

HiPPI: Higher-Order Projected Power Iterations for Scalable Multi-Matching

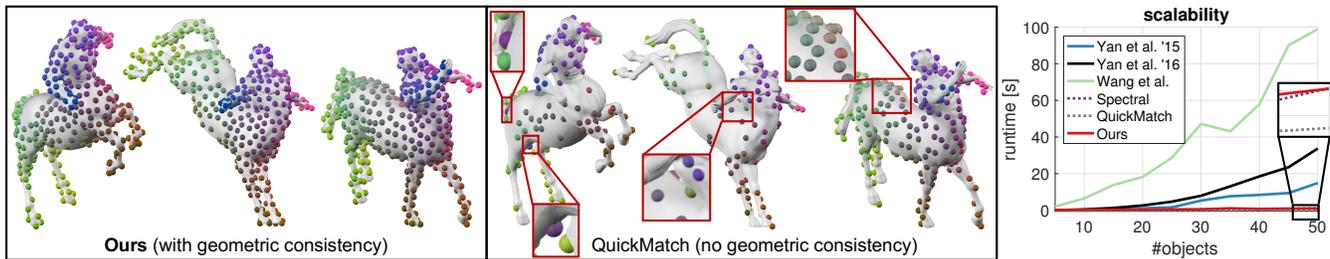
Florian Bernard^{1,2}Johan Thunberg³Paul Swoboda^{1,2}Christian Theobalt^{1,2}¹MPI Informatics²Saarland Informatics Campus³Halmstad University

Figure 1. We propose a novel multi-matching method that is both *scalable* and able to account for *geometric consistency*. Left: the *centaurs* (in different poses) are consistently matched by our approach, as indicated by the coloured dots. Centre: methods that ignore geometric relations (QUICKMATCH [48], SPECTRAL [38]) lead to wrong matchings, as evidenced by mismatching colours in the magnifications. Right: existing methods that account for geometric consistency (Yan et al. '15 [56], Yan et al. '16 [53], Wang et al. [51]) do not scale to large problems (shown are runtimes when matching 5–50 objects, each having 20 points, so that a total of 100–1000 points are matched).

Abstract

The matching of multiple objects (e.g. shapes or images) is a fundamental problem in vision and graphics. In order to robustly handle ambiguities, noise and repetitive patterns in challenging real-world settings, it is essential to take geometric consistency between points into account. Computationally, the multi-matching problem is difficult. It can be phrased as simultaneously solving multiple (NP-hard) quadratic assignment problems (QAPs) that are coupled via cycle-consistency constraints. The main limitations of existing multi-matching methods are that they either ignore geometric consistency and thus have limited robustness, or they are restricted to small-scale problems due to their (relatively) high computational cost. We address these shortcomings by introducing a Higher-order Projected Power Iteration method, which is (i) efficient and scales to tens of thousands of points, (ii) straightforward to implement, (iii) able to incorporate geometric consistency, (iv) guarantees cycle-consistent multi-matchings, and (iv) comes with theoretical convergence guarantees. Experimentally we show that our approach is superior to existing methods.

1. Introduction

Establishing correspondences is a fundamental problem in vision and graphics that is relevant in a wide range of ap-

plications, including reconstruction, tracking, recognition, or matching. The goal of correspondence problems is to identify points in objects (e.g. images, meshes, or graphs) that are semantically similar. While matching points independently of their neighbourhood context is computationally tractable (e.g. via the *linear assignment problem* (LAP) [12]), such approaches are limited to simple cases without ambiguities or repetitive patterns. In order to resolve ambiguities and avoid mismatches in challenging real-world scenarios, it is crucial to additionally incorporate the geometric context of the points, so that spatial distances between pairs of points are (approximately) preserved by the matching. To this end, higher-order information is commonly integrated into the matching problem formulation, e.g. via the NP-hard *quadratic assignment problem* (QAP) [28].

Multi-matching, i.e. finding matchings between *more than two objects* (e.g. an image sequence, a multi-view scene, or a shape collection) plays an important role in various applications, such as video-based tracking, multi-view reconstruction (e.g. for AR/VR content generation) or shape modelling (e.g. for statistical shape models in biomedicine [23]). Computationally, finding valid matchings between more than two objects simultaneously is more difficult compared to matching only a pair of objects. This is because one additionally needs to account for *cycle-consistency*, which means that compositions of matchings over cycles must be the identity matching—even when ignoring higher-order

terms, an analogous multi-matching variant of the *linear assignment problem* that accounts for cycle-consistency results in a (non-convex) quadratic optimisation problem over binary variables, which structurally resembles the NP-hard quadratic assignment problem. When additionally considering higher-order terms in order to account for geometric relations between the points, multi-matching problems become even more difficult. For example, a multi-matching version of the quadratic assignment problem either results in a fourth-order polynomial objective function, or in a quadratic objective with additional (non-convex) quadratic *cycle-consistency constraints*, both of which are to be optimised over binary variables.

Practical approaches for solving multi-matching problems can be put into two categories: (i) methods that jointly optimise for multi-matchings between all objects (e.g. [24, 53, 44, 48, 6]) and (ii) approaches that first establish matchings between points in each pair of objects independently, and then improve those matchings via a post-processing procedure referred to as *permutation synchronisation* [38, 14, 61, 43, 33, 7]. Approaches that jointly optimise for multi-matchings either ignore geometric relations between the points [48], or are prohibitively expensive and thus only applicable to small problems (cf. Fig. 1). In contrast, while synchronisation-based approaches are generally more scalable (e.g. synchronisation problems with a total number of points in the order of 10k-100k can be solved), they completely ignore geometric relations (higher-order terms) during the synchronisation, and thus achieve limited robustness in ambiguous settings (cf. Fig. 1).

The aim of this work is to provide a scalable solution for multi-matching that addresses the mentioned short-comings of previous approaches. Our main contributions are:

- We propose a method that *jointly optimises for multi-matchings* which is *efficient* and thus applicable to *large-scale* multi-matching problems.
- Our method is guaranteed to produce *cycle-consistent multi-matchings*, while at the same time considering *geometric consistency* between the points.
- Our *Higher-order Projected Power Iteration (HiPPI)* method has *theoretical convergence guarantees* and can be implemented in few lines of code.
- We empirically demonstrate that our method achieves beyond state-of-the-art results on various challenging problems, including large-scale multi-image matching and multi-shape matching.

2. Background & Related Work

In this section we review the most relevant works in the literature, while at the same time providing a summary of the necessary background.

Pairwise matching: The linear assignment problem

(LAP) [12] can be phrased as

$$\min_{X \in \mathbb{P}} \langle A, X \rangle, \quad (1)$$

where A is a given matrix that encodes (linear) matching costs between two given objects, and $X \in \mathbb{P}$ is a permutation matrix that encodes the matching between these objects. The LAP can be solved in polynomial time, e.g. via the Kuhn-Munkres/Hungarian method [34] or the (empirically) more efficient Auction algorithm [9, 8]. The quadratic assignment problem (QAP) [28], which reads

$$\min_{X \in \mathbb{P}} \text{vec}(X)^T W \text{vec}(X), \quad (2)$$

additionally incorporates pairwise matching costs between two objects that are encoded by the matrix W . The QAP is a (strict) generalisation of the LAP and in general it is NP-hard [39]. The QAP is a popular formalism for *graph matching* problems, where the first-order terms (on the diagonal of W) account for node matching costs, and the second-order terms (on the off-diagonal of W) account for edge matching costs. Existing methods that tackle the QAP/graph matching include spectral relaxations [30, 17], linear relaxations [47, 46], convex relaxations [59, 41, 37, 21, 1, 26, 20, 6], path-following methods [58, 60, 25], kernel density estimation [50], branch-and-bound methods [5] and many more, as described in the survey papers [39, 31]. Also, tensor-based approaches for higher-order graph matching have been considered [19, 36].

While the requirement $X \in \mathbb{P}$ implies *bijective* matchings, in the case of matching only two objects, the formulations (1) and (2) are general in the sense that they also apply to *partial* matchings, which can be achieved by incorporating *dummy points* with suitable costs. However, due to ambiguities with multi-matchings of dummy points, this is not easily possible when considering more than two objects.

Multi-matching: In contrast to the work [52], where cycle-consistency has been modelled as soft-constraint within a Bayesian framework for multi-graph matching, in [55, 54] the authors have addressed multi-graph matching in terms of simultaneously solving pairwise graph matching under (hard) cycle-consistency constraints. In [56], the authors have generalised *factorised graph matching* [60] from matching a pair of graphs to multi-graph matching. Another approach that tackles multi-graph matching is based on a low-rank and sparse matrix decomposition [57]. In [53], a composition-based approach with a cycle-consistency regulariser is employed. In [26], the authors propose a semidefinite programming (SDP) relaxation for multi-graph matching by (i) relaxing cycle-consistency via a semidefinite constraint, and (ii) lifting the $n \times n$ permutation matrices to $n^2 \times n^2$ -dimensional matrices. In order to reduce computational costs due to the lifting of the permutation matrices, the authors in [6] propose a lifting-free SDP relaxation

for multi-graph matching. More recently, a convex relaxation of multi-graph matching based on message passing has been proposed [45], where cycle-consistency as well as higher-order terms are considered. In [40], the authors propose a random walk technique for multi-layered multi-graph matching. While there is a wide range of algorithmic approaches for multi-graph matching, the aforementioned approaches have in common that they are computationally expensive and are only applicable to small-scale problems, where the total number of points does not significantly exceed a thousand (e.g. 20 graphs with 50 nodes each). In contrast, our method scales much better and handles multi-matching problems with more than 20k points.

In [16], the authors use a two-stage approach with a sparsity-inducing ℓ_1 -formulation for multi-shape matching. While the effect of this approach is that only few multi-matchings are found, our approach obtains significantly more multi-matchings, as we will demonstrate later.

Rather than modelling higher-order relations between points, the recent approach [51] accounts for geometric consistency in 2D multi-image matching problems by imposing a low rank of the (stacked) 2D image coordinates of the feature points. On the one hand, this is based on the (over-simplified) assumption that the 2D images depict a 3D scene under *orthographic* projections, and on the other hand such an extrinsic approach is not directly applicable to distances on non-Euclidean manifolds (e.g. multi-shape matching with geodesic distances). In contrast, our approach is intrinsic due to the use of pairwise adjacency matrices, and thus can handle *general pairwise* information independent of the structure of the ambient space.

Synchronisation methods: Given pairwise matchings between pairs of objects in a collection, synchronisation methods have the purpose to improve these matchings. The motivation of *permutation synchronisation* methods is to achieve *cycle-consistency* in the set of pairwise matchings:

Definition 1. (*Cycle-consistency*)¹

Let $\mathcal{X} = \{X_{ij} \in \mathbb{P}_{m_i m_j}\}_{i,j=1}^k$ be the set of pairwise matchings in a collection of k objects, where each X_{ij} is an element of the set of partial permutation matrices

$$\mathbb{P}_{pq} = \{X \in \{0, 1\}^{p \times q} : X\mathbf{1}_q \leq \mathbf{1}_p, \mathbf{1}_p^T X \leq \mathbf{1}_q^T\}. \quad (3)$$

The set \mathcal{X} is said to be cycle-consistent if for all $i, j, \ell \in [k] := \{1, \dots, k\}$ it holds that:

- (i) $X_{ii} = \mathbf{I}_{m_i}$ (identity matching),
- (ii) $X_{ij} = X_{ji}^T$ (symmetry), and
- (iii) $X_{ij}X_{j\ell} \leq X_{i\ell}$ (transitivity).

For the case of *full* permutation matrices, i.e. in (3) the inequalities become equalities and $p=q$, Pachauri et al. [38]

¹Note that matrix and vector inequalities are understood in an elementwise sense.

have proposed a simple yet effective method to achieve cycle-consistency based on a spectral decomposition of the matrix of pairwise matchings. The authors of [43] provide an analysis of such spectral synchronisations. Earlier works have also considered an iterative refinement strategy to improve pairwise matchings [35].

While the aforementioned synchronisation methods considered the case of full permutations, some authors have also addressed the synchronisation of partial matchings, e.g. based on semidefinite programming [14], alternating direction method of multipliers [61], spectral decomposition followed by k-means clustering [2], or non-negative matrix factorisation [7]. A spectral approach has also been presented in [33], which, however, merely improves given initial pairwise matchings without guaranteeing cycle-consistency.

Rather than explicitly modelling the cubic number of cycle-consistency constraints (cf. Def. 1), most permutation synchronisation methods leverage the fact that cycle-consistency can be characterised by using the notion of *universe points*, as e.g. in [48, 7]:

Lemma 2. (*Cycle-consistency, universe points*)

The set \mathcal{X} of pairwise matchings is cycle-consistent, if there exists a collection $\{X_\ell \in \mathbb{P}_{m_\ell d} : X_\ell \mathbf{1}_d = \mathbf{1}_{m_\ell}\}_{\ell=1}^k$, such that for each $X_{ij} \in \mathcal{X}$ it holds that $X_{ij} = X_i X_j^T$.

Here, the $(m_\ell \times d)$ -dimensional *object-to-universe* matching matrix X_ℓ assigns to each of the m_ℓ points of object ℓ exactly one of d universe points—as such, all points among the k objects that are assigned to a given (unique) universe point are said to be in correspondence.

For notational brevity, it is convenient to consider a matrix formulation of Lemma 2. With m_i being the number of points in the i -th object and $m = \sum_{i=1}^k m_i$, let \mathbf{X} be the $(m \times m)$ -dimensional pairwise matching matrix

$$\mathbf{X} = [X_{ij}]_{i,j=1}^k \in [\mathbb{P}_{m_i m_j}]_{i,j=1}^k \quad (4)$$

and let

$$\mathcal{U} = \{U \in [\mathbb{P}_{m_i d}]_i : U\mathbf{1}_d = \mathbf{1}_m\} \subset \{0, 1\}^{m \times d}. \quad (5)$$

Lemma 2 translates into the requirement that there must be a $U \in \mathcal{U}$, such that

$$\mathbf{X} = UU^T. \quad (6)$$

With this matrix notation it becomes also apparent that one can achieve synchronisation by matrix *factorisation*, such as pursued by the aforementioned spectral approaches [38, 43, 2, 33]. While recently a lot of progress has been made for permutation synchronisation, one of the open problems is how to efficiently integrate higher-order information to model geometric relations between points. We

achieve this goal with our proposed method and demonstrate a significant improvement of the matching accuracy due to the additional use of geometric information.

Power method: The power method is one of the classical numerical linear algebra routines for computing the eigenvector corresponding to the largest (absolute) eigenvalue [22]. Moreover, it is well-known that the eigenvector corresponding to the largest (absolute) eigenvalue maximises the Rayleigh quotient $\frac{x^T Ax}{x^T x}$, which, up to scale, is equivalent to maximising the (not necessarily convex) quadratic objective $x^T Ax$ over the unit sphere. In addition to computing a single eigenvector, straightforward extensions of the power method are the *Orthogonal Iteration* and *QR Iteration* methods [22], which simultaneously compute multiple (orthogonal) eigenvectors and can be used to maximise quadratic objectives over the Stiefel manifold.

Power method generalisations: Moreover, higher-order generalisations of the power method have been proposed, e.g. for rank-1 tensor approximation [18], or for multi-graph matching [44]. However, the latter approach has a runtime complexity that is *exponential* in the number of graphs and thus prevents scalability (e.g. matching 12 graphs, each with 10 nodes, takes about 10 minutes). In [13], the authors propose a *Projected Power Method* for the optimisation of quadratic functions over sets other than the Stiefel manifold, such as permutation matrices. In a permutation synchronisations setting, their method obtains results that are comparable to semidefinite relaxations methods [14, 24] at a reduced runtime. However, due to the restriction to quadratic objective functions, their approach cannot handle geometric relations between points, as they would become polynomials of degree four, as will be explained in Sec. 3. Our method goes beyond the existing approaches as we propose a convergent projected power iteration method for maximising a higher-order objective over the set \mathcal{U} . With that, we can incorporate geometric information between neighbouring points using a fourth-order polynomial, while always maintaining cycle-consistency.

3. Method

The overall idea of our approach is to phrase the multi-matching problem as simultaneously solving k^2 pairwise matching problems with quadratic (second-order) matching scores, where the variables are weighted based on linear (first-order) matching scores. Instead of directly optimising over *pairwise matchings*, we parametrise the pairwise matchings in terms of their *object-to-universe* matchings, cf. Lemma 2 and Eq. (6). Although this has the disadvantage that the quadratic term becomes quartic (a fourth-order polynomial), it has the strong advantages that (i) cycle-consistency is guaranteed to be always maintained, (ii) one only optimises for $m \times d$, rather than $m \times m$ variables in the pairwise case, where commonly $d \ll m$.

3.1. Multi-Matching Formulation

In the following we present our multi-matching formulation. Let $A_i \in \mathcal{S}_{m_i}^+$ denote the $(m_i \times m_i)$ -dimensional symmetric and positive semidefinite *adjacency matrix* of object i (e.g. a matrix that encodes the Gaussian of pairwise Euclidean/geodesic distances between pairs of points). For a given $X_i \in \mathbb{P}_{m_i d}$, the $d \times d$ matrix $X_i^T A_i X_i$ is a row/column reordering of the matrix A_i according to the universe points. As such, we can use the Frobenius inner product $\langle X_i^T A_i X_i, X_j^T A_j X_j \rangle$ for quantifying how well two adjacency matrices A_i and A_j agree after they have been reordered based on X_i and X_j . This term can be understood as the object-to-universe formulation of second-order matching terms when using pairwise matching matrices (analogous to the QAP in Koopmans-Beckmann form [27]). As such, multi-matching with geometric consistency can be phrased as

$$\begin{aligned} \max_{X_1, \dots, X_k} & \sum_{i,j=1}^k \langle X_i^T A_i X_i, X_j^T A_j X_j \rangle \\ \text{s.t} & X_i \in \mathbb{P}_{m_i d} \forall i \in [k]. \end{aligned} \quad (7)$$

Equivalently, for $U = [X_1^T, \dots, X_k^T]^T \in \mathbb{R}^{m \times d}$ and $A \in \mathcal{S}_m^+$ being the (symmetric and positive semidefinite) block-diagonal *multi-adjacency* matrix defined as $A = \text{diag}(A_1, \dots, A_k) \in \mathbb{R}^{m \times m}$, Problem (7) can be written in compact matrix form as

$$\max_{U \in \mathcal{U}} \text{tr}(U^T A U U^T A U). \quad (8)$$

Let $W_{ij} \in \mathbb{R}_+^{m_i \times m_j}$ encode the (non-negative) similarity scores between the points of object i and j (in analogy to linear terms when using pairwise matching matrices), and let $W = [W_{ij}]_{ij} \in \mathbb{R}^{m \times m}$ be a (symmetric) matrix that encodes all similarity scores. With that, we can define the reweighting of the object-to-universe matchings U as WU , where both U and WU have the same dimensionality. The purpose of this reweighting is to amplify matchings in U that have high similarity scores, cf. Fig. 2. By replacing U in Problem (8) with its reweighted matrix WU , we arrive at our final multi-matching formulation

$$\max_{U \in \mathcal{U}} \text{tr}(U^T \overline{W} U U^T \overline{W} U) := f(U), \quad (9)$$

for $\overline{W} := W^T A W$. While (9) is based on the U -matrix and hence intrinsically guarantees cycle-consistent multi-matchings, the objective function is a fourth-order polynomial that is to be maximised over the (binary) set \mathcal{U} .

3.2. Algorithm

In order to solve Problem (9) we propose a Higher-order Projected Power Iteration (HiPPI) method, as outlined in

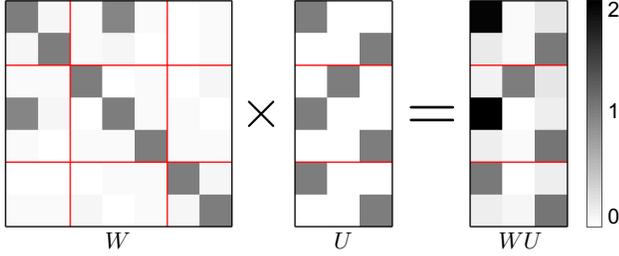


Figure 2. Effect of reweighting U with W . The similarity matrix W indicates that the first point in the first object is similar the second point in the second object (the grey values in W at position (1, 4), and (4, 1)). Hence, those elements in the first column of U that match these points to each other are amplified in WU .

Alg. 1. The main idea is to perform a higher-order power iteration step, followed by a projection onto the set \mathcal{U} . The approach is simple, and merely comprises of matrix multiplications and the Euclidean projection onto \mathcal{U} .

Input: similarities W , multi-adjacency matrix $A \in S_m^+$

Output: cycle-consistent multi-matching $U_t \in \mathcal{U}$

Initialise: $t \leftarrow 0, U_0 \in \mathcal{U}, \bar{W} \leftarrow W^T A W$

1 **repeat**

2 $V_t \leftarrow \bar{W} U_t U_t^T \bar{W} U_t$

3 $U_{t+1} \leftarrow \text{proj}_{\mathcal{U}}(V_t)$

4 $t \leftarrow t+1$

5 **until** $|f(U_t) - f(U_{t-1})| = 0$

Algorithm 1: Higher-order Projected Power Iteration (HiPPI) algorithm.

3.3. Theoretical Analysis

In this section we analyse the properties of our algorithm. The respective proofs can be found in the supplementary material.

Monotonicity: The HiPPI algorithm is guaranteed to monotonically increase the objective $f(U_t)$:

Proposition 3. *In Alg. 1 it holds that $f(U_{t+1}) \geq f(U_t)$.*

In the proof of Proposition 3 we use the following result:

Lemma 4. *Let W be a symmetric and positive semidefinite matrix. For $\text{tr}(V^T W V V^T W V) \leq \text{tr}(U^T W V V^T W U)$ we have $\text{tr}(U^T W V V^T W U) \leq \text{tr}(U^T W U U^T W U)$.*

Convergence: Proposition 3 immediately leads to the convergence of our algorithm, which is stated in the following corollary:

Corollary 5. *The sequence $(f(U_t))_{t=0,1,\dots}$ produced by Alg. 1 converges after a finite number of iterations.*

Remark 6. *As the order of the universe points can be arbitrary, there is a whole family of equivalent solutions: for*

any $U' = UP$ for P being a $d \times d$ permutation, we have that $f(U) = f(U')$, since $U'(U')^T = U P P^T U^T = U U^T$.

Complexity analysis: The update rule in Alg. 1 can be written as $U_{t+1} = \text{proj}_{\mathcal{U}}(\bar{W} U_t U_t^T \bar{W} U_t)$. The matrix multiplications for computing $\bar{W} U_t U_t^T \bar{W} U_t$ have time complexity $\mathcal{O}(m^2 d)$, where the universe size d is generally much smaller compared to the total number of all points m across all the k objects. As can be seen in the proof of Proposition 3, the projection onto \mathcal{U} is a linear programming problem. Finding its optimiser amounts to solving k individual (partial) linear assignment problems, each having sub-cubic empirical average time complexity when using the Auction algorithm [9, 42] (we use the implementation by [8]). Hence, the overall (average) per-iteration complexity is $\mathcal{O}(m^2 d + k d^2 \log(d))$. The memory complexity is $\mathcal{O}(m^2)$ due to the matrix $W \in \mathbb{R}^{m \times m}$, which can be improved by considering *sparse* similarity scores.

4. Experimental Results

In this section we extensively compare our method to other approaches on three datasets. To be more specific, we consider two multi-image matching datasets, Willow [15] and HiPPI, as well as the multi-shape matching dataset Tosca [11]. The datasets are summarised in Table 1.

Dataset	Type	Bij.	#	k	m
HiPPI	images	no	31	[7,100]	[2257,20703]
Willow	images	yes	5	[40,108]	[400,1080]
Tosca	shapes	no	7	[6,20]	[3000,10000]

Table 1. Overview of properties of the datasets. ‘‘Bij.’’ indicates whether the matchings are bijective, ‘‘#’’ is the number of instances per dataset, k is the number of objects, and m is the total number of points summed over all objects.

Similarity scores: The similarity scores between image/shape i and j is encoded in the matrix $W_{ij} \in \mathbb{R}_+^{m_i \times m_j}$, which is defined using *feature matrices* $F_i \in \mathbb{R}^{m_i \times f}$ and $F_j \in \mathbb{R}^{m_j \times f}$ of the respective image/shape, where f is the feature dimensionality. As in [48], the similarity scores are based on a weighted Gaussian kernel, i.e.

$$(W_{ij})_{pq} = \omega_{pq} \exp\left(-\frac{\|(F_i)_p - (F_j)_q\|^2}{2\sigma^2}\right), \quad (10)$$

where ω_{pq} is a weight that depends on the distance between the features and the closest descriptor from the same image. For details we refer the reader to [48]. The particular choice of features for each dataset are described below.

Adjacency matrices: The adjacency matrix $A_i \in S_{m_i}^+$ of image/shape i is based on Euclidean distances between pairs of 2D image point locations in the case of multi-image matching (or geodesic distances between pairs of points on

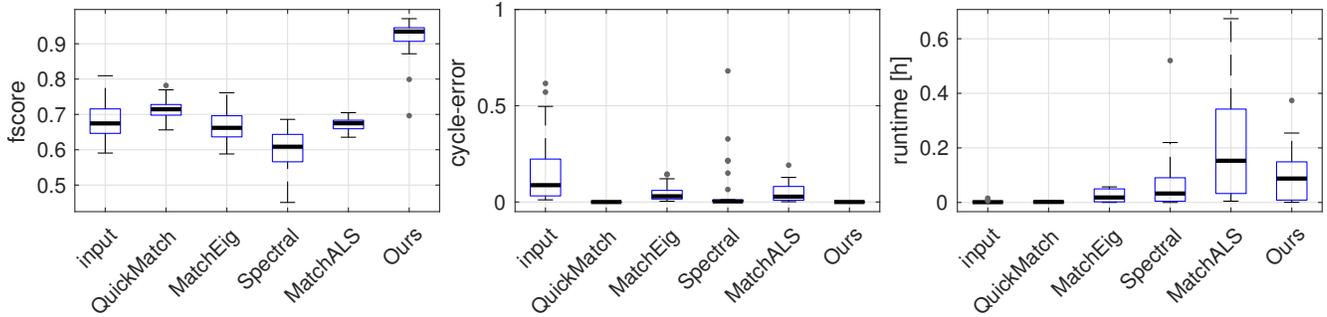


Figure 3. Multi-image matching on the HiPPI dataset. Our method is the only one that considers geometric relations between points, while at the same time being scalable to such large datasets (note that the reported runtimes include the initialisation). We clearly outperform other matching methods as well as synchronisation methods in terms of the fscore (higher is better), while also guaranteeing cycle-consistency.

the 3D shape surface for multi-shape matching). By denoting the distance between the points with indices $p \in [m_i]$ and $q \in [m_i]$ as d_{pq} , the elements of the adjacency matrix are based on a Gaussian kernel, so that

$$(A_i)_{pq} = \exp\left(-\frac{d_{pq}^2}{2\mu\sigma_A^2}\right). \quad (11)$$

We set $\sigma_A = \text{median}(d_{\min})$, where $(d_{\min})_p = \min_{q \neq p} d_{pq}$ for $p \in [m_i]$, and μ is a scaling factor.

Quantitative scores: For the quantitative evaluation we consider the $\text{fscore} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$, and the cycle-error, which is given by the fraction of the total number of cycle-consistency violations in all three-cycles, divided by the total number of matchings in all three-cycles.

4.1. HiPPI Dataset

In this experiment we compare various multi-image matching methods.

Dataset: The HiPPI dataset comprises 31 multi-image matching problems. For each problem instance, a (short) video sequence has been recorded (with resolution 1920×1080 , frame-rate 30 FPS, and duration $> 5s$). In each video, feature points and feature descriptors were extracted using SURF [4] with three octaves. To obtain ground truth matchings, these feature points were tracked across the sequence based on their geometric distance and feature descriptor similarity. To ensure *reliable* ground truth matchings, we have conducted the following three steps: (i) obvious wrong matchings between consecutive images have been automatically pruned, (ii) we have manually removed those features that were incorrectly tracked from the first to the last frame (by inspecting the first and last frame), and (iii) in order to prevent feature sliding in-between the first and last frame, we manually inspected the flow of each remaining feature point and removed wrongly tracked points. Note that steps (ii) and (iii) have been performed by two different persons, which took in total about 24 hours. A multi-matching problem was then created by extracting

evenly spaced frames from a sequence, where in each frame we added a significant amount of outlier points (randomly selected from the previously pruned points, where the number of points is chosen such that in each frame 50% of the points are outliers), and we simulate occlusions (a rectangle of size 0.2×0.2 of the image dimensions) in order to get difficult partial multi-matching problems.

Multi-image matching: We compare our method with QUICKMATCH [48], MATCHEIG [33], SPECTRAL [38] (implemented by the authors of [61] for *partial* permutation synchronisation), and MATCHALS [61]. Other methods that incorporate geometric information [56, 53, 51] do not scale to such large problem instances, see Fig. 1, and thus we cannot compare to them. The universe size d is set to twice the average of the number of points per image [33]. We used QUICKMATCH ($\rho_{\text{den}}=0.5$) for initialising U_0 in our method, and we set $\mu=1$. The results are shown in Fig. 3, where it can be seen that our method achieves a superior matching quality. In contrast to other methods (except QUICKMATCH), our method guarantees cycle-consistency.

4.2. Willow Dataset

The evaluation on the Willow dataset [15] is based on the experimental protocol from [51], where deep features have been used for matching. For this dataset the matchings are bijective, and hence for all methods we set the universe size d to the number of annotated features. Since QUICKMATCH [48] is tailored towards partial matchings, as it implicitly determines the universe size during its internal clustering, we have found that it does not perform so well on this dataset (see Table 2). In Table 2 it can be seen that our method is superior compared to the other approaches. We show some qualitative results in Fig. 4.

4.3. Tosca Dataset

Based on the experimental setup of [16] using the Tosca dataset [11], we compare our method with two other approaches that guarantee cycle-consistency, namely QUICK-

Instance	PAIRWISE	SPECTRAL	MATCHALS	QUICKMATCH	YAN ET AL.	WANG ET AL.	OURS
Car	0.54 (0.4s)	0.65 (0.5s)	0.64 (4.3s)	0.27 (0.2s)	0.57 (22.9s)	0.71 (3.4s)	0.74 (0.9s)
Duck	0.48 (0.6s)	0.66 (0.7s)	0.60 (4.2s)	0.27 (0.2s)	0.54 (21.7s)	0.82 (4.4s)	0.88 (1.0s)
Face	0.94 (3.5s)	0.98 (3.6s)	0.97 (6.9s)	0.91 (0.2s)	0.85 (14.9s)	0.96 (5.0s)	1.00 (3.8s)
Motorbike	0.33 (0.5s)	0.37 (0.6s)	0.35 (3.6s)	0.12 (0.2s)	0.32 (33.3s)	0.65 (8.4s)	0.84 (1.0s)
Winebottle	0.62 (1.3s)	0.82 (1.4s)	0.77 (4.7s)	0.29 (0.2s)	0.71 (20.4s)	0.88 (4.9s)	0.95 (1.6s)

Table 2. Fscores (higher is better) and runtimes for the Willow dataset. All methods are initialised based on pairwise matches with linear costs (except QUICKMATCH [48], which is initialisation-free). OURS ($\mu=10$), YAN ET AL. [53], and WANG ET AL. [51] consider geometric relations between points, amongst which OURS is the fastest (note that the reported runtimes include the initialisation).



Figure 4. A subset of images from four multi-matching problems of the WILLOW dataset (solved using our method). The correspondences are encoded by the colours of the points.

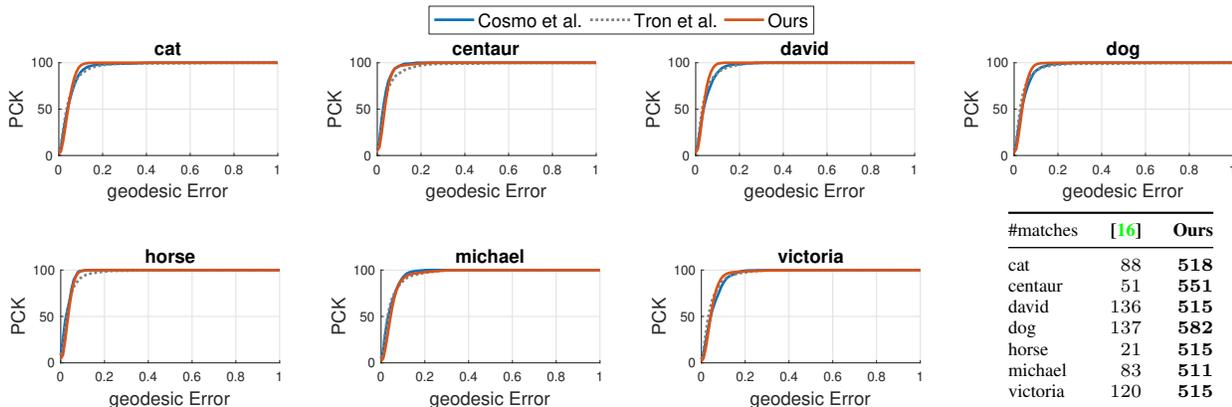


Figure 5. Our method obtains significantly more multi-matches (bottom right, see also Fig. 6) compared to the method of Cosmo et al. [16], while at the same time achieving comparable errors (percentage of correct keypoints, PCK).

MATCH, and the sparse multi-shape matching approach by Cosmo et al. [16]. The feature descriptors on the shape surfaces are based on *wave kernel signatures (WKS)* [3], we use QUICKMATCH ($\rho_{den}=0.2$) as initialisation, and set $\mu=5$.

Multi-shape matching: Quantitative results are shown

in Fig. 5 and qualitative results are shown in Fig. 6. As explained before, QUICKMATCH ignores geometric relations between points, and thus leads to geometric inconsistencies (first row of Fig. 6). While the method of Cosmo et al. [16] is able to incorporate geometric relations between points,

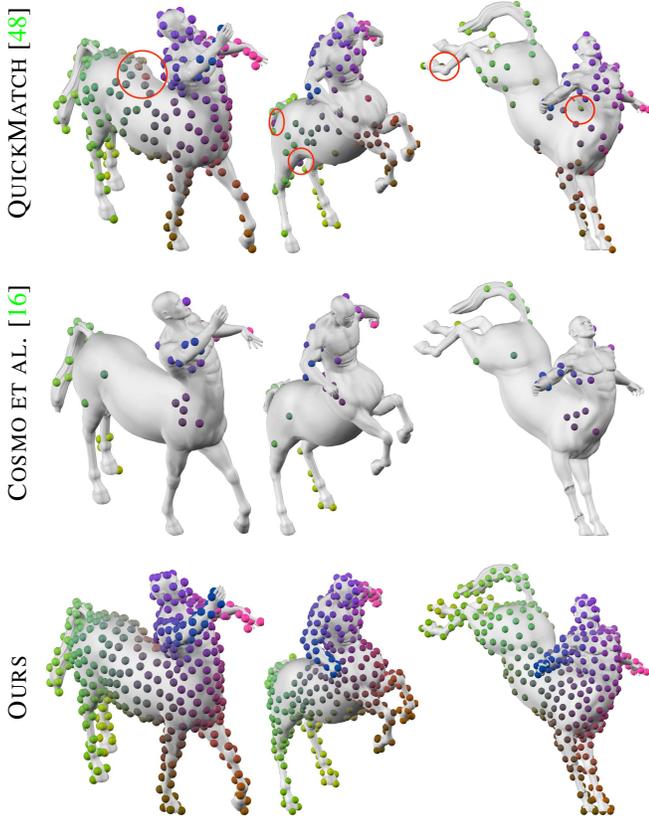


Figure 6. Qualitative results on the Tosca centaur. Dots with same colour indicate matched points. QUICKMATCH (top) does not take geometric consistency into account and thus leads to mismatches (red circles). While Cosmo et al. [16] (centre) obtain only few reliable matchings, leading to *regions without correspondences*, our method (bottom) obtains *significantly more reliable matchings*.

one major limitation of their approach is that only a sparse subset of matchings is found. This may happen even when the shape collection is outlier free [16]. This behaviour can be seen in the second row of Fig. 6, where only few multi-matchings are obtained and hence there are large regions for which no correspondences are found. In contrast, our approach incorporates geometric consistency, produces significantly more multi-matchings (Fig. 6), and results in a percentage of correct keypoints (PCK) that is competitive to the method of Cosmo et al. [16], see Fig. 5.

5. Discussion & Limitations

Positive semidefiniteness of A : Although at first sight it may seem overly restrictive that our monotonicity results of Alg. 1 rely on the multi-adjacency matrix A being positive semidefinite, in practice this is usually not an issue. As explained e.g. in [49], for getting rid of a bias towards far-away points, one usually applies a kernel (e.g. Gaussian or heat kernel) to the Euclidean or geodesic distances between

pairs of points, which directly results in positive semidefinite adjacency matrices [10].

Convergence speed: While we do not provide convergence rates for our algorithm, we experimentally verified that in practice our algorithm only needs few steps, i.e. it took less than 40 iterations in all conducted experiments. We have initially experimented with a Frank-Wolfe (FW) approach for our multi-matching formulation that optimises over the convex hull of \mathcal{U} . Although such an approach would allow to provide convergence rates, we have found that our method was superior compared to FW.

Initialisation: As Problem (9) is non-convex, the result produced by Alg. 1 generally depends on the initialisation U_0 . We have additionally evaluated our method on the HiPPI dataset using random initialisations, where we have obtained an average fscore of 0.90 ± 0.08 (QUICKMATCH on its own resulted in an average fscore of 0.71 ± 0.03 , and our method with QUICKMATCH initialisation achieved 0.92 ± 0.05 , cf. Fig. 3).

General QAPs: Our formulation of the geometric consistency term in Problem (9) corresponds to the Koopmans-Beckmann form of the QAP [27], which is strictly less general compared to Lawler’s form $\text{vec}(X)^T W \text{vec}(X)$ [28]. An interesting direction for future work is to devise an analogous algorithm for solving Lawler’s form.

Partial objects: In our formulation we do not explicitly consider the rejection of outlier points. Hence, our method works best when matching a collection of objects that can properly be matched. We believe that matching collections of partial objects is an interesting direction for future work.

6. Conclusion

We presented a Higher-order Projected Power Iteration approach for multi-matching. Contrary to existing permutation synchronisation methods [38, 14, 61, 43, 2, 33], our method is able to take geometric relations between points into account. Hence, our approach can be seen as a generalisation of permutation synchronisation. Moreover, previous multi-matching methods that consider geometric consistency (e.g. [56, 53]) only allow to solve problems with up to few thousand points. In contrast, we demonstrated that our approach scales to tens of thousands of points. In addition to being able to account for geometric consistency, key properties of our method are convergence guarantees, computational efficiency, simplicity, and guaranteed cycle-consistency. Moreover, we have demonstrated superior performance on three datasets, which highlights the practical relevance of the proposed algorithm.

Acknowledgements: This work was funded by the ERC Consolidator Grant 4DRepLy. We thank Thomas Möllenhoff for helpful input regarding the convergence, and Franziska Müller for providing feedback on the manuscript.

References

- [1] Yonathan Aflalo, Alexander Bronstein, and Ron Kimmel. On convex relaxation of graph isomorphism. *Proceedings of the National Academy of Sciences*, 112(10):2942–2947, 2015. [2](#)
- [2] Federica Arrigoni, Eleonora Maset, and Andrea Fusiello. Synchronization in the Symmetric Inverse Semigroup. In *ICIAP*, 2017. [3](#), [8](#)
- [3] Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *ICCV Workshops*, 2011. [7](#)
- [4] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359, 2008. [6](#)
- [5] Mokhtar S Bazaraa and Alwalid N Elshafei. An exact branch-and-bound procedure for the quadratic-assignment problem. *Naval Research Logistics Quarterly*, 26(1):109–121, 1979. [2](#)
- [6] Florian Bernard, Christian Theobalt, and Michael Moeller. DS*: Tighter Lifting-Free Convex Relaxations for Quadratic Matching Problems. In *CVPR*, 2018. [2](#)
- [7] Florian Bernard, Johan Thunberg, Jorge Goncalves, and Christian Theobalt. Synchronisation of partial multi-matchings via non-negative factorisations. *Pattern Recognition*, 92:146–155, 2019. [2](#), [3](#)
- [8] Florian Bernard, Nikos Vlassis, Peter Gemmar, Andreas Husch, Johan Thunberg, Jorge Goncalves, and Frank Hertel. Fast correspondences for statistical shape models of brain structures. In *Proc. SPIE Medical Imaging*, 2016. [2](#), [5](#)
- [9] Dimitri P Bertsekas. *Network Optimization: Continuous and Discrete Models*. Athena Scientific, 1998. [2](#), [5](#)
- [10] Christopher M Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., 2006. [8](#)
- [11] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes*. Springer Science & Business Media, 2008. [5](#), [6](#)
- [12] Rainer Burkard, Mauro Dell’Amico, and Silvano Martello. Assignment problems. 2009. [1](#), [2](#)
- [13] Yuxin Chen and Emmanuel Candes. The Projected Power Method: An Efficient Algorithm for Joint Alignment from Pairwise Differences. *Communications on Pure and Applied Mathematics*, 71(8), 2018. [4](#)
- [14] Yuxin Chen, Leonidas J Guibas, and Qi-Xing Huang. Near-Optimal Joint Object Matching via Convex Relaxation. In *ICML*, 2014. [2](#), [3](#), [4](#), [8](#)
- [15] Minsu Cho, Karteek Alahari, and Jean Ponce. Learning graphs to match. In *ICCV*, 2013. [5](#), [6](#)
- [16] Luca Cosmo, Emanuele Rodolà, Andrea Albarelli, Facundo Mémoli, and Daniel Cremers. Consistent partial matching of shape collections via sparse modeling. In *Computer Graphics Forum*, volume 36, pages 209–221, 2017. [3](#), [6](#), [7](#), [8](#)
- [17] Timothee Cour, Praveen Srinivasan, and Jianbo Shi. Balanced graph matching. *NIPS*, 2006. [2](#)
- [18] Lieven De Lathauwer, Pierre Comon, Bart De Moor, and Joos Vandewalle. Higher-order power method. *Nonlinear Theory and its Applications, NOLTA95*, 1, 1995. [4](#)
- [19] Olivier Duchenne, Francis Bach, In-So Kweon, and Jean Ponce. A tensor-based algorithm for high-order graph matching. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2383–2395, 2011. [2](#)
- [20] Nadav Dym, Haggai Maron, and Yaron Lipman. DS++ - A flexible, scalable and provably tight relaxation for matching problems. *ACM Transactions on Graphics (TOG)*, 36(6), 2017. [2](#)
- [21] Fajwel Fogel, Rodolphe Jenatton, Francis Bach, and Alexandre d’Aspremont. Convex Relaxations for Permutation Problems. In *NIPS*, 2013. [2](#)
- [22] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 1996. [4](#)
- [23] Tobias Heimann and Hans-Peter Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, 13(4):543–563, 2009. [1](#)
- [24] Qi-Xing Huang and Leonidas Guibas. Consistent shape maps via semidefinite programming. In *Symposium on Geometry Processing*, 2013. [2](#), [4](#)
- [25] Bo Jiang, Jin Tang, Chris Ding, and Bin Luo. Binary Constraint Preserving Graph Matching. In *CVPR*, 2017. [2](#)
- [26] Itay Kezurer, Shahar Z Kovalsky, Ronen Basri, and Yaron Lipman. Tight Relaxation of Quadratic Matching. *Comput. Graph. Forum*, 2015. [2](#)
- [27] Tjalling C Koopmans and Martin Beckmann. Assignment Problems and the Location of Economic Activities. *Econometrica*, 25(1):53, Jan. 1957. [4](#), [8](#)
- [28] Eugene L Lawler. The quadratic assignment problem. *Management science*, 9(4):586–599, 1963. [1](#), [2](#), [8](#)
- [29] Hoai An Le Thi and Tao Pham Dinh. Dc programming and dca: thirty years of developments. *Mathematical Programming*, pages 1–64, 2018. [11](#)
- [30] Marius Leordeanu and Martial Hebert. A Spectral Technique for Correspondence Problems Using Pairwise Constraints. In *ICCV*, 2005. [2](#)
- [31] Eliane Maria Loiola, Nair Maria Maia de Abreu, Paulo Oswaldo Boaventura Netto, Peter Hahn, and Tania Maia Querido. A survey for the quadratic assignment problem. *European Journal of Operational Research*, 176(2):657–690, 2007. [2](#)
- [32] João Maciel and João P Costeira. A global solution to sparse correspondence problems. *TPAMI*, 25(2), 2003. [11](#)
- [33] Eleonora Maset, Federica Arrigoni, and Andrea Fusiello. Practical and Efficient Multi-View Matching. In *ICCV*, 2017. [2](#), [3](#), [6](#), [8](#)
- [34] James Munkres. Algorithms for the Assignment and Transportation Problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1):32–38, Mar. 1957. [2](#)
- [35] Andy Nguyen, Mirela Ben-Chen, Katarzyna Welnicka, Yinyu Ye, and Leonidas J Guibas. An Optimization Approach to Improving Collections of Shape Maps. *Computer Graphics Forum*, 30(5):1481–1491, 2011. [3](#)
- [36] Quynh Nguyen, Antoine Gautier, and Matthias Hein. A flexible tensor block coordinate ascent scheme for hypergraph matching. 2015. [2](#)
- [37] Carl Olsson, Anders P Eriksson, and Fredrik Kahl. Solving Large Scale Binary Quadratic Problems - Spectral Methods vs. Semidefinite Programming. *CVPR*, 2007. [2](#)

- [38] Deepti Pachauri, Risi Kondor, and Vikas Singh. Solving the multi-way matching problem by permutation synchronization. In *NIPS*, 2013. 1, 2, 3, 6, 8
- [39] Panos M Pardalos, Franz Rendl, and Henry Wolkowicz. The Quadratic Assignment Problem - A Survey and Recent Developments. *DIMACS Series in Discrete Mathematics*, 1993. 2
- [40] Han-Mu Park and Kuk-Jin Yoon. Consistent multiple graph matching with multi-layer random walks synchronization. *Pattern Recognition Letters*, 2018. 3
- [41] Christian Schellewald and Christoph Schnörr. Probabilistic subgraph matching based on convex relaxation. In *EMM-CVPR*, 2005. 2
- [42] B L Schwartz. A computational analysis of the Auction algorithm. *European Journal of Operational Research*, 74(1):161–169, 1994. 5
- [43] Yanyao Shen, Qixing Huang, Nati Srebro, and Sujay Sanghavi. Normalized Spectral Map Synchronization. In *NIPS*, 2016. 2, 3, 8
- [44] X Shi, H Ling, W Hu, and J Xing. Tensor power iteration for multi-graph matching. In *CVPR*, 2016. 2, 4
- [45] Paul Swoboda, Dagmar Kainmüller, Ashkan Mokarian, Christian Theobalt, and Florian Bernard. A convex relaxation for multi-graph matching. In *CVPR*, 2019. 3
- [46] Paul Swoboda, Carsten Rother, Hassan Abu Alhaija, Dagmar Kainmüller, and Bogdan Savchynskyy. Study of lagrangean decomposition and dual ascent solvers for graph matching. In *CVPR*, 2017. 2
- [47] Lorenzo Torresani and Vladimir Kolmogorov. A dual decomposition approach to feature correspondence. *TPAMI*, 35(2):259–271, 2013. 2
- [48] Roberto Tron, Xiaowei Zhou, Carles Esteves, and Kostas Daniilidis. Fast Multi-Image Matching via Density-Based Clustering. In *ICCV*, 2017. 1, 2, 3, 5, 6, 7, 8, 12
- [49] Matthias Vestner, Zorah Löhner, Amit Boyarski, Or Litany, Ron Slossberg, Tal Remez, Emanuele Rodola, Alex Bronstein, Michael Bronstein, Ron Kimmel, et al. Efficient deformable shape correspondence via kernel matching. In *3DV*, 2017. 8
- [50] Matthias Vestner, Roei Litman, Emanuele Rodolà, Alexander M Bronstein, and Daniel Cremers. Product Manifold Filter - Non-Rigid Shape Correspondence via Kernel Density Estimation in the Product Space. *CVPR*, 2017. 2
- [51] Qianqian Wang, Xiaowei Zhou, and Kostas Daniilidis. Multi-Image Semantic Matching by Mining Consistent Features. In *CVPR*, 2018. 1, 3, 6, 7
- [52] Mark L Williams, Richard C Wilson, and Edwin R Hancock. Multiple graph matching with Bayesian inference. *Pattern Recognition Letters*, 1997. 2
- [53] Junchi Yan, Minsu Cho, Hongyuan Zha, Xiaokang Yang, and Stephen M. Chu. Multi-graph matching via affinity optimization with graduated consistency regularization. *TPAMI*, 38(6):1228–1242, 2016. 1, 2, 6, 7, 8
- [54] Junchi Yan, Yin Li, Wei Liu, Hongyuan Zha, Xiaokang Yang, and Stephen Mingyu Chu. Graduated Consistency-Regularized Optimization for Multi-graph Matching. *ECCV*, 2014. 2
- [55] Junchi Yan, Yu Tian, Hongyuan Zha, Xiaokang Yang, Ya Zhang, and Stephen M Chu. Joint Optimization for Consistent Multiple Graph Matching. *ICCV*, 2013. 2
- [56] Junchi Yan, Jun Wang, Hongyuan Zha, Xiaokang Yang, and Stephen Chu. Consistency-driven alternating optimization for multigraph matching: A unified approach. *IEEE Transactions on Image Processing*, 2015. 1, 2, 6, 8
- [57] Junchi Yan, Hongteng Xu, Hongyuan Zha, Xiaokang Yang, Huanxi Liu, and Stephen Chu. A matrix decomposition perspective to multiple graph matching. In *ICCV*, 2015. 2
- [58] Mikhail Zaslavskiy, Francis Bach, and Jean-Philippe Vert. A Path Following Algorithm for the Graph Matching Problem. *TPAMI*, 31(12):2227–2242, 2009. 2
- [59] Qing Zhao, Stefan E Karisch, Franz Rendl, and Henry Wolkowicz. Semidefinite programming relaxations for the quadratic assignment problem. *Journal of Combinatorial Optimization*, 2(1):71–109, 1998. 2
- [60] Feng Zhou and Fernando De la Torre. Factorized Graph Matching. *TPAMI*, 38(9):1774–1789, 2016. 2
- [61] Xiaowei Zhou, Menglong Zhu, and Kostas Daniilidis. Multi-image matching via fast alternating minimization. In *ICCV*, 2015. 2, 3, 6, 8