Deformable Registration using Patch-Wise Shape Matching

Francesco Bonarrigo[a], Alberto Signoroni[a], Mario Botsch[b]

[a]*Information Engineering Department, University of Brescia, Italy*
[b]*Computer Graphics & Geometry Processing Group, University of Bielefeld, Germany*

Abstract

We present a novel approach for non-rigid registration of partially overlapping surfaces acquired from a deforming object. To allow for large and general deformations our method employs a nonlinear physics-inspired deformation model, which has been designed with a particular focus on robustness and performance. We discretize the surface into a set of overlapping patches, for each of which an optimal rigid motion is found and interpolated faithfully using dual quaternion blending. Using this discretization we can formulate the two components of our objective function—a fitting and a regularization term—as a combined global shape matching problem, which can be solved through a very robust numerical approach. Interleaving the optimization with successive patch refinement results in an efficient hierarchical coarse-to-fine optimization. Compared to other approaches our as-rigid-as-possible deformation model is faster, causes less distortion, and gives more accurate fitting results.

Keywords: Non-rigid registration, nonlinear shape deformation

1. Introduction

The need to compensate for deformation present on a set of captured data, or between modeled objects, is inbred to all dynamic 3D imaging and graphics applications, such as capturing, modeling, representation, recognition, morphing, and animation. In particular the 3D scanning of physical objects that undergo some kind of non-rigid motion leads to a space-time shape matching problem, which has to be solved to achieve a non-rigid alignment between partial views.

Depending on the application and the scanning system, the type of non-rigid motion the object is undergoing and the distortion caused for local and global features can vary a lot. This mainly depends on the temporal distribution of the single scans. For real-time scanners the deformation of consecutive scans is rather small and their overlap rather large. The acquired data is typically not problematic from a registration perspective. In contrast, when high-end scanners are used to acquire precise high-resolution scans for production quality models, the delay between two consecutive scans is significantly larger, leading to more deformation and less overlap between individual scans, and therefore to a more challenging registration problem.

Large-scale deformations require a more sophisticated regularization of the deformation field in order to prevent implausible deformations and to keep the surface locally rigid during the registration process (Figure 1). Since the small strain assumption does not hold for large deformation, the regularization is typically implemented as some nonlinear physics-inspired strain measure, which is to be minimized simultaneously with the fitting error between the two scans. The resulting nonlinear minimization is challenging in terms of numerical robustness and computational performance.

In this paper we present an approach for accurate and robust deformable registration of models acquired by high-quality, high-resolution 3D scanners. Our method is able to compensate for various kinds of motion and deformation the acquired data can exhibit, which is achieved by a nonlinear, physics-inspired deformation regularization. Our method is based on a space deformation approach, such that it can process arbitrary point-sampled data and is not affected negatively by degenerate input meshes. We discretize the surface into a hierarchy of partially overlapping patches, for each of which a distinct rigid transformation is found by minimizing a global objective function that takes into account both the need for accurate alignment as well as for the regularity of the deformation field. The resulting rigid transformations are then extended to all sample points in a rigid manner using dual quaternion interpolation. As a result of this as-rigid-as-possible deformation
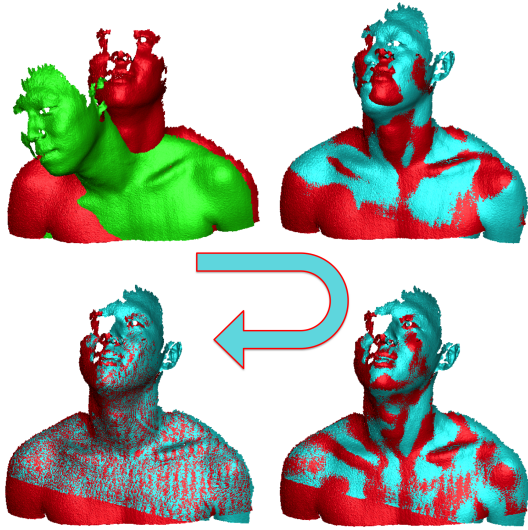
Figure 1: Non-rigid registration of the Torso dataset, where the green scan is to be registered to the red target model. The cyan models show the the deforming surface during the optimization process, until convergence is reached.

our dynamic registration avoids unnecessary distortion and faithfully preserves geometric features. Both the fitting and regularization terms are based on a global shape matching strategy, which results in a unified objective function that can be minimized in a numerically very robust manner. Finally, our hierarchical patch discretization and correspondence refinement further improves stability and accelerates convergence. An example of our deformable registration is shown in Figure 1.

Our approach demonstrates a favorable combination of fitting accuracy, geometric feature preservation, and computational performance, as we evaluate on several challenging datasets and compare to recent state-of-the-art techniques. As such, our approach can have many practical applications for the acquisition of objects that undergo undesired motion, such as acquisition of body parts for orthopedic applications, as well as for the analysis of deformation with respect to a reference surface for manufacturing quality control.

2. Related Work

Non-rigid shape registration belongs to the wide family of shape correspondence problems [1, 2, 3], and its understanding is of fundamental importance for a variety of advanced applications in computer graphics, geometric modeling, and computer vision. In this work we concentrate on non-rigid registration of partial views, which is a fundamental problem in 3D scanning. Since

all acquisition sensors can only provide a limited field of view, scanning a deforming object requires the solution of partial shape matching problems. This defines what has been called the problem of time-varying (non-rigid) registration, which has found different solutions according to the attainable acquisition frame rate for a given scanning technology (real-time scanning or non-interactive scanning). This technological distinction can also act as a classification of non-rigid registration techniques, since the amount of deformation that needs to be taken care of is typically less for scanners with a faster acquisition rate.

Small Deformations

A first family of approaches can be associated to the advancements of real-time scanning technologies [4, 5], where data acquisition is characterized by a regular and dense frame rate (in the order of dozens of frames per second). Recently developed, handy, and affordable devices from the entertainment industry are having a great impact on research in computer vision, computer graphics, and multimedia [6], allowing for 3D object modeling and scene reconstruction [7]. With respect to professional high-end scanners, real-time devices provide only limited spatial resolution and a high amount of geometric noise. For objects that do not move or deform too quickly, subsequent view overlapping is quite high and the deformation between scans is rather small, therefore allowing to assume that subsequent scans are already coarsely aligned.

This inspired the development of non-rigid variants of rigid alignment methods relying on the estimation of correspondence sets [8]. An alternative method that does not have to explicitly compute correspondences has been proposed in [9]. Under the gradual deformation hypothesis, template-free methods that guarantee global consistency of the space-time object reconstruction and prevent erroneous topology changes have been proposed in [10, 11], while the template-based approach [12] has been proposed in the same perspective. Other approaches consider non-rigid alignment as a form of compensation for nonlinearities related to the acquisition device, such as optical distortion and calibration errors [13, 14]. A multiview global alignment solution of these issues has been proposed in [15], while loop closure issues have been addressed in [16].

Large Deformations

A second group of approaches, which is more similar to the problem we wish to address, aims to obtain a high-quality, precise model of a deforming object, for

instance for the sake of reverse engineering or high-fidelity reproduction. The acquisition is usually done at irregular time intervals, therefore a high overlap of consecutive views cannot be assumed, nor can consecutive scans be considered coarsely aligned as in the case of real-time scanners. It is worth noting that the presence of deformation could be considered either as a kind of noise to compensate for, or, on the contrary, as useful additional information that can be attached to a static "consensus model" of the object being acquired.

In order to constrain the registration problem some approaches assume the availability of prior knowledge of either the object shape (such as template-based approaches [17, 18, 19]) or of the deformation model (e.g., assuming articulated motion [20, 21]). Under general deformation assumptions, some approaches based on non-rigid ICP have been proposed [22, 18]. In particular, a reference template is needed in [22], while [18] presents a reliable correspondence selection strategy based on a combination of point-to-point and point-to-plane metrics. The method proposed in [23] estimates a diffeomorphism between data that have undergone large deformations.

Other approaches successfully handle large and general deformations without the need of a-priori assumptions by minimizing a global energy that takes into account both fitting error and deformation smoothness [24, 25, 12]. Our method is most closely related to these latter approaches, as we also employ an objective function consisting of fitting and regularization terms. However, our method differs in a few important aspects, which result in a noticeable improvement over the very successful and quite mature state-of-the-art techniques. We discuss the similarities and differences in Section 5.3, after having presented our method in detail in the following sections.

In general, when the deformation increases, the problem of automatically searching for reliable correspondences between partial views becomes more challenging. Several interesting alternatives to closest-point correspondences have been proposed, based on geometric features and geodesic distances [25], on the Heat Kernel Map [26], Möbius transformation [27], or a blending of several intrinsic maps [28]. The focus of this paper is the deformation model instead of the computation of robust correspondences. We therefore restrict to simple closest-point correspondences and use a stronger regularization to prevent distortions due to wrong correspondences at the early stage of the registration. Even larger deformations of the input scans could be handled by enhancing our method by any of the above-mentioned correspondence computations.

3. Method Overview

Our main objective is to estimate the optimal deformation capable of aligning a scan $\mathcal{S}$ (the surface we want to deform) to a target model $\mathcal{M}$ (the reference surface), while at the same time preserving the shape of the scan as much as possible.

Solving this problem for all the points simultaneously can quickly become an impractical task, depending on the size of $\mathcal{S}$. However, since the deformation of the object typically is a smooth function, we can safely assume that nearby points undergo similar deformations, and discretize the problem at a coarser resolution, solving the registration for a subset of evenly-spaced points, and then interpolating the deformation to the rest of the scan. Similarly to [29] we denote these points as *nodes*, and refer to their set with $\mathcal{N}$. Each node has a radius of influence, chosen such that the nodes' influence regions correspond to a set of partially overlapping patches. Each node $n \in \mathcal{N}$ is then associated a rigid transformation, i.e., a rotation $\mathbf{R}_n \in \mathbb{R}^{3\times3}$ and a translation $\mathbf{t}_n \in \mathbb{R}^3$. Since rotations can be represented by 3 parameters, this yields 6 degrees of freedom (DoFs) per node. During the optimization the density of the set $\mathcal{N}$ will be hierarchically adapted in a coarse-to-fine manner, so that distortion on $\mathcal{S}$ is progressively compensated.

The per-node rigid motions are determined by minimizing an objective function consisting of a fitting term E_{fit}, which drives the deformation of the scan toward the model, and a regularization term E_{reg}, which maintains a smooth and physically plausible deformation throughout the optimization.

For the ***fitting term*** we resort to a set of correspondences $C = \{\mathcal{P}, \mathcal{Q}\}$, where the sample points $\mathbf{p} \in \mathcal{P} \subset \mathcal{S}$ are selected from the scan $\mathcal{S}$, while their corresponding points $\mathbf{q} \in \mathcal{Q} \subset \mathcal{M}$ are found on the target model $\mathcal{M}$. The fitting term penalizes for each node $n \in \mathcal{N}$ the (squared) distance of corresponding points $(\mathbf{p}, \mathbf{q}) \in C_n$ that are within n's influence radius, where $\mathbf{q}$ is being transformed by n's rotation and translation:

$$E_{\text{fit}} = \sum_{n \in \mathcal{N}} \sum_{(\mathbf{p}, \mathbf{q}) \in C_n} \frac{\|\mathbf{R}_n \mathbf{p} + \mathbf{t}_n - \mathbf{q}\|^2}{|\mathcal{N}| \cdot |C_n|}. \qquad (1)$$

The rigid transformations of neighboring nodes n and m are not completely independent, but will instead be similar due to correspondence points $(\mathbf{p}, \mathbf{q})$ located in the partial overlap of their influence regions, i.e., $C_n \cap C_m$. Nevertheless, the ***regularization term*** explicitly penalizes the difference between neighboring nodes' transformations, as measured by the distance of points $\mathbf{p}$ in their
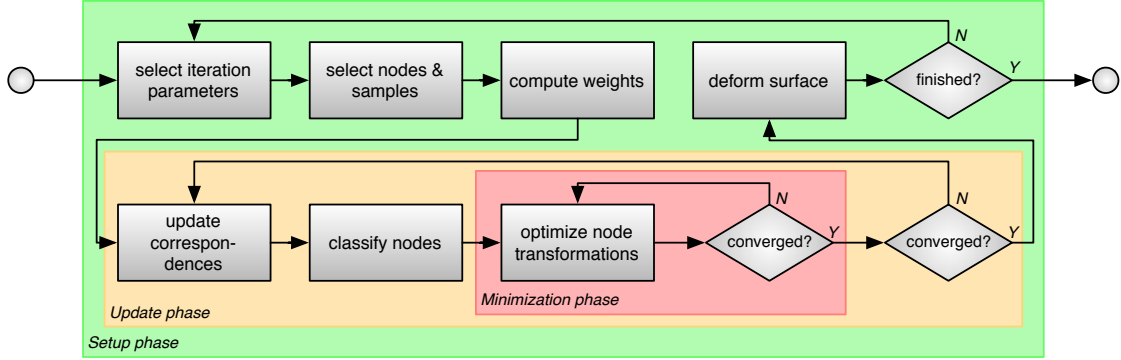
Figure 2: The proposed deformable alignment pipeline, split into three nested phases: Setup, Update and Minimization.

overlap region $\mathcal{P}_n \cap \mathcal{P}_m$ after being transformed by the rigid motions of n and m, respectively:

$$E_{\text{reg}} = \sum_{n \in \mathcal{N}} \sum_{m \in \mathcal{N}} \sum_{\mathbf{p} \in \mathcal{P}_n \cap \mathcal{P}_m} \frac{\|\mathbf{R}_n \mathbf{p} + \mathbf{t}_n - \mathbf{R}_m \mathbf{p} - \mathbf{t}_m\|^2}{|\mathcal{N}|^2 \cdot |\mathcal{P}_n \cap \mathcal{P}_m|}. \quad (2)$$

The global energy is a weighted sum of fitting and regularization

$$E_{\text{total}} = w_{\text{fit}} E_{\text{fit}} + (1 - w_{\text{fit}}) E_{\text{reg}}, \quad (3)$$

where the fitting weight w_{fit} is increased during the optimization so that the surface is deformed in a stiff-to-smooth manner.

Our framework can be considered as a sequence of operations distributed on three nested loops (see Figure 2), where in each loop some of the optimization features or parameters are selected and considered fixed throughout the execution of the inner loops. The three phases are outlined below, and the involved techniques are described in detail in the following section.

The **setup phase** is responsible for adjusting the optimization parameters (Section 4.1) and for selecting the set of nodes $\mathcal{N}$ and correspondence samples $\mathcal{P}$ (Section 4.2). At this step we also compute the weights used for interpolating the nodes' transformations to all the points of the scan $\mathcal{S}$ (Section 4.3). The setup phase is repeated until a given number of iterations is reached.

The **update phase** updates the target points $\mathcal{P}$ for correspondences $C = (\mathcal{P}, \mathcal{Q})$ according to the current deformation estimate of $\mathcal{S}$ (Section 4.4). It also performs a classification of the nodes (Section 4.5) based on the updated correspondences, which allows to speed up the minimization and avoid potential instabilities. The update phase is terminated when the correspondences C do not improve any further.

In the **minimization phase** the energy (3) is solved while keeping the set of parameters, nodes, and correspondences fixed (Section 4.6). Similarly in spirit to

[30], we can interpret all the nodes as rigid objects, which are coupled to the model $\mathcal{M}$ through (1) and to each other through (2). In this view, the minimization of (3) is a global shape matching problem, which can be solved robustly as shown in [31, 32].

4. Detailed Description and Implementation Notes

4.1. Parameter selection

The proposed approach relies on three main parameters. Besides the first one, the other two require an initial and final value each, so that at a given iteration the current values for these parameters are linearly interpolated.

The first parameter is the **number of iterations** for which the setup phase, i.e., the outer loop, should be repeated. This value may require tuning according to the complexity of the deformation that has to be handled: more demanding cases may need a greater number of iterations, so that the scan slowly adapts to the target model.

The **node spacing** represents the minimum Euclidean distance between two nodes (and indirectly defines the spacing of correspondence samples, Section 4.2). While its initial value should be tuned according to the metric size of the processed scan, its final value controls the smallest deformation that can be handled. During the optimization this value decreases, so that the number of nodes increases. This way the optimization initially takes care of large, low-frequency deformations, while handling small, high-frequency ones at the end, in a coarse-to-fine manner.

The **fitting weight** w_{fit} represents the weight associated to the fitting term E_{fit} in the objective function (3). Similar to other approaches, this parameter is monotonically increased during the optimization (typically from
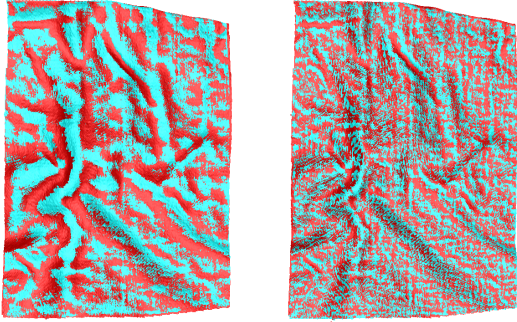
Figure 3: Alignment between scan $\mathcal{S}$ (cyan) and target model $\mathcal{M}$ (red) obtained with same number of nodes $\mathcal{N}$, but different numbers of correspondence samples $\mathcal{P}$. Left: *sample spacing = node spacing*. Right: *sample spacing = node spacing*/4.

0.1 to 0.9), causing the surface to remain stiff during the initial optimization stages, while getting softer toward the end, in a stiff-to-smooth fashion.

4.2. Node and sample points selection

At every iteration of the setup phase we select a progressively denser set $\mathcal{N}$ of regularly-spaced nodes on the scan $\mathcal{S}$, controlled by the node spacing parameter, which can be implemented efficiently through simple farthest point sampling. The node spacing also controls the sampling of the correspondence points $\mathcal{P}$. A key difference of our method to other approaches, such as [25] and [12], is that we chose the density of correspondence samples as a multiple of the node density, which leads to a more accurate estimation of the deformation field (see Figure 3). In particular, in all our examples we chose the spacing of correspondence samples to be 1/4 of the node spacing. An example of how the sets $\mathcal{N}$ and $\mathcal{P}$ are distributed on the scan $\mathcal{S}$ is shown in Figure 4.

4.3. Blending of rigid transformations

The deformation of points $\mathbf{p} \in \mathcal{S}$ is computed by interpolating the nodes' rigid transformations. The interpolation weights can be computed by an arbitrary smooth weighting function with local support radius r. In our case a simple inverse distance weighting turned out to be sufficient, where the weight $w_n(\mathbf{p})$ of point $\mathbf{p}$ with respect to node n is computed as

$$w_n(\mathbf{p}) \;=\; \left(1 - \frac{\|\mathbf{n} - \mathbf{p}\|}{r} \right)_+ .$$

Here, r is the influence radius of node n, typically chosen as 1.25 times the node spacing, and $(\cdot)_+$ clamps negative values to zero. The weights are normalized such that $\sum_{n \in \mathcal{N}} w_n(\mathbf{p}) = 1$.

Once the blending weights have been computed, we can interpolate the transformations of the nodes to all other points of the scan. The approaches of Huang et al. [25] and Li et al. [24, 12] employ linear blend skinning [33] for this task, which is known to produce artifacts when interpolating strongly different transformations. In contrast, we employ dual quaternion skinning [34], which produces a more accurate, artifact-free interpolation of rigid transformations by inherently preserving the orthogonality of rotation matrices. Due to the more complex formulation of the dual-quaternion interpolation, we cannot integrate the interpolation of node transformations into the minimization of correspondence errors, as has been done in the embedded deformation approach [24, 12] (for simple *linear* interpolation of node transformations). Instead, our dual-quaternion interpolation is performed after node transformations have been computed, similar to [25].

4.4. Correspondence update

Every time new correspondence samples $\mathcal{P} \subset \mathcal{S}$ have been selected (Section 4.2) or have been deformed during the registration process (Section 4.3), we have to update their corresponding points Q on the target model $\mathcal{M}$.

To this end, we find for each sample $\mathbf{p} \in \mathcal{P}$ its closest point $\mathbf{q} \in \mathcal{M}$. To avoid erroneous correspondences, we discard any point couple $(\mathbf{p}, \mathbf{q})$ if their Euclidean distance or the angle between their normals is above a certain threshold [35]. These thresholds are linearly interpolated from rather large initial values to more stringent final values, such that towards the end of the optimization, when the surfaces are very close to each other, outliers can be effectively rejected. While the thresholds on the distance of corresponding points depends on the model dimensions and the amount of noise and deformation present, the normal deviation thresholds are $60°$ in the beginning and $30°$ at the end for all our examples. The updated correspondence set then becomes $C = (\mathcal{P}, Q)$.

4.5. Nodes classification

To guarantee a numerically robust computation of each node's rigid transformation by minimizing the objective function (3), each node has to be properly constrained. This can happen either through sufficient valid constraints for the fitting term (1), or through a sufficient overlap with neighboring (constrained) nodes in the regularization term (2).

Nodes that do not have a sufficient number of valid correspondences are excluded from the fitting term. If
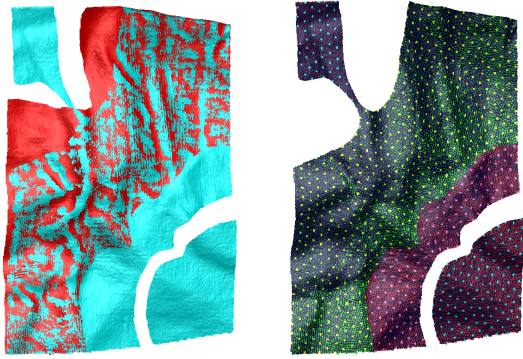
Figure 4: An extreme (synthetic) version of the Pillow dataset. Left: The scan (cyan) and the target model (red) after alignment. Right: Classification of nodes $\mathcal{N}$ (yellow, cyan, pink) and correspondence samples $\mathcal{P}$ (green, red) on the scan $\mathcal{S}$. Valid correspondence samples are green, invalid ones are red. Yellow nodes have sufficiently many valid correspondences, cyan nodes are not (sufficiently) connected, and pink nodes do not have sufficient correspondences, but are sufficiently connected to a constrained region. Both the ill-connected nodes (upper-left) and isolated nodes (lower-right) are correctly labeled as disconnected.

such nodes still have a sufficient overlap with other nodes, they can be used in the regularization term. Otherwise, they have to be completely discarded from the optimization. Since in our approach the densities of nodes $\mathcal{N}$ and correspondence samples $\mathcal{P}$ are coupled (Section 4.2), we can simply use the fixed thresholds $|\mathcal{C}_n| > 20$ and $|\mathcal{P}_n \cap \mathcal{P}_m| > 20$ for these tests.

The node classification starts by collecting all nodes with sufficient correspondences (yellow in Figure 4). From each of those a breadth first search is started, proceeding to all neighboring nodes with sufficient overlap. At the end of the process, all unmarked nodes are considered disconnected and are temporarily excluded from the current optimization step (cyan in Figure 4). We found this node classification to be crucial for avoiding numerical instabilities, in particular when working with raw scans that did not undergo any sort of preprocessing and therefore might contain outliers as well as ill-constrained (or un-constrained) regions.

4.6. Numerical optimization

In the numerical minimization of (3), we can consider each node as an individual rigid object (having 6 DoF for rotation and translation). These rigid parts are coupled to each other through the regularization term (2), which is closely related to the PriMo approach [32, 30], where it has proven to be a physically plausible deformation model. In contrast to the interactive PriMo application, in our approach the deformation is driven by the correspondences, i.e., by the fitting term (1) that couples the nodes of the scan to the target model. Our optimization can be solved robustly and efficiently by the Gauss-Newton approach described in [31, 32]. In each Gauss-Newton iteration, rigid transformation updates are linearized as affine transformations (helical screw motions), and the optimal updates are found by solving a *sparse* $6|\mathcal{N}| \times 6|\mathcal{N}|$ linear system. The resulting screw motions are projected back to the rigid motion manifold and used to update the current node transformations. This kind of nonlinear optimization was shown to be very robust and reasonably efficient in [32, 30], and worked reliably in our application.

5. Results and discussion

The described approach has been implemented in C++ and tested on a 2.4 GHz Intel Dual Core laptop with 4 GB of RAM. The code can exploit multi-core processing. Several datasets were considered, each composed by a pair of surfaces, previously referred to as the scan $\mathcal{S}$ and the target model $\mathcal{M}$, where the former presents both some degree of overlap and deformation with respect to the latter. The Sole (a shoe sole subject to elastic bending), Mould (a foam rubber mould subject to external deforming forces), and Pillow datasets (a pillow taken in two different creased states) were acquired with a high-resolution, structured-light professional scanner, while the Torso dataset (kindly provided by the authors of [24]) was obtained through the fast scanner described in [5]. Both the Sole and Mould datasets are characterized by a smooth, low-frequency warping, however the Sole is a problematic dataset for any ICP-based technique (such as ours) due to the presence of repeated patterns, which causes many local minima in the objective function and increases the risk of obtaining a suboptimal alignment. In addition we tested our approach on some face scans, which represent more noisy and partial surfaces with complex deformations.

5.1. Alignment performance

For all the considered datasets, our approach succeeded in estimating a physically plausible deformation, as testified by the alignment shown in Table 1 and Figure 5. The final average distance between the scans is consistent among all datasets with the exception of the Torso model, where a greater final distance is justified by the more noisy surface compared to the other high-resolution scans. Thanks to our physically plausible regularization term in the error functional, the extension of the deformation field to the rest of the surface leads to a plausible warping, even for areas where no overlap with the target surface exists.

This is shown for the Sole dataset, where the wedge heel is correctly bent according to the rest of the data, as well as for the Torso dataset, where the upper part of the face has been warped correctly, even though it was completely missing on the target model.

More examples with noisy surfaces, incomplete data, and partial overlap are shown Figure 6, where a scan of a neutral face is registered to scans of different facial expressions. As the corresponding statistics in Table 2 demonstrate, our method succeeds to accurately register the different scans in reasonable time. For the example "neutral→smile" a higher resolution could be used to reduce the deviation in the mouth area, however, the other regions are sufficiently accurate. Hence, this problem could well be solved using adaptive (instead of uniform) refinement, which we leave for future work. Nevertheless, these examples demonstrate the capability of our method for building a blend shape facial rig by matching a neutral scan (or a template facial model) to recorded facial expressions.

5.2. Timing breakdown

We now present a breakdown of the involved computations of our approach with respect to the conceptual operations described in the pipeline presented in Figure 2. While the exact time balance may vary depending on the dataset characteristics and node density, the following general trend can be observed. The most time-consuming part involves the solution of the minimization problem, which accounts for 40% of the total time, followed by the correspondence update phase, which requires 25% of the total time. Next is the application of the deformation field to the entire surface (15%), while the remaining 20% of the time is distributed among the other operations (parameter selection, selection of nodes and samples, weight computation, node classification, I/O operations).

5.3. Comparison to other techniques

In order to further assess our approach, we compare to the two approaches most similar to ours, namely the ones of Huang et al. [25] and Li et al. [12] (which improves upon [24]). Since our approach shares certain details with [25] and [12], but differs in other crucial aspects, we start with a theoretical discussion of the similarities and differences between the methods, before providing an experimental comparison.

Methodological comparison. The approach of Huang et al. [25] selects uniformly distributed nodes on the deforming scan, and infers their deformation from a set of robust correspondences computed by taking into account both geodesic distance and feature signature consistency. During the optimization, neighboring nodes undergoing the same rigid transformation are clustered and are assigned a common transformation. However, since these clusters are only allowed to grow, the final phase of the optimization may be hampered since sub-cluster alignment accuracy would be required to improve the fitting quality. Their error functional consists of a fitting and a regularization term, both weighted with constant factors.

Similarly to [25] we compute a hierarchy of nodes, but instead of a fine-to-coarse node clustering, our node selection strategy actually is a coarse-to-fine optimization. This is particularly useful toward the end of the processing, when optimization of the deformation field at node level is necessary to reach complete convergence. Moreover, our adaptive balancing between fitting and regularization is more flexible and further contributes to reaching more accurate results.

The method of Li et al. [12] (which improves upon [24]) also employs an adaptive weighting for fitting and regularization in order to reach a higher fitting accuracy. Their regularization is based on the embedded deformation approach [29], which also discretizes the scan by a set of nodes that build the so-called deformation graph. Each node has an associated affine transformation, which is kept close to a rigid motion by a special term in their energy formulation. As a consequence, their method uses 12 DoFs per node (affine motion), whereas we use 6 DoFs per node only (rigid motion). This consequently reduces memory consumption and computational costs for our numerical optimization. In [12] the resulting node transformations are then linearly interpolated to all points of the scan, which for highly differing matrices might lead to the well-known artifacts of linear blend skinning (as described in [34]). In contrast, our use of dual quaternion skinning should help to avoid these artifacts and reduce distortion during the registration process.

In contrast to [25], however, Li et al. perform an adaptive refinement of the deformation graph in region of highly differing local transformation, which somewhat reduces the linear interpolation problems and allows for more localized updates of the deformation field. In comparison, our hierarchical but uniform discretization is less flexible, and an adaptive refinement is a promising subject of future work.

Our method differs from both approaches by selecting a greater number of correspondences compared to the node density, which leads to superior alignment accuracy, as demonstrated in Section 4.2 and Figure 3.

Model	$\|\mathcal{S}\|$	$\|\mathcal{N}\|$	initial dist. [mm]	final dist. [mm]	iters	time [s]
Sole	360 k	1986	20	0.035	10	77
Mould	1200 k	1095	100	0.034	4	37
Torso	135 k	953	600	0.142	50	58
Pillow	200 k	2423	5	0.031	10	55

Table 1: Registration performance for the models of Figure 5. $\|\mathcal{S}\|$ is the complexity of the scan, $\|\mathcal{N}\|$ the maximum number of nodes reached during optimization. Reported distances correspond to the average distances in the initial and final state, respectively. The iterations numbers correspond to the outer-most loop (setup-phase).
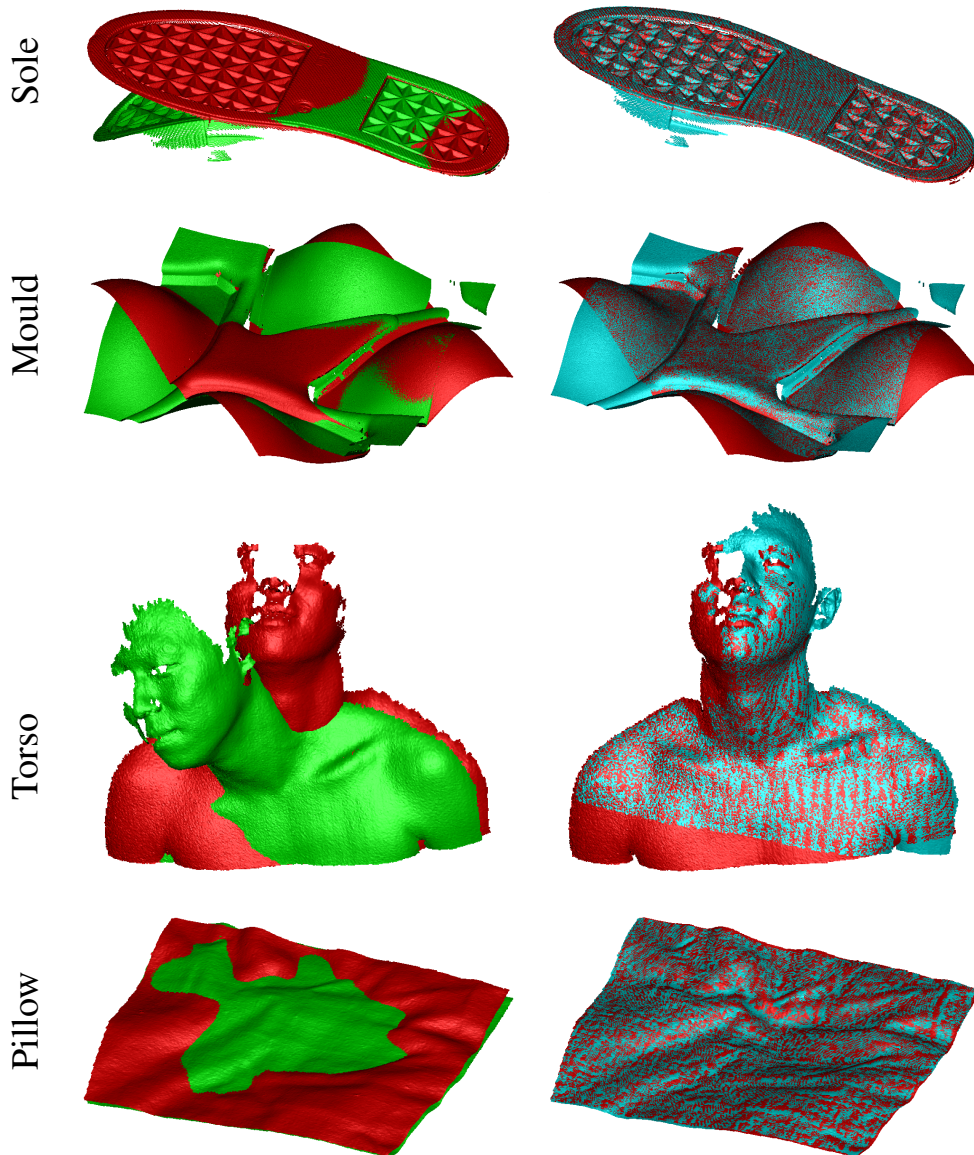

Figure 5: Registration results. The left column shows the original scan (green) and the target model (red). The right column show the deformed scan (cyan) and the target model (red).

Model	$|\mathcal{S}|$	$|\mathcal{N}|$	distance [mm]	time [s]
Neutral → Happy	114k	2257	0.066	137
Neutral → Smile	114k	2207	0.092	123
Neutral → Sad	114k	2191	0.072	131
Neutral → Asym.	114k	2220	0.072	109

Table 2: Registration performance of the face scans of Figure 6. $|\mathcal{S}|$ is the complexity of the scan, $|\mathcal{N}|$ the maximum number of nodes reached during optimization. Reported distance corresponds to the average distances in the final state.
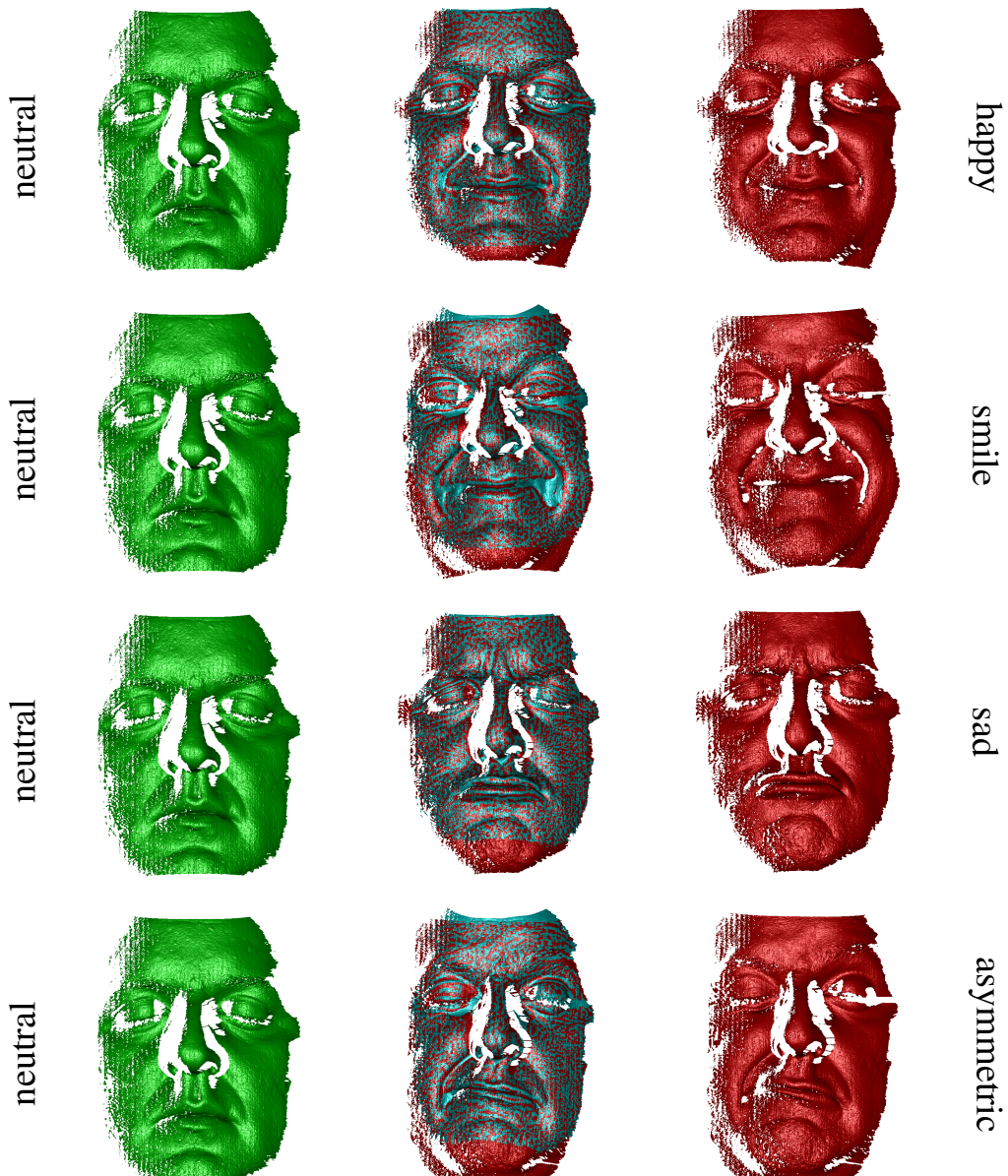

Figure 6: Registration of face scans. The left column shows the neutral source mesh, the right column the target expressions, and the center the overlay of deformed neutral scan (cyan) and target (red).

Model	Approach	Distance [mm]	Distortion [$\cdot 10^{-5}$]	$|\mathcal{N}|$	Time [s]
Sole	Huang et al. 08	0.248	9.34	400	4
	Li et al. 09	0.041	2.53	730	60
	Proposed ($\mathcal{D}$)	0.040	2.48	725	35
Mould	Huang et al. 08	0.094	2.52	400	2
	Li et al. 09	0.045	0.47	709	54
	Proposed ($\mathcal{D}$)	0.036	0.29	710	27
Torso	Huang et al. 08	0.698	29.6	400	4
	Li et al. 09	0.178	8.72	268	38
	Proposed ($\mathcal{D}$)	0.194	8.46	266	35
Pillow	Huang et al. 08	0.240	29.2	400	3
	Li et al. 09 (*)	(0.049)	2.68	419	62
	Proposed ($\mathcal{D}$)	0.086	2.36	421	13

Table 3: Comparison between Huang et al. [25], tested on a single-core 3.2 GHz CPU, Li et al. [12], tested on a 2.6 GHz Intel Core i7, and our approach (downgraded for comparison, $\mathcal{D}$), tested on a 2.4 GHz Intel Dual Core. (*) Note that the method of Li et al. failed for the Pillow example (shifted alignment, see Figure 7).

Experimental comparison. The authors of both approaches kindly agreed to apply their techniques to our datasets, and provided us with the deformation results as well as with the maximum number of nodes employed and the computational performance. In order to provide a fair experimental comparison we decided to adapt (downgrade) our algorithm to the approach that obtained the best results between the two. In particular, we decreased the maximum number of nodes employed by our algorithm to roughly match the number of nodes of the best algorithm. In addition, since both other methods do not exploit multi-core processing, we deactivated this feature in our algorithm. In the following, we indicate this adapted configuration with the symbol ($\mathcal{D}$) to distinguish its results from the ones obtained by employing our optimal working configuration, already presented in Table 1 and Figure 5. The results of the comparative assessment are organized in Table 3 and Figure 7.

Non-rigid registration is always a compromise between fitting accuracy and geometric distortion. Zero fitting error can be achieved by simply projecting the scan onto the target model, but this would cause high distortion. Performing only a rigid registration avoids any distortion, but leads to a high fitting error. To allow for a more meaningful comparison we therefore also report the isometric distortion in Table 3. We measure distortion (in a simplified manner) as the squared relative difference of edge length between the original scan (length L_e) and the deformed scan (length l_e), averaged over all edges $e \in \mathcal{E}$ of the scan:

$$\frac{1}{|\mathcal{E}|} \sqrt{\sum_{e \in \mathcal{E}} \frac{(l_e - L_e)^2}{L_e^2}}$$

This measure does (intentionally) not penalize bending, but is sensitive to stretching and shearing. Obviously a certain amount of distortion is required to match the scan to the target model. However, for two methods reaching a similar fitting accuracy, the one causing fewer distortion should be the preferred solution.

From a visual inspection of the results (Figure 7), as well as the alignment error and geometric distortion reported in Table 3, we conclude that the approach of Huang et al. [25] yields the least accurate results for the given datasets. These results seem to confirm the considerations with regard to the fine-to-coarse node clustering: Although it can grant satisfying results for objects undergoing homogeneous, smooth deformations, it can be a limiting factor when the deformation becomes more heterogeneous, or when small portions of data need to be accurately registered. While its computational performance greatly outperforms the other approaches, the alignment performance proved to be unsatisfactory for the given datasets, regardless of the number of nodes employed (as testified by the Torso and Pillow datasets, for which said technique employs a greater and a comparable number of nodes as the other approaches).

We compared our approach ($\mathcal{D}$) to the results obtained by Li et al. [12], employing a similar number of nodes as the ones declared by their technique. While for the Sole and Mould datasets our approach seemingly reaches better visual results, in the Torso case Li's method achieves a better deformation. For the Pillow dataset Li's method fails to reach any plausible result, while our approach achieves a meaningful alignment. In general, our alignment distances are similar to those of Li et al., with the exception of the Pillow dataset, which
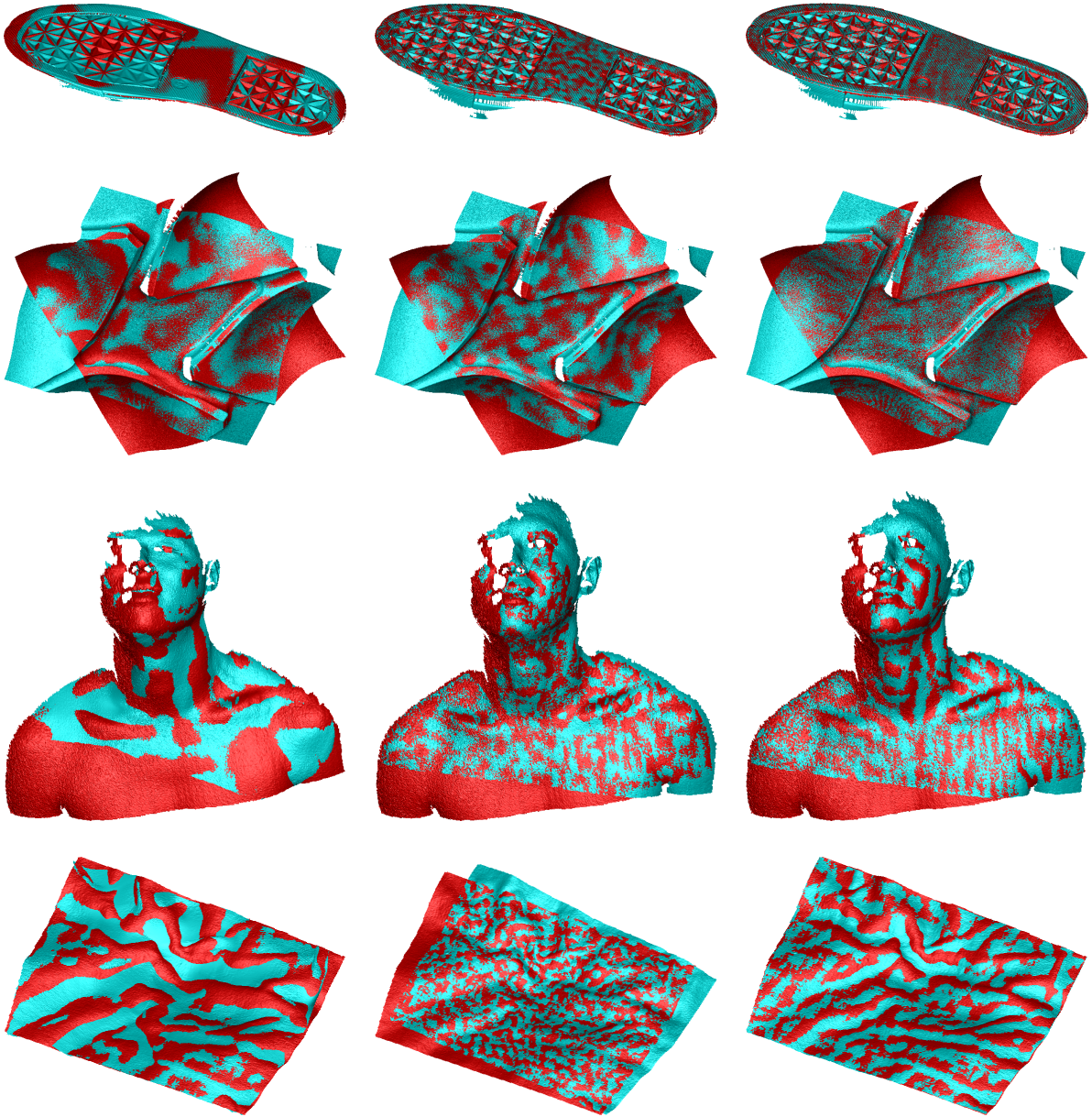
Figure 7: Comparison to other approaches by means of visual inspection of the deformed data (cyan) and the target model (red) (same datasets as in Figure 5). The results of each technique are shown in a separate column, from left to right: Huang et al. [25], Li et al. [12], and our approach (downgraded for fair comparison, $\mathcal{D}$). Datasets from top to bottom: Sole, Mould, Torso, Pillow.

should not be considered because of the alignment failure. In terms of distortion our method performs slightly but consistently better, which might be due to our use of exact per-node rotations and dual quaternion interpolation. With respect to computational performance, our technique is about twice as fast as Li's method for all the datasets with the exception of the Torso, for which it is just slightly faster. We assume that this is because

the Torso dataset is a smaller one, requiring the smallest number of nodes. For more complex deformations the difference between our 6 DoFs to their 12 DoFs per node seem to have a stronger influence.

In conclusion our method compares favorably to the methods of Huang et al. [25] and Li et al. [12], in particular when it comes to the precise registration of high-resolution range scans.
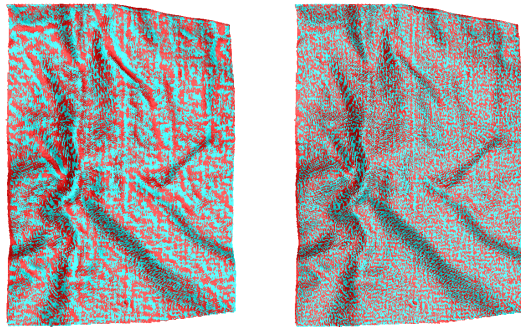
Figure 8: Alignment results for varying node densities. Left: $|\mathcal{N}|$ = 2400, 55 seconds. Right: $|\mathcal{N}|$ = 9600, 10 minutes.

5.4. Limitations and future work

The proposed approach improves upon closely related techniques in terms of fitting accuracy, geometric distortion, and computational performance (for bigger datasets). However, there are also some limitations. One major limitation is related to the fact that the method cannot handle topological changes, such as an opening mouth. This is mostly due to the adopted formulation for the smoothing energy that constraints neighbor nodes to deform similarly, as well as due to the interpolation policy adopted. A possible way to compensate for this would be to include the interpolation weights in the error functional, modifying their computation so that they can convey a topology-aware neighborhood information.

Another issue of the current approach is related to its scalability. As Figure 8 testifies, our approach could achieve an even better alignment of the Pillow dataset, but the increased number of nodes noticeably increase its computational cost. A possible solution would be a more efficient error-driven node placement or refinement: Through the estimate of an average error associated to each node, one could locate which areas of the surface need a greater node density, and only refine the discretization there, rather than imposing a uniform node density on the entire scan. This would reduce the overall number of nodes, therefore speeding up the computational performance.

Finally, we would like to extend our approach in such a way as to treat multiple-view deformable registration.

6. Conclusions

In this work we presented an improved solution for the non-rigid registration of partial views which is suitable for deforming object acquisition and modeling, especially for high-resolution surfaces acquired through professional 3D scanners. The method demonstrated its capability of suitably handling different degrees of deformation for all the considered challenging datasets. The selection of a denser set of correspondences with respect to the number of node patches defined on the surface granted our approach an improved accuracy with respect to state-of-the-art techniques for high-resolution datasets, while the adoption of a fast and robust optimization scheme, as well as a lightweight correspondence selection procedure granted satisfactory computational performance. With the complex of the described features, our technique is capable to provide physically-plausible, as-rigid-as-possible deformations with accurate and effective detail-aware registration properties which are valuable for many demanding and emerging applications.

References

[1] W. Chang, H. Li, N. Mitra, M. Pauly, M. Wand, Geometric registration for deformable shapes, in: Eurographics 2010 Course Notes, 2010.

[2] W. Chang, H. Li, N. J. Mitra, M. Pauly, S. Rusinkiewicz, M. Wand, Computing Correspondences in Geometric Data Sets, in: Eurographics 2011 Course Notes, 2011.

[3] O. van Kaick, H. Zhang, G. Hamarneh, D. Cohen-Or, A Survey on Shape Correspondence, Computer Graphics Forum 30 (6) (2011) 1681–1707, ISSN 1467-8659.

[4] S. Rusinkiewicz, O. Hall-Holt, M. Levoy, Real-time 3D model acquisition, ACM Transactions on Graphics 21 (3) (2002) 438–446, ISSN 0730-0301.

[5] T. Weise, B. Leibe, L. Van Gool, Fast 3D Scanning with Automatic Motion Compensation, in: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), ISSN 1063-6919, 1–8, 2007.

[6] Z. Zhang, Microsoft Kinect Sensor and Its Effect, IEEE Multi-Media 19 (2) (2012) 4–10, ISSN 1070-986X.

[7] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, A. Fitzgibbon, KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera, in: Proc. of ACM Symposium on User Interface Software and Technology (UIST), ISBN 978-1-4503-0716-1, 559–568, 2011.

[8] D. Hahnel, S. Thrun, W. Burgard, An extension of the ICP algorithm for modeling nonrigid objects with mobile robots, in: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI), 915–920, 2003.

[9] N. J. Mitra, S. Flöry, M. Ovsjanikov, N. Gelfand, L. Guibas, H. Pottmann, Dynamic geometry registration, in: Proc. of Symposium on Geometric Processing, ISBN 978-3-905673-46-3, 173–182, 2007.

[10] T. Popa, I. South-Dickinson, D. Bradley, A. Sheffer, W. Heidrich, Globally Consistent Space-Time Reconstruction, Computer Graphics Forum 29 (5) (2010) 1633–1642, ISSN 1467-8659.

[11] M. Wand, B. Adams, M. Ovsjanikov, A. Berner, M. Bokeloh, P. Jenke, L. Guibas, H.-P. Seidel, A. Schilling, Efficient reconstruction of nonrigid shape and motion from real-time 3D scanner data, ACM Transactions on Graphics 28 (2) (2009) 15:1–15:15, ISSN 0730-0301.

[12] H. Li, B. Adams, L. J. Guibas, M. Pauly, Robust single-view geometry and motion reconstruction, ACM Transactions on Graphics 28 (5) (2009) 175:1–175:10, ISSN 0730-0301.

[13] L. Ikemoto, N. Gelfand, M. Levoy, A hierarchical method for aligning warped meshes, in: Proc. 3-D Digital Imaging and Modeling (3DIM), 434–441, 2003.

[14] M. Wand, P. Jenke, Q. Huang, M. Bokeloh, L. Guibas, A. Schilling, Reconstruction of deforming geometry from time-varying point clouds, in: Proc. of Symposium on Geometric Processing Symp. on Geometric Processing, ISBN 978-3-905673-46-3, 49–58, 2007.

[15] B. J. Brown, S. Rusinkiewicz, Global non-rigid alignment of 3-D scans, ACM Transactions on Graphics 26 (3) (2007) 21:1–21:9, ISSN 0730-0301.

[16] T. Weise, T. Wismer, B. Leibe, L. V. Gool, Online loop closure for real-time interactive 3D scanning, Computer Vision and Image Understanding 115 (5) (2011) 635–648, ISSN 1077-3142.

[17] V. Blanz, T. Vetter, A morphable model for the synthesis of 3D faces, in: Proceedings of SIGGRAPH 99, Computer Graphics Proceedings, Annual Conference Series, ISBN 0-201-48560-5, 187–194, 1999.

[18] M. Pauly, N. J. Mitra, J. Giesen, M. Gross, L. J. Guibas, Example-based 3D scan completion, in: Proc. of Symposium on Geometric Processing, 23–32, 2005.

[19] A. M. Bronstein, M. M. Bronstein, R. Kimmel, Generalized multidimensional scaling: A framework for isometry-invariant partial surface matching, Proceedings of National Academy of Sciences (PNAS) 103 (5) (2006) 1168–1172.

[20] B. Allen, B. Curless, Z. Popović, The space of human body shapes: Reconstruction and parameterization from range scans, ACM Transactions on Graphics 22 (3) (2003) 587–594, ISSN 0730-0301.

[21] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, J. Davis, SCAPE: Shape completion and animation of people, ACM Transactions on Graphics 24 (3) (2005) 408–416, ISSN 0730-0301.

[22] B. Amberg, S. Romdhani, T. Vetter, Optimal Step Nonrigid ICP Algorithms for Surface Registration, in: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), ISSN 1063-6919, 1–8, 2007.

[23] M. F. Beg, M. I. Miller, A. Trouvé, L. Younes, Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms, International Journal of Computer Vision 61 (2) (2005) 139–157.

[24] H. Li, R. W. Sumner, M. Pauly, Global correspondence optimization for non-rigid registration of depth scans, in: Proc. of Symposium on Geometric Processing, 1421–1430, 2008.

[25] Q.-X. Huang, B. Adams, M. Wicke, L. J. Guibas, Non-rigid registration under isometric deformations, in: Proc. of Symposium on Geometric Processing, 1449–1457, 2008.

[26] M. Ovsjanikov, Q. Merigot, F. Memoli, L. Guibas, One Point Isometric Matching with the Heat Kernel, Computer Graphics Forum (Proc. SGP) 29 (5) (2010) 1555–1564.

[27] Y. Lipman, T. Funkhouser, Möbius voting for surface correspondence, ACM Transactions on Graphics 28 (3) (2009) 72:1–72:12.

[28] V. G. Kim, Y. Lipman, T. Funkhouser, Blended intrinsic maps, ACM Transactions on Graphics 30 (4) (2011) 79:1–79:12.

[29] R. W. Sumner, J. Schmid, M. Pauly, Embedded deformation for shape manipulation, ACM Trans. Graph. 26 (3) (2007) 80:1–80:7, ISSN 0730-0301.

[30] M. Botsch, M. Pauly, M. Wicke, M. Gross, Adaptive Space Deformations Based on Rigid Cells, Computer Graphics Forum 26 (3) (2007) 339–347.

[31] H. Pottmann, S. Leopoldseder, M. Hofer, Simultaneous Registration of Multiple Views of a 3D Object, Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (3A) (2002) 265–270.

[32] M. Botsch, M. Pauly, M. Gross, L. Kobbelt, PriMo: coupled prisms for intuitive surface modeling, in: Proc. of Symposium on Geometric Processing, ISBN 3-905673-36-3, 11–20, 2006.

[33] J. P. Lewis, M. Cordner, N. Fong, Pose space deformation: a unified approach to shape interpolation and skeleton-driven deformation, in: Proceedings of ACM SIGGRAPH 2000, Computer Graphics Proceedings, Annual Conference Series, ISBN 1-58113-208-5, 165–172, 2000.

[34] L. Kavan, S. Collins, J. Žára, C. O'Sullivan, Geometric skinning with approximate dual quaternion blending, ACM Transactions on Graphics 27 (4) (2008) 105:1–105:23, ISSN 0730-0301.

[35] S. Rusinkiewicz, M. Levoy, Efficient variants of the ICP algorithm, in: Proc. of International Conference on 3D Digital Imaging and Modeling (3DIM), 145–152, 2001.