_1, \cdots, \mathbf{c}_m\}$ where $\mathbf{c}_k$ is the current location of the kth cage-handle in global coordinates system (Fig.3). Our input cage polytope structure is augmented by the

Discrete Laplace-Beltrami Operator $\mathcal{L}(\cdot)$ with non-uniform cotangent weights to enforce as-low-as possible distortion of the irregular cage connectivity [Desbrun et al. 1999]. In addition, associated cage differential δ-coordinates encode each cage-handle relatively to its neighborhood in the enclosing cage connectivity [Botsch and Sorkine 2008]. The geodesic-aware volumetric relationship between the cage subspace and the static template is encapsulated by bi-harmonic rigging, performed at the default pose. All of these components help to achieve scalable template mesh registration.

3.2 Target Point Cloud Reconstruction

Watertight Space Carving. We assume the observed scene is filmed using a synchronized multi-camera network converging toward the observed surface in motion. The single real-actor is emerged in a controlled environment in order to facilitate the extraction of silhouette images thanks to the chroma-key background. We reconstruct a geometry entity using voxel-based shape-from-silhouette and watertight surface extraction process. Even if the accuracy of the carved surface is dependent of the volumetric resolution, we obtain a photo-consistent manifold mesh by starting from a uniform 128^3 voxel grid. We define the cone $\mathcal{H}_l$ generated by the silhouette $\mathcal{I}_l^t$ as the set of 3D points that fully reproject inside the given silhouette. The reconstructed visual hull $\mathbb{V}^t$ defined by the silhouette set can be written as the following cones intersection: $\mathbb{V}^t = \bigcap_l \mathcal{H}_l = \left\{ x \in \mathbb{R}^3 : \forall l, \Pi_l(x) \in \mathcal{I}_l^t \right\}$. A spacetime isosurface $\mathbb{S}^t = \left\{ (x) \in \mathbb{R}^3, f(x,t) = 0 \right\}$ can be extracted using a polygonization process of the boundary voxel crust. As a first step in surface extraction, we polygonize the boundary of connex boolean voxels that facing toward the exterior to generate the desired winged-edge mesh connectivity and reconstructed topology. Consequently, the final optimized output surface is then crack-free watertight.
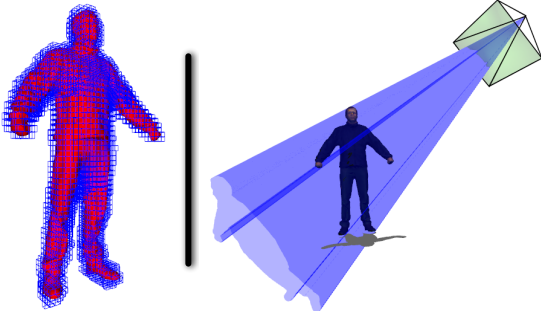
Figure 5: *Voxel Carving and Visual Cone. Space-carving driven by the silhouette fusion is compared to the static input laser scanned on the first frame (**left hand side**). According a given calibrated camera and its foreground-segmented silhouette, the visual cone is traced thought the silhouette bins contour. The silhouette defines a back-projected generalized cone that contains the actual object (**right hand side**).*

Silhouette-Aware Oriented Point Clouds. In a second step, we optimize the surface fairness to obtain a pleasant smooth silhouette-aware geometry by attenuating the grid artifact. Then, the extracted surface is denoised by applying an iterative Laplacian smoothing where each reconstructed surface vertex $\mathbf{p}_j$ are moved using the following operator $\mathbf{p}_j \leftarrow \mathbf{p}_j - \Delta^2(\mathbf{p}_j)/d$ where d is the valence of $\mathbf{p}_j$. Then, normals associated to each point of the cloud can be easily estimated as the surface gradient. As far as all reconstructed targets are temporally connectivity-inconsistent, the resulting connectivity is without of interest for the rest of our algorithm. Next, we derive an oriented target point cloud from the optimized reconstructed surface at each frame. Our system does not exploit the target connectivity or any surface flow assumption. As time-varying geometric output, we obtain a sampled sequence of l temporally-inconsistent but oriented point clouds $\mathcal{P} = \{\mathcal{P}_1, \cdots, \mathcal{P}_b\}$ that are partially overlapping. The point cloud $\mathcal{P}_t = \left\{ \left(\mathbf{p}_j, \tilde{\mathbf{m}}_j \right) \right\}$ defines a set of 3D points $\mathbf{p}_j \in \mathbb{R}^3$ at the time step t with corresponding normals $\tilde{\mathbf{m}}_j \in \mathbb{R}^3$. If more accurate surface

is required, it also possible to refined the surface using multi-view stereo. Given the fact our technique bets on an intermediate reconstructed geometric entity, our the following registration process is also potentially effective for low-cost time-of-flight or structured light sensors. This algorithm stage converted calibrated silhouettes into a reliable 3D geometric scalar field with its orientation field.
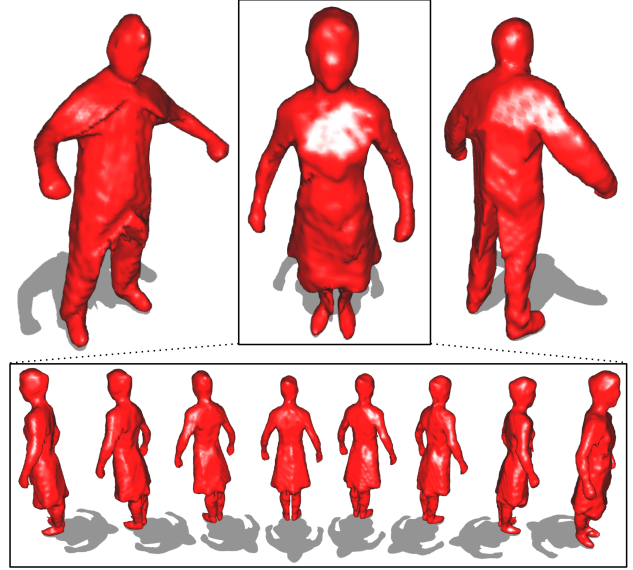
Figure 6: *Carved Visual Hulls. In the perspective of relying on a geometric target prior, a smoothed watertight surface is reconstructed as the isosurface offered by the fusion of silhouette maps (**top row**). Free-viewpoint rendering of the reconstructed shape for a single frame (**bottom row**).*

3.3 Normal-Guided Pairwise Correspondences

Pairwise Correspondences. Since the correspondence information is unknown, the problem demands an alignment algorithm. We leverage the problem of temporal matching by the mean of the *Iterative Closest Points* paradigm. To determine reliable data-driven temporal correspondences with no prior knowledge, candidate geometric feature correspondences are established automatically by approximating k-nearest neighbors queries for all given targets points [Arya et al. 1998]. We use a similar fuzzy-yet-robust geometric strategy than Budd *et al.* [2010] to infer a minimal set of compatible feature correspondences updated at each intra-frame iteration. We determine a suitable heuristically-optimal k-nearest neighbor to establish unsupervised correspondences between the source template and target point cloud. Our strategies significantly improve the data-driven fitting.

We preprocess the template in its current pose into a Bounding Volume Hierarchy structure [Wald et al. 2007] that optimizes the nearest neighbor search via a space partitioning. Then, the nearest-neighbor queries descend into the tree faster than in kd-tree by minimizing root-to-leaf traversal time. More importantly, we prefer to use the BVH for approximating k-nearest neighbors since this structure can be rebuilt every frame faster than classical kd-tree to handle changing template geometry. We opt for a top-down construction of the BVH, following ideas of Wald *et al.* [2007]. Since we construct soft geometric constraints in order to be injected inside an iterative subspace solver, approximating non-exact nearest neighbor search is well-suited for the tackled optimization strategy.

Normal-Guided Pruning. In a first *target-to-source* pass (Fig.8a), unsupervised geometric correspondences are established from target-to-source location because the target is invariant during intra-frame iteration. Consequently, for each $\mathbf{p}_t \in \mathcal{P}$ of the point cloud we compute a limited set of nearest neighbor candidates. For
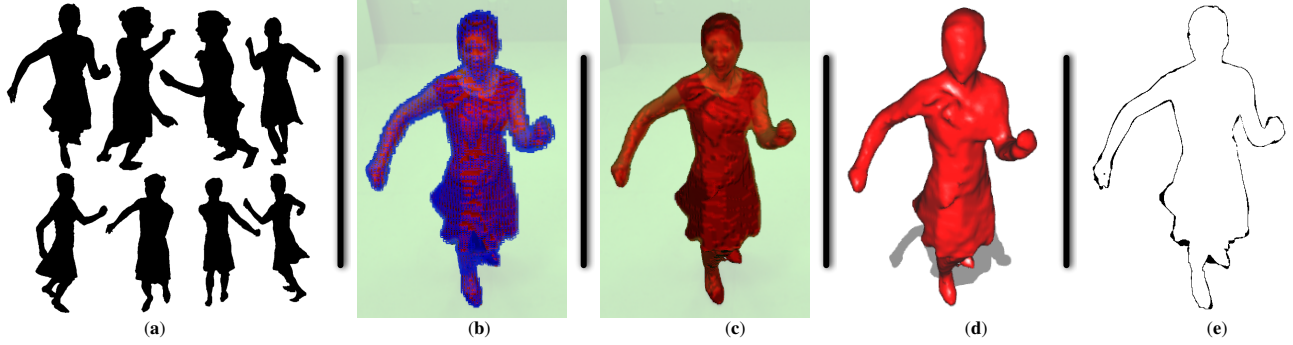
Figure 4: *Video-based Reconstruction of Silhouette-Aware Geometric Targets. Extraction of multi-view silhouette as binary maps (**a**). Voxel-based shape obtained from sparse set of 8 converging multi-view silhouettes (**b**). Watertight surface reconstruction (**c**) and surface optimization (**d**) reprojected in casual images. We compare the silhouette-awareness of the reconstruction by applying the XOR operator between the recorded silhouette map and the silhouette image rendered from the visual hull. This silhouette overlap metric demonstrates an acceptable pixel discrepancy (**e**).*

the selection of pair, we use a nearest-neighbor approximation using BVH structure.

$$J = \left\{ \left(\mathbf{v}_i, \mathbf{p}_j \right) \ : \ \forall \mathbf{p}_j \in \mathcal{P}_t \ ; \ \exists \mathbf{v}_i \in Knn\left(\mathbf{p}_j \right) \right\}$$

where $Knn(\cdot) \subset \mathcal{V}$ is the k-nearest neighbors subset of the source template for a given target point query. The resulting set J is composed of collection of compatible pair $\left(\mathbf{v}_i, \mathbf{p}_j \right)$ verifying the following normal divergence criterion. Conflicting outliers correspondences are eliminated by a normal-guided pruning routine. The compatible correspondences must satisfy a cost measuring how much the smoothed normal divergence lies under a given threshold:

$$\tilde{\mathbf{n}}_i \cdot \tilde{\mathbf{m}}_j < \cos(45°)$$

where $\tilde{\mathbf{m}}_j$ is the smooth normal direction of a given target point, and $\tilde{\mathbf{n}}_i$ the smooth normal direction of a given source template vertex. Each pair $(\mathbf{v}_i, \mathbf{p}_j)$ are locally weighted according to the following weighting function measuring the divergence in terms of distance and normals:

$$\chi_j(i) = \frac{\max(\tilde{\mathbf{n}}_i \cdot \tilde{\mathbf{m}}_j, 0)}{d(\mathbf{v}_i, \mathbf{p}_j)^2}$$

where $d(\cdot)$ is the Euclidean distance. The numerator is the clamped dot product of source and target vertex normals. The denominator is the squared Euclidean distance of the source and target vertices, acting as the inverse distance weighting.
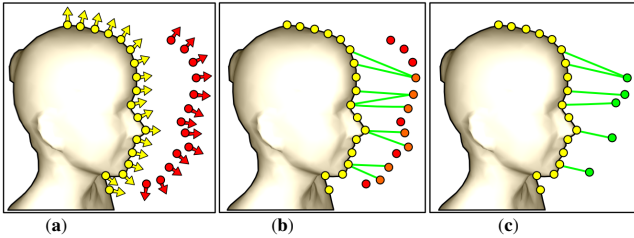

Figure 8: *ICP-like Correspondences. Unsupervised geometric correspondences are established from target to source location, exploiting normal and Euclidean information (**a**). Reliable data-driven geometric temporal correspondences are established automatically with no prior knowledge by approximating k-nearest neighbors queries for all given targets points (example shown only for correspondences for orange colored target) (**b**) Conflicting outliers correspondences are rejected by a normal-guided pruning mechanism to infer a minimal set of compatible feature correspondences (**c**).*

In a second *source-to-target* pass (Fig.8c), we rearrange the correspondences to construct the ultimate compact set $\mathcal{S}$. To avoid constraints conflicts, a unique target location $\mathbf{q}_k \in \mathbb{R}^3$, for the kth current template vertex, is obtained by averaging compatible target points in the current point cloud of $\mathcal{P}_t$ to proceed, and the resulting pairwise correspondence $s_k \in \mathcal{S}$ is constructed. In more details, if the kth template mesh vertices is softly constrained by a set of compatible target points (at least one), its final target location $\mathbf{q}_k$ for a given constrained template mesh vertices is defined as the weight centroid of compatible targets. The unknown desired target

location is formulated as the centroid of weighted k-nearest multi-variate neighbors for a given query. We weight the contribution of all neighbors using a common inverse Euclidean distance function, multiplied by a clamped dot product of pairwise normals.

For each constrainsted vertex $\mathbf{v}_i$ appearing at least in one pair of J, the corresponding target location $\mathbf{q}_i \in \mathbb{R}^3$ (and associated normal $\tilde{\mathbf{r}}_i \in \mathbb{R}^3$) is the weighted average of compatible target point, obtained by the standard Shepard interpolation procedure:

$$\mathbf{q}_i = \frac{\sum_{j \in J} \chi_j(i) \cdot \mathbf{p}_j}{\sum_{j \in J} \chi_j(i)} \quad \text{and} \quad \tilde{\mathbf{r}}_i = \frac{\sum_{j \in J} \chi_j(i) \cdot \tilde{\mathbf{m}}_j}{\sum_{j \in J} \chi_j(i)}$$

After this outliers pruning and aggregation stage, each as-colinear-as possible oriented points correspondence is locally weighted by $\gamma_k = \max(\tilde{\mathbf{n}}_k \cdot \tilde{\mathbf{r}}_k, 0)$ defining the clamped dot product of smoothed pairwise normals between input source and generated target location. The final set of reliable locally-weighted correspondences $\mathcal{S}$ is denoted by:

$$\mathcal{S} = \{s_k : (\mathbf{v}_k, \mathbf{q}_k, \gamma_k)\}$$

3.4 Iterative Elasto-Plastic Optimization

Iterative Subspace Solver. For each noisy point cloud, our cage-handle curve registration process alternates between deformation optimization and correspondences successively, in order to approximate surface motions subtle by clothing. At each intra-frame iteration $u+$, we initialize the cage geometry with location $\mathbf{c}^u$ resulting of the previous iteration u. Then, we update $\mathcal{L}(\cdot)$, the corresponding δ-coordinates and we infer $\mathcal{S}$. Finally, driven by the correspondences propagated in the subspace, a new per-iteration configuration $\mathbf{c}^{u+}$ is estimated by solving the following the global quadratic variational objective function:

$$\underset{\{\mathbf{c}_1^{u+}, \cdots, \mathbf{c}_m^{u+}\}}{\operatorname{argmin}} \left\{ \underbrace{\alpha \cdot \sum_{j=1}^{m} \left\| \mathcal{L}^u\left(\mathbf{c}_j^{u+}\right) - \delta_j^u \right\|_2^2}_{\text{Higher-level Laplacian}} + \underbrace{\beta \cdot \sum_{s_k \in \mathcal{S}^u} \gamma_k \cdot \left\| \mathbf{q}_k - \sum_{j=1}^{m} w_{kj} \cdot \mathbf{c}_j^{u+} \right\|_2^2}_{\text{ICP-like Target Constraints}} \right\}$$

where $w_{kj} : \Omega \mapsto \mathbb{R}^+$ is the biharmonic weight associated to a given cage-handle j with respect to the kth template vertex. Our biharmonic coordinates computation follows core technical ideas proposed in [Jacobson et al. 2011]. Our rigging weights are the unique solution, obtained by finite-element approximation to the fourth-order elliptic biharmonic equation with Dirichlet boundary conditions. For the sake of simplicity, we express the diffusing bi-Laplacian kernel $\nabla^4 \equiv \nabla^2\nabla^2$ onto the volumetric cells discretization of the cage interior domain, at the binding step only:

$$E(w_{kj}) = \int_{\Omega} \left\| \nabla^4 w_{kj} \right\|^2 dx$$

The cage-Laplacian regularization ensures the system to be overdetermined and enforces temporally smoothness prior by injecting
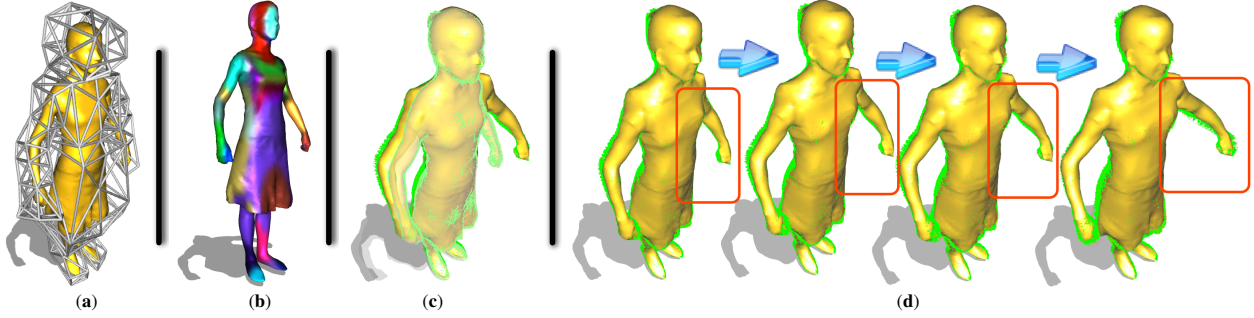
Figure 7: *Handle-Aware Rigidities for Detail-Preserving Registration. An input generic cage is roughly arranged in order to fairly enclose a laser-scanner template having a complex skirt topology (**a**). Cage-handles with biharmonic fields produce a sense handle-aware overlapping control over the surface. Each colored submesh corresponds to the proximity-rigidity influence (**b**). The starting pose of the skeleton-free template is superposed with its registered configuration after few movements of samba dance (**c**). Our iterative elasto-plastic registration optimization preserves the wrinkling details while allowing the template evolution to reach large deformation roughly, such as arm waving motion, framed in the orange squares (**d**).*

weak penalty assumption. At each iteration, the vertex positions of enclosed shape are updated by a dense surface deformation field resulting of cage-based deformation with new cage parameters. Besides, we restitute the registered cage-based shape geometry $\widetilde{\mathcal{V}} = \left\{ \widetilde{\mathbf{v}}_1^{u+}, \cdots, \widetilde{\mathbf{v}}_n^{u+} \right\}$ directly by linear blending of estimated cage parameters, via the following linear non-rigid skinning operator:

$$\forall i \in [1, n]; \quad \widetilde{\mathbf{v}}_i^{u+} = \sum_{j=1}^{m} w_{ij} \cdot \mathbf{c}_j^{u+}$$

Even if the convergence is reached generally after less than one hundred of iterations, we fix a maximum number of iteration to stop the iterative process just in case the surface evolution does not lie under a specified *Root Mean Square* tolerance. The general procedure for a complete frame can be seen as solving a non-linear system with an inexact Gauss-Newton method. During successive deformations, the cage connectivity remains unchanged.
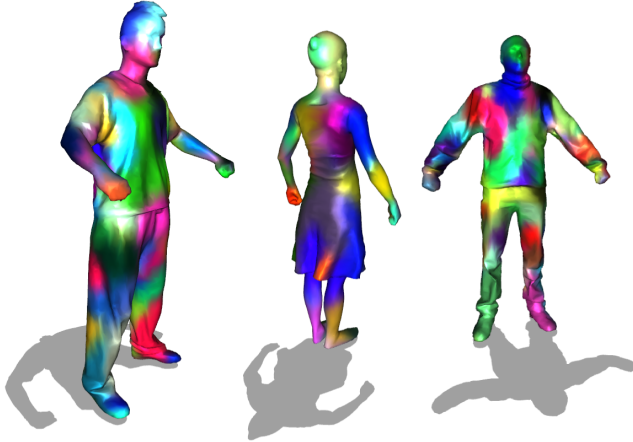

Figure 9: *Handle-Aware Local Flexural Rigidities. Cage-handles with biharmonic mapping convey a sense of handle-aware soft-rigidities influence over the surface. Thus, biharmonic fields produce a local overlapping control. Each colored sub-mesh corresponds to a region of interest where the influence is highest for a given handle. The proximity-rigidity influence smoothly decays non-linearly from the center of implicit patch center to the farthest vertex on the whole mesh, if responding to the unbounded heat-like diffusion.*

Sparse Matrix Form. Our well-conditioned Poisson-like minimization problem can be mathematically solved efficiently and simply using a global re-weighted least-squares formulation, in a cascading fashion. So, at each iteration we rewrite this system into its succeeding over-determined matrix-form expression: $\mathbf{WA} \cdot \mathbf{X}' = \mathbf{WB}$, where $\mathbf{X}'$ is a $3m$-dimensional unknown objective vector. The left-hand side matrix $\mathbf{A}$ is constructed by stacking the Laplacian elements of the cage and involved biharmonic weights. The $3m$-dimensional right-hand side $\mathbf{B}$ is formed by stacking target location

of all correspondences and the differential coordinates of the cage. The diagonal matrix $\mathbf{W}$ compiles the weights α, β and γ. Then the normal equation can also be expressed by $\mathbf{A}^T \mathbf{WA} \cdot \mathbf{X}' = \mathbf{A}^T \mathbf{WB}$. Thus, the global location of cage-handles can be therefore found by solving the following classical closed-form expression in inter-active time: $\mathbf{X}' = ((\mathbf{WA})^T \mathbf{WA})^{-1} (\mathbf{WA})^T \mathbf{WB}$. Note that the most time-consuming part of this solving process will be in the computation of the square matrix $(\mathbf{WA})^T \mathbf{WA})^{-1}$. In practice, we estimate the optimal least-squares cage location $\mathbf{X}' = \left\{ \mathbf{c}_1^{u+}, \cdots, \mathbf{c}_m^{u+} \right\}$ with off-the-shell Conjugate Gradients algorithm to avoid forming the normal equations and to exploit sparsity of the left-hand side. Then, the optimal solution is attained by a steepest-descent black-box, thought walking in the direction of the gradient.

Biharmonic Overlapping Fields Inspiring by the illuminating ideas of [Au et al. 2007], we understand that the biharmonic coordinates defined a set of per-handle fields that overlap over the enclosed mesh with meaningful properties. In particular, the so-expected multi-rings local rigidity constraint in non-rigid ICP problem is here implicitly leverage by the combined rigidity field that could be written as $\Psi = \sqrt{\sum_i \|\nabla \varphi_i\|^2}$ for free. This formula corresponds to the square root sum of the squared gradient magnitude of biharmonic fields. We denote by φ_i the biharmonic fields associated to each i cage-handle. This rigidity field can be reused further for shape understanding and shape analysis. Assigns heavy rigidity influences to region that are spatially close to the handles but decays geodesically far from the handles. This properties is fundamental in our approach to maintain local rigid motion of the surface during registration. Directly related to *Biharmonic Overlapping Fields*, the cage-Laplacian energy term allows the actual combined rigidity fields to stay smooth during inverse modeling, but most important it brings non-uniforming flexibility over the enclosed mesh, needed for the surface registration with large stretching, and that seems impossible to realize previously with other kind of articulated-based ICP techniques.

Output Space-Time Surface. The output of our algorithm is a collection of temporal cage-handle curves, allowing the real-time reconstruction of time-varying mesh, sharing the same fixed connectivity derived from the original input genus-zero template. In other words, the overall algorithm output is a low-dimensional non-rigid surface motion signal $s(t) = (\mathbf{c}_1^t, \cdots, \mathbf{c}_m^t)$ describing the resulting shape in motion into the biharmonic projective subspace basis where $\mathbf{c}_j^t \in \Omega$ is a 3-vector indicating the location of the jth cage-handle in global Cartesian coordinates system at the time step t. Storing the resulting sparse-and-detached spacetime curves with the pre-computed biharmonic coordinates are sufficient to generate a regular sampled sequence of non-rigid temporally consistent 4D
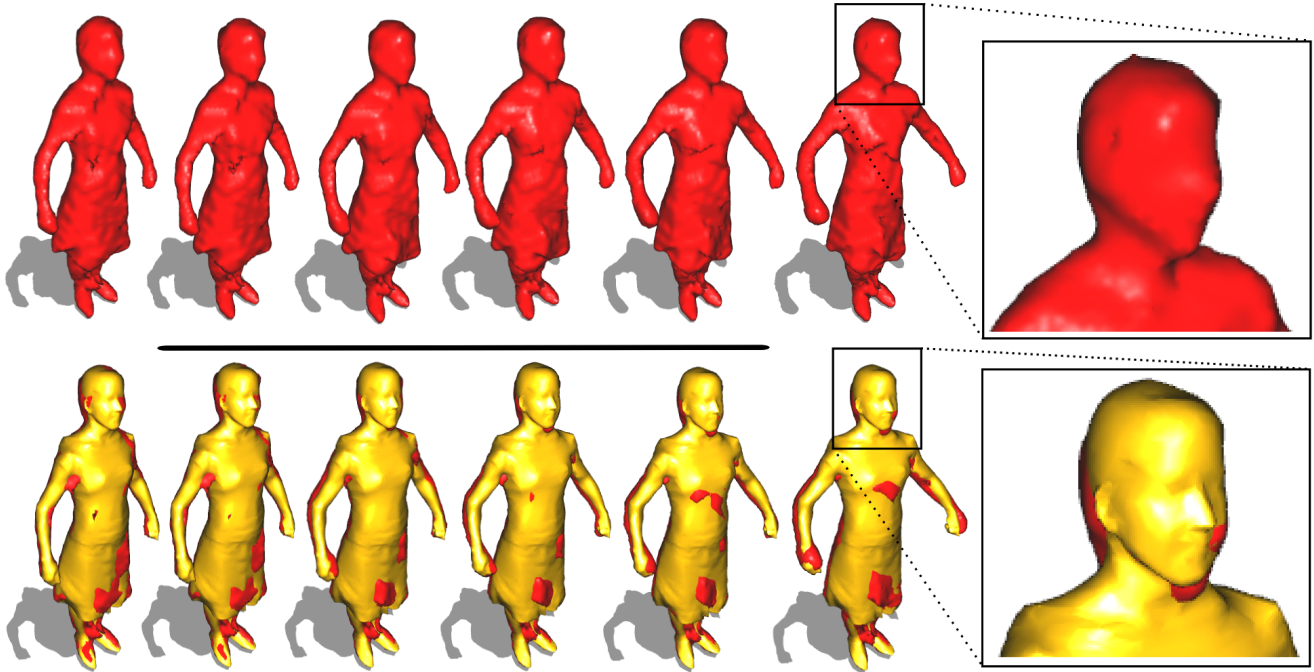
Figure 10: *Small-scale Detail-Preserving. The temporal sequence of unregistered reconstructed meshes from sparse silhouettes is shown in red color (**top row**). Surimposed in yellow color, the cage-based surface fitting is driven by unsupervised temporal correspondences automatically established with the reconstructed targets (**bottom row**). Our technique enforces sharp shape details, corresponding to the nose and eyes of the input template, over topology inconsistency of targets as shown in the close-up frames.*

space-time surface meshes $\mathcal{A} = \{\mathcal{M}_0(F, V_0), \cdots, \mathcal{M}_t(F, V_t)\}$ on-the-fly. The so-expected consistent global connectivity F offered by the input laser-scanned template and V the estimated per-frame cage-based geometry.

3.5 Weight-Control Update Rules

Decomposing the non-linear registration into successive infinitesimal linear deformation steps yields a careful balance of the energy terms using an automatic weighting system. Our heuristic rules for adjusting weight controls are motivated by the need of stronger relaxation of the shape prior across iteration for higher flexibility and monotone convergence to a local minimum. This relaxation is plausible because the confidence in correspondences increase, and convergence is closer along iterations. As seen in Fig.11, the close-form formulation of our rules are obtained by experiments and provided as follows. Thus, the data-term weight-control is initialized at $\beta = 0.01$ at each new frame and increases along intra-frame iterations by following an exponential geometric growth as parametric update-rule to promote the constraint-guided energy:

$$\beta(u) = e^{0.01 \cdot u} - 1.0$$

where 0.01 is the growth rate and u is the current iteration number. Concurrently, the weight-control α enforcing the global shape prior is setup to 1.0 at each frame and the importance slightly decreased across intra-frame iteration to gradually relax the deformation stiffness prior, following an exponential decay to penalize the shape resistance behavior:

$$\alpha(u) = 0.99 \cdot e^{-0.001 \cdot u}$$

where -0.001 is the decay constant, and u is the current iteration number. The value of α is clamped between 0 and 0.8, and the one for β between 1.0 and 0.85 to maintain a coherent balance between energy terms by keeping the shape prior always superior to the data term. We remark that the convergence speed and fitting accuracy are closely dependent of these weight-control update rules.

4 Experimental Results

In this section, we briefly describe our promising results and we discuss related observations. We have implemented a stand-alone prototype to demonstrate the robustness and usefulness of our techniques. All of our experiments are executed on a workstation machine with 4Go of memory and a GeForce 8800 GTX graphic card. We train our algorithm by aligning several real-world datasets of [Vlasic et al. 2008] as shown on Fig.10-13. In other to assess the final results, the accompanying video illustrates the capabilities of the techniques to produce automatic animator-friendly registration. Figure 1 illustrates the behavior of our techniques. In all of our experiments, we use the same generic humanoid-type cage of abstraction, represented as a user-specified closed 2-manifold mesh.
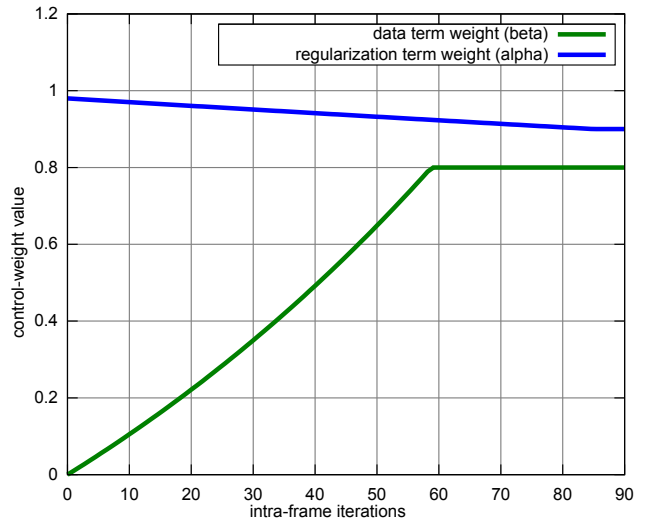

Figure 11: *Weight-Control Update Rules. The cage stiffness is relaxed and the deformation prior is enforced across iterations. The evolution of α-weight is depicted in green color, and the evolution of β-weight across iterations is depicted in blue color.*

The generic cage structure composed of approximatively 141 cage-handles is roughly adapted around the input template by a non-skilled artist in a shrink-wrapping fashion. The enclosed static tem-

plate model is reconstructed using a whole-body laser scanner and sampled into 10002 vertices with 20K triangular faces. To evaluate the computational process, we measure the timing statistic of each step in the processing cycle. At each iteration, the target point cloud reconstruction takes 20*sec.*, optimizations included. Output reconstructed point clouds has a approximately between 17K vertices and 30K vertices. The system spends 126*msec.* to match the source template with the target point cloud, and 9.2*sec.* to prune and rearrange correspondences (for a total average of 4000 final correspondences at the first iteration, according to our pruning strategy). We fix the number of source vertex candidates to 10 for the k-nearest neighbors search for each points in the target. The separable solver process takes 9*sec.* to retrieve cage-handles locations and a constant time of 38*msec.* is required to generate the corresponding enclosed mesh. Our algorithm spends the vast majority of processing efforts in the non-linear optimization process. Even if the results are yet encouraging with few intra-frame iterations and non-optimal adjustments, much more iterations and finest weight-control are required for accurate results. We would like to point out the importance to have a good weighting system to drive a robust deformation. Finally, comparison with other similar skeleton-free techniques is not directly feasible because implementation are not publicly available and quite complex to reproduce.

5 Discussion

In this section, we discuss the intriguing properties of the proposed low-dimensional registration technique. In particular the benefits of reusable registered curves are detailed. Finally, current inherent limitations are clearly uncovered.

Abstraction and Generality. The shape-aware abstraction coupled with near-conformal rigging has sufficient flexibility to potentially drive the scalable fitting of versatile geometry template from multimodal sensors. In particular, our subspace-based registration is not limited to single-component template thanks to its detached form and allows the registration of versatile geometric input templates such polygon soup, degenerated geometries such as uncleaned non-manifold or disconnect multi-component template. The intermediate markerless visual hull reconstruction is well-adapted for texture-less surface. In addition, our technique is well-adapted to other geometric fusion of sensor information. Building animation-friendly registration framework on top of the low-dimensional cage-time structure avoids heavy downsampling or processing of dense point clouds.

Low-Dimensional Registration. The cage parametrization is a well-suited low-dimensional subspace for capturing the latent space of human activities with cloth dynamics. Our model reduction can deal with the registration of dynamical effects in linear low-dimensional models. Furthermore, the purely geometric tessellation of cage allows as-conformal-as possible low-rank data reduction, and this is at the main importance to skin target point clouds. Rather than focusing exclusively on surface registration, we pay also attention to the registration of template-independent sparse temporal curve to compress the data in as-conformal-as possible manner. Building on previous observations of [Au et al. 2007], the Laplacian relationship of cage-handles mixed with their biharmonic influences convey a meaningful sense of handle-aware Laplacian-of-Biharmonics scalar fields, controlling the local elastic-rigidity properties over the enclosed surface. Then, the registered surface is restricted to evolve along the geodesics imposed by the space of natural ambient warping.

Elasticity vs. Plasticity. The kinematic-free cage connectivity offers sufficient degrees of freedom, and the associated rigging functions maintain a controllable proximity influence. On top of that, the Laplacian operator on the cage acts as a linear elasticity tensor between cage-handles, and consequently portrays the global elasticity of overlapping biharmonic deformation fields. Meanwhile, cage-handles enriched with biharmonic fields impose meaningful overlapping elasto-rigid influence control and enforce the model registration to be locally smooth and softly rigid along the geodesics, without prior of strong piecewise rigidity. We refer to the theory of plasticity to invoke the location change of the underlying animatable structure. In few words, the plasticity is the characteristic of a given surface to be deformed definitively under external flexural stress. The balance between the linear elasticity and the flexural tensor is adjusted by control-weights controlling the modulus of elasticity and rigidity.

Benefits of Reusable Registered Curves. Our ICP-like subspace optimization offers a controllable, meaningful, and animator-friendly tool in order to acquire temporally consistent surface with an unsupervised spacetime correspondences strategy. The non-rigid nature of the parameterization allows volume changes of the template, necessarily to approximate surface motion and volume variations subtle by clothing. Besides, our iterative updating framework naturally overcomes the problem of large translation and rotation problem, without the use of user-specified feature correspondences. Since the cage-based paradigm is about ambient space deformation, our registration technique is versatile and can be employed to enclose disconnected polygons soup, particles or mesh-less points as source template. Therefore, the estimated cage configuration can be reused by artists to perform scalable shape manipulation over the registered shape, shape interpolation or other type of cage-based transfer applications. The main benefit of our framework remains in its iterative form, preserving the life-likeness of captured data.
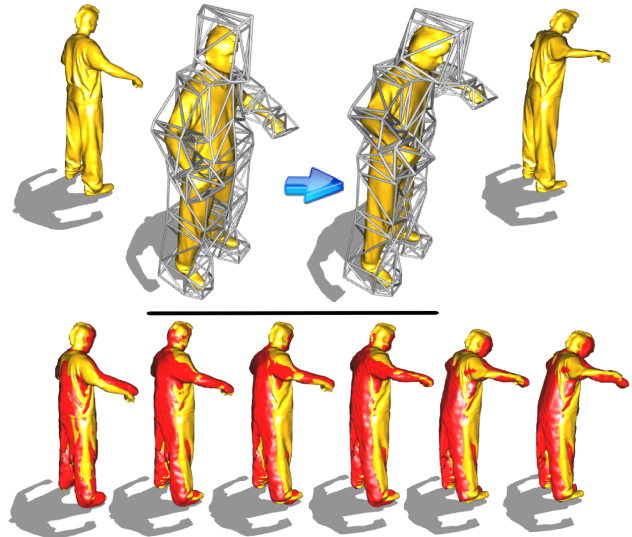

Figure 12: *Detail-Preserving Registration. Our registration process co-registers an enjailed surface and associated cage-handles trajectories (**top row**) over reconstructed noisy point clouds while retaining baked-in template surface details (**bottom row**). The template-based priors is fundamental to overcome ghost-phantom geometry features.*

Current Limitations. The inherent limitation of cage-based deformation is the well-known burden to arrange cage-handles one-by-one in a shape-aware manner. This adaptation is realized at the default pose without strong prior knowledge of the motion or skilled artists. Furthermore, it is worth to mention the time-consuming process with memory overhead related to the biharmonic coordinates computation. Actually, this detached underlying parametrization may fail under drastic topological changes. More severely, our approach is subject to other limitations such as non-scaling invariance under extra-large motion.

At the bind pose, the need to setup a well normal-offseted cage for the enclosed mesh is hard to design because geometry interferences often occur during the cage adaptation, and may affect the computation of rigging weights. Our method is subject to error accumulation under remaining unfiltered correspondence outliers and may provoke failures over long sequence. Even if we combat the inherent problem of ICP to get fall into local minima, our cage-based registration can conduct to suboptimal results and over-fitting. During our experiments, we notice the difficulty to precisely control the surface bending when the embedding is too coarse on highly-deformed zones. The extremely dimension reduction implied by the abstraction of the meaningful cage may lead to a solution that does not necessarily correspond to a fair registration in the presence of locally small deformations. Finally, our template-based approach does not handle changes in topology in that current state. As a *chicken-and-egg* situation, the underlying cage and input template present the disadvantage of having a fixed topology.

6 Conclusions and Future Work

We introduce a new class of approach called cage-based performance capture, offering a simultaneous registration of surface and cage capture. In this paper, we proposed the first approach to turn markerless multi-view video into a high-fidelity dynamic cage-based surface that evolve with cross-connectivity along time. Our reduced deformable model restricts the dynamic geometric processing to the subspace domain and allows physically plausible deformation. The registration benefits from an iterative feedback between revised correspondences inference and cage-based surface fitting.

Registration Highlights. Our approach is a first insightful step toward the hard problem of automatic template-based registration for highly non-rigid dynamic shape using low-dimensional cage-based encoding. At the first sight, the low-dimensional nature of the underlying parameterization allows us to simultaneous register and compress unstructured input animated geometry. On second thoughts, the implicit Laplacian-of-Biharmonics shape control prevents triangle fold-over or the candy-wrapper effects during the registration process under unsupervised correspondences. Similarly, the main advantage of our iterative optimization remains in the simultaneous cross-reconstruction of dynamic shape and skin-detached registration of reusable curves expressing the clothed-body deformations. Thus, our system registers non-rigid shape variations while preserves the life-likeness of captured data and acquires spatiotemporal consistent surface parameters with an unsupervised correspondence strategy.

Future Work. Registering cage-based surfaces opens new direction in the outstanding area of video-based mesh tracking, by allowing the use of non-rigid underlying structures able to track articulated-free animatable surface. Thus, we will focus to realize fine-tuned non-rigid registration for multi-view shape completion. Additionally, we hope to develop tailored ICP-like for determining registration over longer sequences. Other compelling future work would be to considerate the scenario of incomplete, noisy point clouds or versatile targets inputs. We wish to generalize our point clouds fitting technique to a generic human template that necessarily does not correspond to the observed shape. Several avenues remain for longer-term future work. At the bottom, we hope to extend our current pipeline to a templateless cage-free registration approach at video-rate from freely-moving cameras, filming multi-characters interacting in the scene. Toward this direction, we are interested in the temporally coherent completion of life-like dynamic shapes without orientation field estimation or intermediate geometric reconstruction from casual images.

Subsequently, the proposed scheme invites us to pursue ongoing efforts in designing accurate numerical objective function to converge with less iterations. Rethinking the registration phase by incorporating photometric cues or different kind of linear constraints will improve the data-driven fitting. A more careful study of confidence weights and comparison with other cage-based coordinates could also be intriguing to run. Ultimately, a major challenge for years to come in this area is to propose an effective non-iterative procedure for off-hours cage-based recording of topology-evolving surface. Assuming any knowledge of a template offers the possibility to introduce an animator-friendly surface registration that undergoes large deformations with drastic topological changes.

7 Acknowledgment

References

ABDELMUNIM, H., AND FARAG, A. 2011. Elastic shape registration using an incremental free form deformation approach with the icp algorithm. In *CRV*.

AMBERG, B., ROMDHANI, S., AND VETTER, T. 2007. Optimal step nonrigid icp algorithms for surface registration. In *CVPR*.

ANGUELOV, D., KOLLER, D., SRINIVASAN, P., THRUN, S., H.-C.PANG, AND DAVIS, J. 2005. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *Advances in Neural Information Processing Systems NIPS 2004)*.

ARYA, S., MOUNT, D. M., NETANYAHU, N. S., SILVERMAN, R., AND WU, A. Y. 1998. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *J. ACM 45* (November).

AU, O. K.-C., FU, H., TAI, C.-L., AND COHEN-OR, D. 2007. Handle-aware isolines for scalable shape editing. *ACM Trans. Graphic. 26*.

BALLAN, L., AND CORTELAZZO, G. M. 2008. Marker-less motion capture of skinned models in a four camera set-up using optical flow and silhouettes. In *3DPVT*.

BEN-CHEN, M., WEBER, O., AND GOTSMAN, C. 2009. Spatial deformation transfer. In *SCA'09*.

BEN-CHEN, M., WEBER, O., AND GOTSMAN, C. 2009. Variational harmonic maps for space deformation. *ACM Trans. Graph. 28*.

BESL, P. J., AND MCKAY, N. D. 1992. A method for registration of 3-d shapes. *TPAMI*.

BOTSCH, M., AND SORKINE, O. 2008. On linear variational surface deformation methods. *TVCG*.

BUDD, C., AND HILTON, A. 2010. Temporal alignment of 3d video sequences using shape and appearance. In *CVMP'10: 9th Conference on Visual Media Production*.

BUDD, C., HUANG, P., AND HILTON, A. 2011. Hierarchical shape matching for temporally consistent 3d video. In *3DIMPVT*.

CHANG, W., AND ZWICKER, M. 2008. Automatic registration for articulated shapes. *SGP'8) 27*.

CHANG, W., AND ZWICKER, M. 2011. Global registration of dynamic range scans for articulated modelreconstruction. *ACM Transactions on Graphics*.

CHEN, G., SU, H., JIANG, J., AND WU, W. 2010. Safe polyhedral visual hulls. In *Advances in Multimedia Modeling*.

CHEN, L., HUANG, J., SUN, H., AND BAO, H. 2010. Cage-based deformation transfer. *Comput. Graph. 34*.

CONNOR, M., AND KUMAR, P. 2010. Fast construction of k-nearest neighbor graphs for point clouds. *TGCV*.

CORAZZA, S., MÜNDERMANN, L., GAMBARETTO, EMILIANOAND FERRIGNO, G., AND ANDRIACCHI, T. P. 2010. Markerless motion capture through visual hull, articulated icp and subject specific model generation. *Int. J. Comput. Vision 87*.

DE AGUIAR, E., STOLL, C., THEOBALT, C., AHMED, N., SEIDEL, H.-P., AND THRUN, S. 2008. Performance capture from sparse multiview video. In *SIGGRAPH'08*.

DESBRUN, M., MEYER, M., SCHRÖDER, P., AND BARR, A. H. 1999. Implicit fairing of irregular meshes using diffusion and curvature flow. In *SIGGRAPH'99*.

ECKSTEIN, I., PONS, J.-P., TONG, Y., KUO, C.-C. J., AND DESBRUN, M. 2007. Generalized surface flows for mesh processing. In *SGP*.

GALL, J., STOLL, C., DE AGUIAR, E., THEOBALT, C., ROSENHAHN, B., AND SEIDEL, H.-P. 2009. Motion capture using joint skeleton tracking and surface estimation. In *CVPR*.

GRAUMAN, K., SHAKHNAROVICH, G., AND DARRELL, T. 2003. A bayesian approach to image-based visual hull reconstruction. In *CVPR*.

GUENNEBAUD, G., AND GROSS, M. 2007. Algebraic point set surfaces. In *ACM Trans. Graph.*

HORNUNG, A., AND KOBBELT, L. 2006. Robust reconstruction of watertight 3d models from non-uniformly sampled point clouds without normal information. In *SGP'06*.

HUANG, Q.-X., ADAMS, B., WICKE, M., AND GUIBAS, L. J. 2008. Non-rigid registration under isometric deformations. *Comput. Graph. Forum*.

HUANG, P., BUDD, C., AND HILTON, A. 2011. Global temporal registration of multiple non-rigid surface sequences. In *CVPR 2011*, 3473–3480.

JACOBSON, A., BARAN, I., POPOVIĆ, J., AND SORKINE, O. 2011. Bounded biharmonic weights for real-time deformation. *ACM Trans. Graph. 30*.

JOSHI, P., MEYER, M., DEROSE, T., GREEN, B., AND SANOCKI, T. 2007. Harmonic coordinates for character articulation. *ACM Trans. Graph. 26*.

JU, T., SCHAEFER, S., AND WARREN, J. 2005. Mean value coordinates for closed triangular meshes. *ACM Trans. Graph. 24*.

JU, T., ZHOU, Q.-Y., VAN DE PANNE, M., COHEN-OR, D., AND NEUMANN, U. 2008. Reusable skinning templates using cage-based deformations. *ACM Trans. Graph. 27*.

KIM, H., SAKAMOTO, R., KITAHARA, I., TORIYAMA, T., AND KOGURE, K. 2006. Robust foreground segmentation from color video sequences using background subtraction with multiple thresholds. In *Third Korea-Japan Joint Workshop on Pattern Recognition*.

KOLEV, K., BROX, T., AND CREMERS, D. 2012. Fast joint estimation of silhouettes and dense 3d geometry from multiple images. *IEEE TPAMI 34*.

LANDRENEAU, E., AND SCHAEFER, S. 2010. Poisson-based weight reduction of animated meshes. *Comput. Graph. Forum*.

LAURENTINI, A. 1994. The visual hull concept for silhouette-based image understanding. *IEEE TPAMI*.

LI, H., SUMNER, R. W., AND PAULY, M. 2008. Global correspondence optimization for non-rigid registration of depth scans. *SGP'08 27*.

LI, H., ADAMS, B., GUIBAS, L. J., AND PAULY, M. 2009. Robust single-view geometry and motion reconstruction. *ACM Trans. on Graph.*.

LI, H., LUO, L., VLASIC, D., PEERS, P., POPOVIĆ, J., PAULY, M., AND RUSINKIEWICZ, S. 2012. Temporally coherent completion of dynamic shapes. *ACM Trans. Graph.*.

LIPMAN, Y., KOPF, J., COHEN-OR, D., AND LEVIN, D. 2007. Gpu-assisted positive mean value coordinates for mesh deformations. In *SGP '07*.

LIPMAN, Y., LEVIN, D., AND COHEN-OR, D. 2008. Green coordinates. *ACM Trans. Graph. 27*.

LIU, Y., DAI, Q., AND XU, W. 2010. A point-cloud-based multiview stereo algorithm for free-viewpoint video. *TVCG*.

LIU, Y., STOLL, C., GALL, J., SEIDEL, H.-P., AND THEOBALT, C. 2011. Markerless motion capture of interacting characters using multiview image segmentation. In *CVPR'11*.

LOW, K.-L. 2004. Linear least-squares optimization for. point-to-plane icp surface registration.

MATUSIK, W., BUEHLER, C., AND MCMILLAN, L. 2001. Polyhedral visual hulls for real-time rendering. In *Eurographics Workshop on Rendering Techniques*.

MONTENEGRO, A. A., VELHO, L., CARVALHO, P. C. P., AND SOSSAI, J. J. 2006. Polygonization of volumetric reconstructions from silhouettes. *SIBGRAPI*.

MÜNCH, D., COMBÈS, B., AND PRIMA, S. 2010. A modified ICP algorithm for normal-guided surface registration. In *Proceedings of SPIE 7623*.

OFIR WEBER, R. P., AND GOTSMAN, C. 2012. Spatial deformation transfer. In *Computer Graphics Forum*.

POPA, T., SOUTH-DICKINSON, I., BRADLEY, D., SHEFFER, A., AND HEIDRICH, W. 2010. Globally consistent space-time reconstruction. *SGP*.

SAGAWA, R., AKASAKA, K., YAGI, Y., HAMER, H., AND GOOL, L. V. 2009. Elastic convolved icp for the registration of deformable objects. In *3DIM*.

SHARF, A., ALCANTARA, D. A., LEWINER, T., GREIF, C., SHEFFER, A., AMENTA, N., AND COHEN-OR, D. 2008. Space-time surface reconstruction using incompressible flow. *ACM Trans. Graph. 27*.

SORMANN, M., ZACH, C., BAUER, J., KARNER, K., AND BISHOF, H. 2007. Watertight multi-view reconstruction based on volumetric graph-cuts. In *Proceedings of the 15th Scandinavian conference on Image analysis*, SCIA'07.

STARCK, J., AND HILTON, A. 2007. Surface capture for performance-based animation. *IEEE Comput. Graph. Appl. 27*.

STOLL, C., KARNI, Z., RSSL, C., YAMAUCHI, H., AND SEIDEL, H.-P. 2006. Template deformation for point cloud fitting. In *SPBG*.

SÜSSMUTH, J., ZOLLHÖFER, M., AND GREINER, G. 2010. Animation transplantation. *Comput. Animat. Virtual Worlds 21*, 173–182.

TEVS, A., BERNER, A., WAND, M., IHRKE, I., BOKELOH, M., KERBER, J., AND SEIDEL, H.-P. 2012. Animation cartography: intrinsic reconstruction of shape and motion. *ACM Trans. Graph. 31*, 2.

TONG, J., ZHOU, J., LIU, L., PAN, Z., AND YAN, H. 2012. Scanning 3d full human bodies using kinects. *TVCG*.

VLASIC, D., BARAN, I., MATUSIK, W., AND POPOVIĆ, J. 2008. Articulated mesh animation from multi-view silhouettes. *ACM Trans. Graph.*.

VOGIATZIS, G., TORR, P. H. S., AND CIPOLLA, R. 2005. Multi-view stereo via volumetric graph-cuts. In *CVPR*.

VOLODINE, T., FLOATER, M. S., AND ROOSE, D. 2007. Volumetric snapping: Watertight triangulation of point clouds. In *GRAPP'07*.

WALD, I., BOULOS, S., AND SHIRLEY, P. 2007. Ray tracing deformable scenes using dynamic bounding volume hierarchies. *ACM Trans. Graph. 26*, 1.

WALD, I. 2007. On fast construction of sah-based bounding volume hierarchies. In *Proceedings of the 2007 IEEE Symposium on Interactive Ray Tracing*, RT '07.

WAND, M., JENKE, P., HUANG, Q.-X., BOKELOH, M., GUIBAS, L. J., AND SCHILLING, A. 2007. Reconstruction of deforming geometry from time-varying point clouds. In *SGP*.

WEILAN LUO, T. Y., AND AIZAWA, K. 2010. Articulated human motion capture from. segmented visual hulls and surface reconstruction. *APSIPA*.

WINDHEUSER, T., SCHLICKEWEI, U., SCHMIDT, F. R., AND CREMERS, D. 2011. Large-scale integer linear programming for orientation-preserving 3d shape matching. In *SGP*.

WITKIN, A., AND KASS, M. 1988. Spacetime constraints. In *SIGGRAPH '88*.

XIAN, C., LIN, H., AND GAO, S. 2009. Automatic generation of coarse bounding cages from dense meshes. In *Shape Modeling International*.

XIAN, C., LIN, H., AND GAO, S. 2012. Automatic cage generation by improved obbs for mesh deformation. In *Vis. Comput*.

YEH, I.-C., LIN, C.-H., SORKINE, O., AND LEE, T.-Y. 2010. Template-based 3d model fitting using dual-domain relaxation. *IEEE TVCG*.

YOUS, S., LAGA, H., KIDODE, M., AND CHIHARA, K. 2007. Gpu-based shape from silhouettes. In *GRAPHITE '07*.

ZHENG, Q., SHARF, A., TAGLIASACCHI, A., CHEN, B., ZHANG, H., SHEFFER, A., AND COHEN-OR, D. 2010. Consensus skeleton for non-rigid space-time registration. *Computer Graphcis Forum 29*, 2.