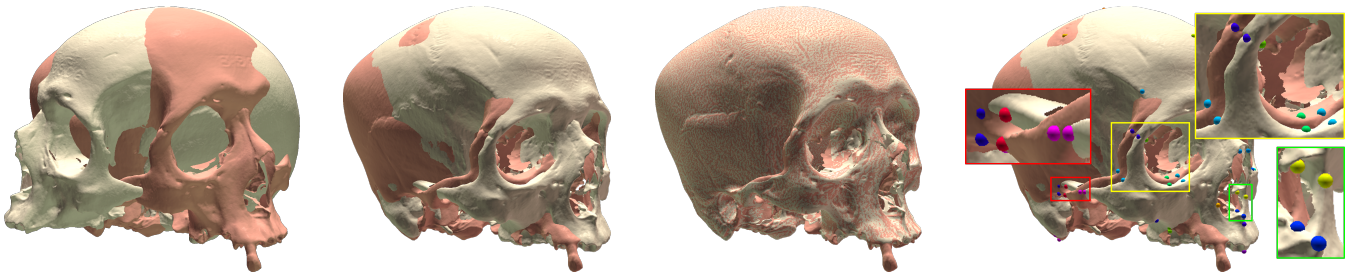
EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2017)
R. Sablatnig and B. Štular (Editors)



# Automatic Transfer of Landmarks on Human Skulls using GPU-based Non-rigid Registration

Matthias Innmann[1] Philipp Erhardt[1] Daniel Schütz[2] Günther Greiner[1]

[1]Universität Erlangen-Nürnberg
[2]Universität Kiel



**(a)** *Initial poses* **(b)** *Rigid alignment* **(c)** *Non-rigid registration* **(d)** *Original and transferred landmarks*

**Figure 1:** *Our algorithm to automatically transfer landmarks: After the acquisition, the initial pose is arbitrary (a). We first align the geometry rigidly (b) and then remove deviations by non-rigid registration (c). After that, all landmarks can easily be transferred (d).*

**Abstract**

*In this work, we present a novel approach to automatically transfer landmarks from a template mesh of a human skull to other meshes obtained via 3D scanning. Since previous methods rely on user input or only work on a subset of the data, these algorithms are not suited for large databases. Our system is designed to work for arbitrary meshes of human skulls, i.e. having artifacts like incomplete geometry or being non-watertight. Since the input data has no common orientation, we first apply a rigid coarse registration followed by a refinement. Afterwards, the remaining geometric deviation is removed by non-rigidly deforming one mesh into the other. With this precise geometric mapping, arbitrary landmarks can be transferred easily. To ensure efficient computation, we use a highly optimized GPU implementation to solve arising optimization problems. We apply our method to a dataset consisting of 1200 models acquired via structured light scanning and evaluate its accuracy on a subset of these models.*

## 1. Introduction

When analyzing the ancient populations and their movements across the world, statistics of several landmarks on the human skull play an important role [HW06] [GF16] [RBS*].

Nowadays, this analysis is often done in virtual environments, i.e. using an acquired 3D model of the skull. Advantages of virtual exploration include reproducibility, the possibility to work remotely and the availability of a huge set of software tools for measuring, editing, modeling, etc. In order to use software for 3D data processing, the physical object is scanned using standard 3D capturing methods, e.g. via structured light.

To obtain the statistics of the chosen landmarks, they have to be transferred to every acquired skull model. In general, this is a tedious and time-consuming task, since many landmarks, e.g. 60, have to be placed manually by the user. For a large database, e.g. consisting of $> 1000$ models, this task will take several weeks or months to complete and will also be prone to mistakes.

In this work, we present an algorithm that is able to transfer landmarks of a template mesh to an arbitrary input mesh without user interaction. The meshes do not need to be watertight, can contain duplicate or missing geometry and can have arbitrary initial orientations. After the processing, the landmarks can be placed at arbitrary positions on the template, since we establish dense correspondences from one model to the other. As our highly optimized algorithm runs the costly tasks on the GPU, the processing time for one mesh is only 3.7 seconds on average.



DOI: 10.2312/gch.20171304

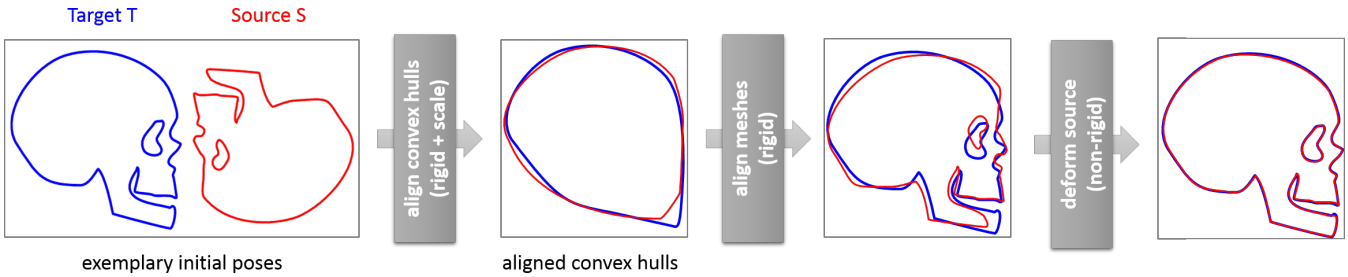



**Figure 2:** *Method overview: from an arbitrary initial pose, we first align the convex hulls of the models via a rigid transform including scale. After that, the rigid alignment is refined using the complete meshes. To remove the remaining geometric deviation, we perform a non-rigid deformation to warp the source mesh into the target. Our algorithm works on 3D data, but for illustration purposes we use 2D figures.*

## 2. Related Work

### 2.1. Transfer of Landmarks

Since several landmarks play an important role in the field of landmark-based geometric morphometrics and physical anthropology [HW06] [GF16] [RBS*], there has been some prior work.

Cheng et al. [CLL12] identify the Frankfurt plane and the mid-sagittal plane of the human skull to define important landmarks. However, it is only possible to find landmarks on these two planes.

Zhang et al. [ZCL13] also use non-rigid registration techniques to establish dense correspondences between two models. Although this approach seems quite similar, they need to identify anatomical landmarks as hard-constraints to ensure robust results. In case of noisy or missing geometry, these anatomical landmarks may not be observable in the 3D geometry or may not even exist at all.

### 2.2. Non-rigid Registration

The field of non-rigid registration offers a wide variety of previous work, which would be out of scope to cover sufficiently in this work. A survey of different techniques can be found in [TCL*13].

One of the most popular approaches to model non-rigid deformation is the as-rigid-as-possible (ARAP) energy [SA07], originally designed to work on triangular meshes. It is an elegant approach that operates directly on the surface, therefore does not require any skeleton or embedding cage. The global deformation is smooth, while local geometric details are preserved. Since its energy is defined locally, the problem generates a very sparse matrix that can be solved efficiently. In previous work in the context of real-time 3D scanning, this energy has been applied to *track* known [ZNI*14] or even *scan* unknown deforming objects [IZN*16].

## 3. Method Overview

The input to our method is a template mesh of a human skull and a set of different meshes of human skulls. The goal is to place landmarks on the template, while the algorithm automatically transfers these marks to the whole set of other models (i.e., a database). The models are usually acquired via a common 3D scanner, e.g. using structured light or laser scanning. Therefore, they may contain geometric artifacts, e.g. holes, overlapping geometry, etc. and will not be watertight. Solving these issues for a whole dataset of, e.g. 1200 models, would be an enormously time-consuming preprocessing task. We design our algorithm to work in the presence of these problems.

We will refer to the template mesh as **target T** and the mesh of the database as **source S**.

Our pipeline mainly consists of three steps (see also Figure 2) that do not require any user input:

1. Compute a coarse rigid alignment between **T** and **S** using their convex hulls (Section 4.1).
2. Refine the rigid alignment via point-to-plane ICP (Section 4.2).
3. Remove the remaining difference by non-rigidly deforming **S** (Section 5).

Having successfully warped the source into the target, the valid parts of the meshes, i.e. no holes, are now geometrically identical. Therefore, it is now trivial to transfer any landmark to the deformed source and thus also to the undeformed source. The user can place a landmark onto the target, which is instantly transferred to the source.

In order to efficiently solve the non-linear optimization problem with up to 6M variables that arises in Section 5, we use a hand-crafted GPU implementation that is able to solve these problems in a few seconds.

## 4. Rigid Alignment

The goal of this step is to find the optimal rigid alignment $[\mathbf{R},\mathbf{t}]$ between the **target T** and the **source S**.

Given a set of $N$ correspondences $\{(\mathbf{p}_i,\mathbf{q}_i)\}_{i=1,\dots,N}$, the point-to-plane error metric defines the following optimization problem:

$$\arg\min_{\mathbf{R},\mathbf{t}} \sum_{i=1}^{N} \langle \mathbf{R}\mathbf{p}_i + \mathbf{t} - \mathbf{q}_i, \mathbf{n}_i \rangle^2, \tag{1}$$

where $\mathbf{n}_i$ is the normal vector at point $\mathbf{q}_i$.

We use the point-to-plane error metric, since it yields better convergence properties than the point-to-point error metric in our experiments, which is consistent with literature [RL01].

In both of the following steps (Section 4.1 and 4.2), we employ a standard ICP approach that is repeated until convergence:

1. Find a set of correspondences $\{(\mathbf{p}_i,\mathbf{q}_i)\}_{i=1,\dots,N}$.
2. Solve Equation 1.

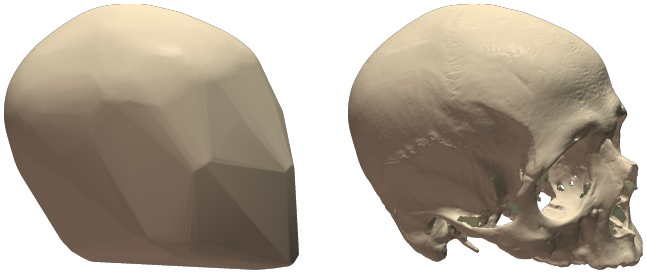
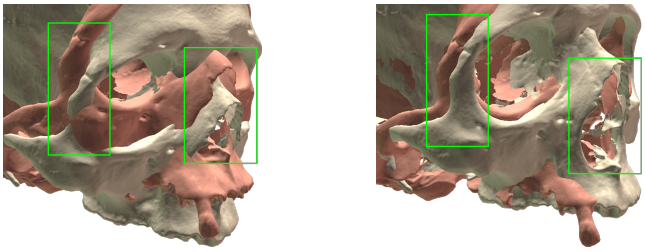


**Figure 3:** *The convex hull (left) of a skull model (right) is used to compute a coarse alignment.*



**(a)** *Coarse alignment* **(b)** *Fine registration*

**Figure 4:** *Refining the rigid transformation.*

### 4.1. Coarse Alignment

Since ICP on the complete meshes will fail, i.e. it will only find a local minimum, if the deviations are too large, we first align the convex hulls of the meshes to achieve a coarse registration.

The idea is that the convex hull of every skull is very similar, because the anatomy of human skulls is strongly related.

First, we scale the convex hulls such that their volumes are equal. Second, we compute an initial solution for the optimal translation $\mathbf{t}$ by aligning the centroids of the convex hulls.

Given the fact that the skulls show similar structure, we exploit this by computing the principal axes of the meshes via principal component analysis (PCA). There exist 24 possible combinations of valid right-handed coordinate systems, 6 for choosing the $x$-axis, 4 remain for the $y$-axis, the $z$-axis is then uniquely defined. For each of these possibilities as initial transformation, we apply point-to-plane ICP: We estimate the best rigid transform $[\mathbf{R},\mathbf{t}]$ that minimizes Equation 1 for a set of nearest-neighbor correspondences. For every point on the convex hull of $\mathbf{S}$, we choose the nearest neighbor (in terms of Euclidean distance) on the convex hull of $\mathbf{T}$ as correspondence. A benefit of the convex hull is the significant reduction in terms of complexity: While the complete meshes have 500K to 1M vertices, the convex hulls only have several thousands of vertices.

We discard the worst half of the results, i.e. 12 combinations, and repeat this process until only one combination is left.

### 4.2. Fine Registration

Having estimated a coarse alignment of $\mathbf{S}$ and $\mathbf{T}$ using their convex hulls, we wish to refine this transformation (see Figure 4). To this end, we employ ICP on the full geometry, e.g. up to 1M vertices.

For every vertex of $\mathbf{S}$, we find the nearest-neighbor (Euclidean distance) on $\mathbf{T}$ and choose this point as correspondence. Depending on the geometry of the meshes and the initial alignment, we need 1 to 15 ICP iterations until convergence.

## 5. Non-rigid Alignment

After having found the best rigid alignment, we remove the remaining deviation of the meshes by applying non-rigid ICP. To this end, we use the well-known ARAP deformation energy [SA07]. It ensures that local geometric details are kept while the overall deformation is smooth.

The energy between the undeformed mesh $S$ and the deformed mesh $S'$ is defined as:

$$E(S,S') = \sum_{i=1}^{M} \sum_{j\in N(i)} \left\| \mathbf{p}'_i - \mathbf{p}'_j - \mathbf{R}_i(\mathbf{p}_i - \mathbf{p}_j) \right\|_2^2, \qquad (2)$$

where $N(i)$ is the set of neighbors at vertex $\mathbf{p}_i$.

In each step of the non-rigid ICP, we first find a set of correspondences $\{(\mathbf{p}_i,\hat{\mathbf{p}}_i)\}_{i=1,\dots,N}$ by employing a radius search. I.e., we store all correspondence candidates inside a sphere with fixed radius $r$ and assign the following cost function $C$:

$$C(\mathbf{p},\hat{\mathbf{p}},\mathbf{n},\hat{\mathbf{n}}) = (1 - \langle \mathbf{n},\hat{\mathbf{n}} \rangle) \cdot w_{\text{normal}} + \frac{\|\mathbf{p}-\hat{\mathbf{p}}\|_2}{r} \cdot w_{\text{distance}}, \qquad (3)$$

where $\mathbf{p},\hat{\mathbf{p}}$ are the positions on source and target, respectively. $\mathbf{n},\hat{\mathbf{n}}$ are the corresponding normal vectors at these points. In our experiments, we found that $r = 0.07$ provides a good trade-off between complexity and accuracy. We chose $w_{\text{normal}} = 6, w_{\text{distance}} = 1$ to ensure that local geometric features guide the correspondence association. To further reduce computational complexity, we only use every 40th vertex on $\mathbf{S}$ for correspondence search. From the candidate set, we chose the pair with minimum cost $C$. In the last iteration, every vertex is used for nearest-neighbor correspondence association to ensure smooth results.

Having computed a set of correspondences, we establish the following global optimization problem:

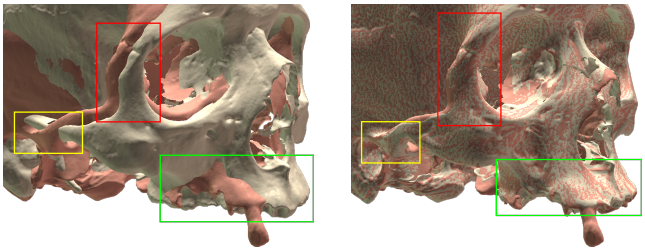
$$E_{\text{tot}} = w_{\text{fit}} \cdot E_{\text{fit}} + w_{\text{reg}} \cdot E_{\text{reg}}, \qquad (4)$$

where $E_{\text{fit}}$ represents the fitting constraint for the set of correspondences $\{(\mathbf{p}'_i,\hat{\mathbf{p}}_i)\}_{i=1,\dots,N}$ $E_{\text{fit}} = \sum_{i=1}^{N} \langle \mathbf{p}'_i - \hat{\mathbf{p}}_i, \mathbf{n}_i \rangle^2$ and $E_{\text{reg}}$ is the ARAP regularization constraint (Equation 2). In our implementation, we use $w_{\text{fit}} = 0.5$ and $w_{\text{reg}} = 1$.

To solve the optimization problem (Equation 4), we follow the authors of [SA07] and [IZN*16] and decompose the problem into two sub-problems that are solved in a flip-flop manner:

1. For fixed rotations $\mathbf{R}_i$, the problem is linear in the positions.
2. For fixed positions $\mathbf{p}'_i$, the optimal rotations $\mathbf{R}_i$ can be found via singular value decomposition at each vertex:

$$\sum_{j\in N(i)} (\mathbf{p}_i - \mathbf{p}_j)(\mathbf{p}'_i - \mathbf{p}'_j)^T = \mathbf{U}_i\mathbf{S}_i\mathbf{V}_i^T, \qquad \mathbf{R}_i = \mathbf{U}_i\mathbf{V}_i^T.$$



**(a)** *Rigid alignment* **(b)** *Non-rigid alignment*

**Figure 5:** *We use non-rigid ICP to remove remaining deviations.*

Since the optimization problem contains up to 6M variables, we solve the arising sparse linear system for the positions $\mathbf{p}_i'$ using a hand-crafted preconditioned conjugate gradient (PCG) solver on the GPU. Our implementation is able to solve the problem on average in less than 2 seconds (see Table 1).

## 6. Landmark Transfer

Since the geometry of the source $\mathbf{S}$ is now completely warped into the target $\mathbf{T}$, landmarks can simply be transferred by projecting the landmark onto the other mesh.

### 6.1. Outlier Removal

Depending on the geometry of $\mathbf{S}$, there may exist points on $\mathbf{T}$ that have no valid correspondence on $\mathbf{S}$, e.g. in case of missing geometry (missing teeth, zygomatic bones, jaw etc.) (e.g. see Figure 7).

We identify these outliers by analyzing the distances of corresponding landmark pairs $(\mathbf{s}_i, \mathbf{t}_i)$. Successfully transferred landmarks have a small Euclidean distance $\|\mathbf{s}_i - \mathbf{t}_i\|_2$, whereas incorrect pairs lead to larger distances. In our experiments, we chose a threshold of 1 cm for pruning.

## 7. Results

We quantitatively evaluate the results of our proposed method on a large dataset of human skulls. The 1200 human skulls (individuals from the Southern Pacific islands), which are stored in the natural history museums in Paris, London, and Vienna, have been scanned with a structured-light scanner (David SLS-2) by one of us in 2014. For the quantitative evaluation, the models have been annotated by an expert with 16 precisely chosen landmarks on each skull.

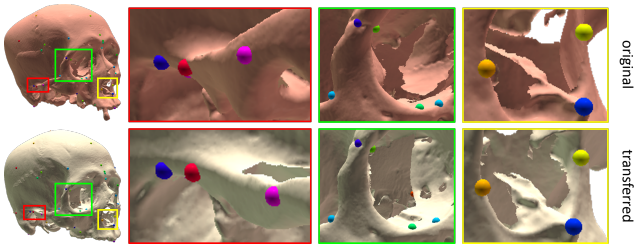



**Figure 6:** *Automatically transferred landmarks.*

Table 1 shows the runtimes for the computation of the convex hull (*hull*), the rigid registration (*rigid*) and the non-rigid alignment (*non-rigid*) as well as error measurements on our dataset.

| | Runtime (s) | | | | Error (mm) | |
|---|---|---|---|---|---|---|
| | hull | rigid | non-rigid | total | w/o pruning | w/ pruning |
| min | 0.07 | 0.19 | 0.54 | 0.98 | $8.01 \cdot 10^{-5}$ | $8.01 \cdot 10^{-5}$ |
| max | 1.16 | 4.03 | 5.10 | 9.57 | 45.3 | 33.6 |
| avg | 0.46 | 0.97 | 1.98 | 3.71 | 3.78 | 3.76 |

**Table 1:** *Runtimes and landmark errors on our dataset, measured on a system with an Intel i7-6700 CPU and an NVIDIA GTX 1080.*

Figure 7 shows the error distribution of transferred landmarks. 50 % of the landmarks have an error $< 2.8$ mm, 90 % below 8 mm.

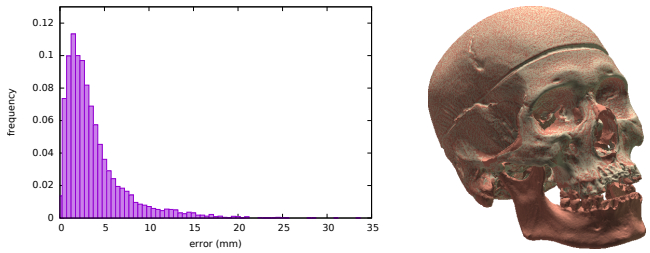



**Figure 7:** *Left: Error distribution of transferred landmarks (with outlier removal), Right: Successful registration result in case of missing geometry (lower jaw missing in source).*

## 8. Conclusion

In this work, we developed a system for automatic transfer of arbitrary landmarks on human skulls. The input to our method is a template mesh of a skull and a database of skull meshes. The database meshes do not have to satisfy complex requirements like watertight topology, they may also contain holes or miss geometry, e.g. teeth. Since the models will have arbitrary orientation after acquisition, we first compute a rough alignment based on the convex hulls and principal component analysis. We refine the alignment using classic point-to-plane ICP and remove remaining deviations in the geometry by non-rigidly deforming the database model to best fit the template. After that, the transfer of arbitrary landmarks is straightforward. We take care of missing geometry by analyzing the landmark pair distances and prune outliers.

In cases where the transfer precision is not sufficient for anthropological research, the landmarks can easily be corrected by hand. The algorithm can also be used as initialization for further processing and optimization.

To efficiently solve the arising optimization problems, we use a hand-crafted implementation on the GPU that provides an average processing time of 3.7 seconds for one mesh. Thus, our system allows for efficient application on large datasets, e.g. consisting of $> 1000$ meshes, with moderate overall runtime.

Our method alleviates tedious, time-consuming and error-prone work in the field of anthropology, where researchers often have to manually place many meaningful landmarks, e.g. 60, on each skull.

## References


[CLL12] CHENG Y., LEOW W. K., LIM T. C.: Automatic identification of frankfurt plane and mid-sagittal plane of skull. In *2012 IEEE Workshop on the Applications of Computer Vision (WACV)* (Jan 2012), pp. 233–238. doi:10.1109/WACV.2012.6162994. 2

[GF16] GALLAND M., FRIESS M.: A three-dimensional geometric morphometrics view of the cranial shape variation and population history in the new world. *American Journal of Human Biology 28*, 5 (2016), 646–661. 1, 2

[HW06] HARVATI K., WEAVER T. D.: Human cranial anatomy and the differential preservation of population history and climate signatures. *The Anatomical Record 288*, 12 (2006), 1225–1233. 1, 2

[IZN*16] INNMANN M., ZOLLHÖFER M., NIESSNER M., THEOBALT C., STAMMINGER M.: *VolumeDeform: Real-Time Volumetric Non-rigid Reconstruction*. Springer International Publishing, Cham, 2016, pp. 362–379. URL: http://dx.doi.org/10.1007/978-3-319-46484-8_22, doi:10.1007/978-3-319-46484-8_22. 2, 3

[RBS*] RASSKASOVA A., BEREZINA N., SOFICARU A. D., ET AL.: Resolving relationships between several neolithic and mesolithic populations in northern eurasia using geometric morphometrics. *American Journal of Physical Anthropology*. 1, 2

[RL01] RUSINKIEWICZ S., LEVOY M.: Efficient variants of the icp algorithm. In *Proceedings Third International Conference on 3-D Digital Imaging and Modeling* (2001), pp. 145–152. doi:10.1109/IM.2001.924423. 2

[SA07] SORKINE O., ALEXA M.: As-rigid-as-possible surface modeling. In *EUROGRAPHICS SYMPOSIUM ON GEOMETRY PROCESSING* (2007). 2, 3

[TCL*13] TAM G. K., CHENG Z.-Q., LAI Y.-K., LANGBEIN F. C., LIU Y., MARSHALL D., MARTIN R. R., SUN X.-F., ROSIN P. L.: Registration of 3d point clouds and meshes: a survey from rigid to nonrigid. *IEEE transactions on visualization and computer graphics 19*, 7 (2013), 1199–1217. 2

[ZCL13] ZHANG K., CHENG Y., LEOW W. K.: *Dense Correspondence of Skull Models by Automatic Detection of Anatomical Landmarks*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 229–236. URL: http://dx.doi.org/10.1007/978-3-642-40261-6_27, doi:10.1007/978-3-642-40261-6_27. 2

[ZNI*14] ZOLLHÖFER M., NIESSNER M., IZADI S., REHMANN C., ZACH C., FISHER M., WU C., FITZGIBBON A., LOOP C., THEOBALT C., ET AL.: Real-time non-rigid reconstruction using an rgb-d camera. *ACM Transactions on Graphics (TOG) 33*, 4 (2014), 156. 2