Correspondence problem under isometric transformation as nested linear assignment problems

Anonymous Author(s)
Affiliation
Address
email

Abstract

Three-dimensional motion tracking has enjoyed various applications in computer graphics and other fields. One of the technical issues common in optical motion capture systems is time-consuming manual data correction for partially missing points and mis-tracking of objects. There has been several existing algorithms for recovering a point-to-point correspondence between two non-aligned point-sets at two distinct time frames. The many of these algorithms explicitly estimate transformation from one set of points to the other, but they often result in sub-optimal solution. Unlike this existing approach, we propose to exploit the distance matrix that is invariant under isometric transformation. This distance-based approach takes a form of quadratic assignment problem, that is efficiently approximated by a nested form of linear assignment problems. In numerical experiments, the proposed nested Hungarian algorithm found the one-to-one correspondence between point-sets with missing points and mis-tracking more accurately than the alternative algorithms.

1 Motion tracking by optical motion capture systems

Three-dimensional motion tracking has enjoyed various applications in the field of computer graphics and biomechanics. In computer graphics applications, such as video games, films, and virtual reality, the motion dataset of human subjects is used to produce realistic motion of the virtual character. In biomechanics, sports science, and their related research fields, recording human or animal movements provides the primary data source for the basic research on kinematic characteristics of their movements.

The motion capture system records the shape/posture of a subject at each time point as a collection of points in the three-dimensional coordinates, (x, y, z), each represents the location of a marker attached on the subject. Among various motion capture technologies, optical motion capture systems, composed of multiple infra-red ray cameras and reflective markers, is of a standard choice to capture fine-grained motion patterns due to its high frame rate (over 240 frames per second). In principle, however, it has the technical limitation such that markers would be frequently dropped from the sample due to occluding objects. Since infra-red rays cannot pass through the subject's body, the optical systems cannot detect reflective markers being occluded by the subject's body or something occluders. The other problem is that the markers are often mis-tracked or switched over, when two markers pass similar spatial and temporal point.

These two types of errors, dropping and mis-tracking, cause considerable cost in data cleaning, often made manually, before any data analysis of the data collected by the optical motion capture system. In practice, data cleaning, interpolation of the missing markers and identification of markers, is often manually performed and eyeball checked. It is time-consuming (e.g., 40 markers captured for 10 minutes in 240 Hz produce 5,760,000 points), and it is often a bottle-neck of the work flow in any use of motion data. The goal of this study is to propose an automated preprocessing algorithm for this class of problems.

In 1970s, the researchers of computer vision have formulated the motion tracking as the so-called *correspondence problem* [10]. Since then, the correspondence problem is one of active research topics in computer vision. The correspondence problem is to find a one-to-one correspondence between the two given sets of points (point-sets). In the motion tracking case, each point-set includes multiple points, where each marker at a time frame is represented as a point in the 3D coordinates, and motion tracking is to find a one-to-one correspondence between points across the two point-sets at distinct time frames.

A naive algorithm for the correspondence problem would be to match a nearest-neighbor point in a point-set to each point in the other one, under the assumptions that the two point-sets have the same number of points, and the point-set has points quite close to one of them in the other point-set. In practice, this algorithm may work well, if this assumption holds (that is likely, if two point-sets are the consecutive samples at high frame rate). With some markers missing in either/ both of the point-sets, however, this algorithm can fail to find a one-to-one correspondence.

In practice, however, a number of markers keep missing for seconds or even longer interval, for example, when the subject's limbs occlude the markers attached on his/her torso. With such long-term missing data points, the naive algorithm mentioned above cannot work properly. A class of algorithms, e.g., [1], is motivated to find a partial point-to-point correspondence under a certain assumption on the global consistency of the object. For instance, the subject's body is supposed to be one or more rigid bodies, and a one-to-one correspondence between two point-sets is searched on the basis of point-to-point proximity under a certain isometric (distance-preserving) transformation.

In this paper, we present our new algorithm to find a partial correspondence between two point-sets under the assumption that the object is a rigid body. Our algorithm is motivated to exploit the characteristics of the point-set, invariant under isometric transformation. Specifically, our algorithm uses the $n \times n$ and $m \times m$ distance matrices of the point-set with n points and the other point-set with m points, respectively. These distance matrices are invariant under any isometric transformation up to a permutation, and thus, the correspondence problem is reduced to find a permutation for each of these distance matrices. This computational process is closely approximated by a set of linear assignment problems [4]. As each linear assignment problem is solved efficiently by the well-known Hungarian algorithm [7, 11], we call our algorithm *nested Hungarian algorithm*.

In Section 2, we briefly illustrate the correspondence problem. In Section 3, we introduce the special case of correspondence problem of a rigid body, and present the nested Hungarian algorithm. In Section 4, we evaluate the proposed algorithm in a series of numerical case studies by comparing it with the existing algorithms.

2 Correspondence for real-world 3D shapes as point-sets

2.1 The correspondence problem

Suppose that there are n reflective markers, each attached on a body part of a subject, and the markers are fixed on the same points on the body over time. We assume the subject's body is one or more rigid bodies, but not completely deformable matter.

Let $\mathcal{A} \subset \mathbb{R}^3$ and $\mathcal{B} \subset \mathbb{R}^3$ be two point-sets, and each has the n markers (as points) in a three-dimensional Euclidean space. We identify each marker in the point-set by the indices $i = 1, \dots, n$, and denote the i^{th} point in the point-set $\mathcal{A}$ by $\boldsymbol{a}_i \in \mathcal{A}$ (and the point-set $\mathcal{B}$ by $\boldsymbol{b}_i \in \mathcal{B}$). The point-sets $\mathcal{A}$ and $\mathcal{B}$ represent the two set of markers recorded at two distinct time frames, but the identity of each marker is not always preserved by the indices across the two point-sets.

For now, we suppose there is no missing marker: Every marker in the point-set $\mathcal{A}$ has exactly one corresponding marker in the point-set $\mathcal{B}$. We will relax this condition later. Denote by the matrix with the vector $\boldsymbol{a}_i$ at the i^{th} row by $A = (\boldsymbol{a}_i) \in \mathbb{R}^{n\times 3}$ (and with $\boldsymbol{b}_i$ at the i^{th} row by $B = (\boldsymbol{b}_i) \in \mathbb{R}^{n\times 3}$).

A one-to-one correspondence between the point-sets $\mathcal{A}$ and $\mathcal{B}$ is a bijective map $p : \{1, \dots, n\} \to \{1, \dots, n\}$, that satisfies a certain condition. In a matrix form, such a bijective map is equivalent to a permutation matrix $P \in \{0, 1\}^{n\times n}$, where the sum of all elements in every row and column is

1. For given A and B, the corresponding problem under an unknown isometric transformation is to find a pair of permutation P and isometric transform $f : \mathbb{R}^{n\times 3} \mapsto \mathbb{R}^{n\times 3}$ such that

$$Pf(A) \sim B, \tag{1}$$

where the operator $X \sim Y$ represents a class of similarity between matrices X and Y. Equation (1) implies that there are two inter-dependent problems in the "chicken-and-egg" relationship [1]. If a permutation P is given, one can find the isometric tranformation f satisfying (1). If f is given, one can find the P. However, if all P and f are simultaneously unknown, it is not trivial to find P and f of a desired condition.

2.2 Transformation-based and distance-based approach

In past literature, there are at least two major classes of approach to the correspondence problem (1). The first approach is to explicitly estimate the isometric transformation f as well as P in an iterative manner, which we call *transformation-based* approach. The second one is to estimate only the permutation P by using the distance (dissimilarity) matrices constructed from A and B, which we call *distance-based* approach. The transformation-based one is computationally cheaper, but it may result in sub-optimal solution, due to its iterative procedure for dual-minimization of both P and f [12]. The distance-based approach essentially avoids the chicken-and-egg problem and need to find only permutation P, as the distance matrices are invariant under any isometric transformation. It is, however, computationally extremely costly, as its exact search is known as a NP-complete problem.

In this study, we propose a computationally feasible algorithm as an approximation of the distance-based approach.

3 Correspondence under isometric transformations

3.1 Distance-based formulation of the correspondence problem

The key observation of the distance-based approach is the distance matrix $D(A) = \left(d^A_{i,j}\right)_{i,j\in\{1,\ldots,n\}} \in \mathbb{R}^{n\times n}$, where $d^A_{i,j} = \|A_i - A_j\|$ and $A_i \in \mathbb{R}^3$ is the i^{th} row of A, is invariant under isometric transformation f: i.e., $D(f(A)) = D(A)$. Thus, the correspondence problem (1) is reformulated to find a permutation $P \in \mathbb{R}^{n\times n}$ such that

$$PD(A)P^\top \sim D(B). \tag{2}$$

Introducing an error function, such as the summed squared error $E = \|PD(A)P^\top - D(B)\|$ where $\|X\| = \sum_{i,j} X^2_{i,j}$, it is a quadratic assignment problem to find the minimizer P of E, that is known as a NP-complete problem [3].

3.2 Nested linear assignment problems

Therefore, we reformulate the problem (2) by a hierarchical linear assignment problem. To illustrate our idea, consider the special case $PD(A)P^\top = D(B)$ with an exact correspondence between $\mathcal{A}$ and $\mathcal{B}$. In this case, with such permutation P, there is some permutation $Q \in \mathbb{R}^{n\times n}$ independent of P satisfies

$$PD(A)Q^\top = D(B), \tag{3}$$

and $Q = P$ is unique, if there is no exchange of any two rows in A that preserves the distance matrix $D(A)$. This means that there is permutation Q such that

$$(d^A_{j,1}, \ldots, d^A_{j,n})^\top = Q(d^B_{i,1}, \ldots, d^B_{i,n})^\top, \tag{4}$$

for all i, j such that $P_{i,j} = 1$ or i^{th} point in $\mathcal{B}$ corresponds with j^{th} point in $\mathcal{A}$. For each pair (i, j), Q in (4) is solved by a linear assignment problem

$$\hat{Q} := \arg\min_Q \operatorname{tr}\left(D_{i,j}Q\right), \tag{5}$$

where $D_{i,j} := (\|d^A_{j,k} - d^B_{i,l}\|)_{k,l\in\{1,\ldots,n\}} \in \mathbb{R}^{n\times n}$. Then the permutation P is given by another linear assignment problem

$$\hat{P} := \arg\min_P \operatorname{tr}\left(CP\right), \tag{6}$$

ALGORITHM 1: Nested Hungarian algorithm

Input: Two point-sets $A \in \mathbb{R}^{n\times 3}$ and $B \in \mathbb{R}^{m\times 3}$
Output: Correspondence P
$D(A) = (\boldsymbol{d}_i^A)$ is the distance matrix of A
$D(B) = (\boldsymbol{d}_j^B)$ is the distance matrix of B
$C \in \mathbb{R}^{n\times m}$
for $i = 1$ **to** n **do**
    for $j = 1$ **to** m **do**
        Compute $\left[\hat{Q}, c_{ij}\right]$ using $\mathrm{Hungarian}(D_{i,j})$
        Set the (i, j) element of C to c_{ij}
    end
end
Compute $[\hat{P}, c]$ using $\mathrm{Hungarian}(C)$

where $C = (c_{i,j}) \in \mathbb{R}^{n\times n}$ with $c_{i,j} := \min_Q \mathrm{tr}\,(D_{i,j}Q)$. This observation holds only if there is no exact match $PD(A)P^\top = D(B)$, but it reduces the original quadratic assignment problem (2) to the nested linear assignment problem of (5) and (6). As the linear assignment is solved by the Hungarian algorithm calling the function $O(n^3)$ times [4], this nested linear assignment problem is solved by $O(n^5)$ times of the function calls.

In summary, this procedure mentioned above is implemented by the nested Hungarian algorithm described in the psudo-code (Algorithm 1) for the point-set $\mathcal{A}$ with n points and $\mathcal{B}$ with m points. In the pseudo-code, the function $\mathrm{Hungarian}(\cdot)$ is an implementation of the Hungarian algorithm, computing the minimal total cost $c_{ij}$and the correspondence $\hat{Q}$. Note that the nested **for** loops can be computed independently. Thus, the construction of C in Algorithm 1 can be parallelized, and then its computational complexity will be reduced to $O(n^3)$.

4 Related works

In the following section, we will report numerical studies comparing the nested Hungarian algorithm with the representative existing algorithms. Here we briefly overview them. Some of the following algorithms for shape/feature matching is not developed specifically for the correspondence problem of interest, but for more general problems than it. Here we choose the existing algorithms according to their applicability to the correspondence problem under isometric transformation.

As an algorithm taking the transformation-based approach, Basl [1] has proposed Iterative closest point (ICP) algorithm that finds a locally minimal correspondence by iteratively searching the nearest-neighbor and the least-square rotation and translation. However, this algorithm often results in a local minimum, that is not sufficiently good in practice [12]. An extension of ICP algorithm [5] for non-rigid shape matching approximates non-rigid deformable subjects as a patchwork of small rigid segments.

As an algorithm taking the distance-based approach, Maciel [9] and Berg [3] have proposed an algorithm to match image patch of a set of pixels, by formulating it as a constrained concave programming problem. Leordeanu [8] has proposed the spectral matching algorithm, that is motivated by the graph/network analysis techniques. It solves a Google's Page rank-like problem on a graph of points (as nodes) $\mathcal{A} \cup \mathcal{B}$ and then extracts a node-to-node correspondence from the visiting frequency of the nodes. This spectral matching algorithm is computationally efficient and reported its good performances in practice [12].

5 Experiments

We tested our algorithm by comparing it with existing algorithms. Our motivation here is how accurately our algorithm finds the underlying one-to-one correspondence, rather than the running speed of algorithms. Thus, we analyzed the accuracy of estimated correspondence P under several perturbations, including measurement noise and isometric transformations (rotation and translation). For the notational simplicity, we identify point-set $\mathcal{A}$ and its matrix representation A in this section.

5.1 Implementation of existing algorithms

We compared our algorithms with the existing ones proposed in [1, 9, 3, 8]. We implemented Itereative closet point algorithm [1] using the k-d tree [2] and the least-square algorithm [6]. For the shape-in-image matching algorithms [9, 3, 8], instead of the pixel-based point-sets in high-dimensional feature space, we just used the 3D coordinates in the Euclidean space, and their dissimilarity matrix was constructed from distances within A, within B, and between A and B.

5.2 Tolerance against noise and invariance under isometric transformations

To test these algorithms, we define the accuracy by the ratio of the number of datasets for which each algorithm found the exactly correct correspondence to the number of all datasets. Specifically, (1) a point-set A of size $n \times 3$ are uniformly randomly generated within $[0, 1]^{n\times3}$. Then, (2) generate the opposite point-set by $B = AR^{\Delta} + \mathbf{1}_n T + E$ for a given matrix A, where $\mathbf{1}_n \in \mathbb{R}^{n\times1}$ is the vector with every element being 1, $R^{\Delta} \in \mathbb{R}^{3\times3}$ is randomly generated rotation matrix, $T \in [0, \tau]^{1\times3}$ is a translation vector with uniformly random values, and $E \in \mathbb{R}^{n\times3}$ is element-wise noise with each element drawn from the normal distribution $N(0, \epsilon)$ with the variance ϵ^2. The rotation matrix R is generated via QR decomposition of uniform random matrix, and R^{Δ} is R to the power of Δ. To evaluate the accuracy, each algorithm was tested for 100 randomly-generated A and B for each combination of Δ, τ, and ϵ. Here we set the number of points in the point-set $\mathcal{A}$ to be $n = 10$.

Figure 1(a) shows the accuracy against the power $\Delta = 0, 0.1, \ldots, 1.2$ to rotation given $\tau = \epsilon = 0$ and (b) the accuracy against translation $\tau = 0, 0.1, \ldots, 1.2$ given $\Delta = \epsilon = 0$. Our algorithm 'Proposed' shown as red markers produced the correct correspondence P exactly 100% for all the rotations and translations. This result shows that the algorithm is robust for a wide range of isometric transformations. This robustness comes from the nature of our algorithm based on the distance matrix, is the invariant under isometric transformations. We found 'Berg2005' [9, 3] also performing as good as ours against rotation and translation, whose result overlaps underneath of the 'Proposed' one in Figure 1 (a) and (b). Iterative closet point algorithm 'Basl1992' [1] showed lower accuracy against larger rotations and translations, and the result shows its limitation under isometric transformations.

Figure 1(c) shows tolerance against measurement noise $\epsilon = 0, 0.01, \ldots, 0.12$ given $\Delta = \tau = 0$ and (d) against noise ϵ combined with rotation $\Delta = 1$ and translation $\tau = 1$. In the conditions of noise effect alone (Figure 1(c)), our proposed algorithm performed worse than 'Berg2005' [9, 3] and 'Basl1992' [1] as the scale of noise increased. This is probably due to the two algorithms 'Berg2005' [9, 3] use the distances between points within each point-set A or B, as well as distances between the point-sets A and B. In contrast, our proposed algorithm is based on the distances within each point-set A or B. However, combining noise and transformations shown in Figure 1(d), our proposed algorithm outperformed others, since our algorithm less affected by isometric transformations. Thus, these analyses with noise and isometric transformations revealed both advantage and disadvantage of our algorithm.

5.3 Tolerance against missing values: partial subset correspondence

Next, we tested our algorithm against missing values. Specifically, we generated two point-sets A and B in the same procedure above. Then, after generated A and B, 7 rows (7 points) of A were dropped as missing values. By this, all the corresponding points of A still exist in B but some points of B has no corresponding points in A, i.e., the problem of injective partial correspondence. In this analysis, the accuracy is defined by the ratio of the number of datasets for which each algorithm could find the exactly correct correspondence between non-missing points between A and B to the number of all datasets. Here we supposed that the all 3 remaining points of A are on the same rigid body of the subject, but the other 7 omitted points of A can be on arbitrary rigid bodies other than this one. In other words, this experiment was intended for the cases finding a partial (subset) correspondence between a single rigid body A and multiple rigid bodies B, or finding a corresponding subset of A in B.

Figure 2(a)–(d) show the results of partial correspondence under the same conditions as Figure 1(a)–(d). In Figure 2(a) rotation only and Figure 2(b) translation only, our proposed algorithm achieved 100% of accuracy. But, for these conditions, 'Berg2005' [9, 3] greatly decreased its accuracy of correct response, in contrast to the full correspondence case (see Figure 1(a) and (b)). In (c) noise

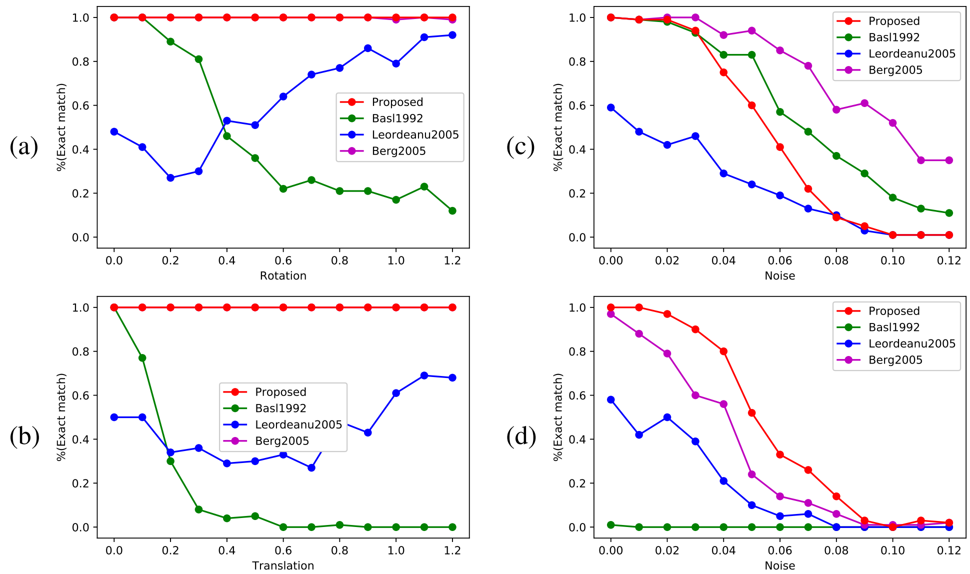

Figure 1: Rates of producing correct correspondence in several conditions. The number of points in $\mathcal{A}$ and $\mathcal{B}$ is both 10, and no missing value. Types of perturbations differ: (a) rotation only, (b) translation only, (c) noise only, (d) combination of noise, rotation and translation.

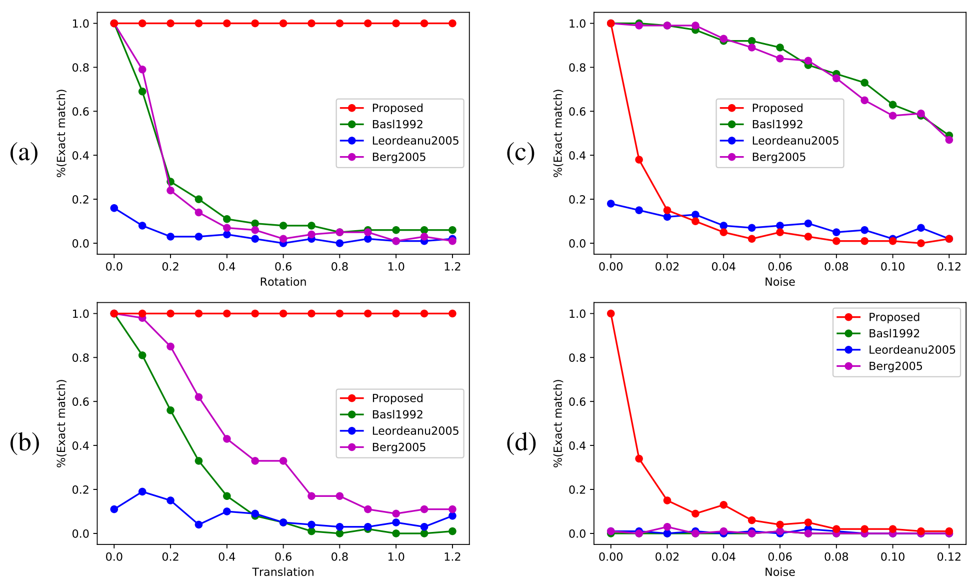

Figure 2: Rates of producing correct "partial" correspondence in several conditions. The number of points in $\mathcal{A}$ is 3 and in $\mathcal{B}$ is 10. There are 7 missing values. Types of perturbations differ: (a) rotation only, (b) translation only, (c) noise only, (d) combination of noise, rotation and translation.

only, our proposed algorithm hardly produced correct partial correspondence under largely noisy conditions, but 'Berg2005' [9, 3] and 'Basl1992' [1] kept higher accuracy overall. It is probably due to the use of distances between point-sets. However, again in (d) combination of noise and transformations, our proposed algorithm relatively worked better than other existing algorithms. Again, this analysis on the partial correspondence shows the advantage of using the invariant under isometric transformations.

6 Discussion

In applications to realistic data recorded by optical motion capture systems, the relationship between the smallest distance between points and the scale of noise matters. The precision of a recent optical motion capture system [13] achieves up to 0.5 mm. When using a standard marker placement, the smallest distance between two markers is about 100 mm. In the experiments in the previous section, the smallest distance between two points is in average about 0.165. Thus, the corresponding precision (scale of noise) in the previous experiments is about 0.000825 ($= 0.5 \times 0.165/100$). At this scale of noise $\epsilon = 0.000825$ with random Euclidean transformations ($\Delta = \tau = 1$), the accuracy or the rate of exactly correct response over 97% in average was achieved by our proposed algorithm for finding partial (subset) correspondence (as Figure 2(d)). However, at the same scale of noise, the other algorithms achieved less than 2% of correct response. Thus, when to find full and partial correspondence between two point-sets at two largely-apart time frames, the proposed nested Hungarian algirithm alone is effectively applicable to a recorded data with optical motion capture systems.

Our algorithm currently has limitation in finding a correspondence between two multiple rigid bodies. Extension of our algorithm for two multiple rigid bodies is our on-going future work.

References

[1] P. J. Basl and N. D. McKay. A method for registration of 3-d shapes. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 14, pages 239–256. 1992.

[2] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.

[3] A. C. Berg, T. L. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondences. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 26–33. 2005.

[4] R. E. Burkard, M. Dell'Amico, and S. Martello. *Assignment Problems*. Society for Industrial and Applied Mathematics, 2009.

[5] Q.-X. Huang, B. Adams, M. Wicke, and L. J. Guibas. Non-rigid registration under isometric deformations. *Computer Graphics Forum*, 27(5):1449–1457, 2008.

[6] W. Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica*, A32:922–923, 1976.

[7] H. W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics*, 2(1–2):83–97, 1955.

[8] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2, pages 1482–1489. 2005.

[9] J. Maciel and J. P. Costeira. A global solution to sparse correspondence problems. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 25, pages 187–199. 2003.

[10] D. Marr. *Vision*. Cambridge: MIT Press, 1982.

[11] J. Munkres. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1):32–38, 1956.

[12] O. van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or. A survey on shape correspondence. *Computer Graphics Forum*, 30(6):1681–1707, 2011.

[13] Vicon Motion Systems. Plug-in gait reference guide. In *Vicon Nexus 2.5 Documentation*, pages 1–95. 2016.