Multi-View Point Registration via Alternating Optimization*

Junchi Yan[12], Jun Wang[5], Hongyuan Zha[34], Xiaokang Yang[1], Stephen M. Chu[2]
yanjunchi@sjtu.edu.cn, j.wang@alibaba-inc.com, zha@cc.gatech.edu, xkyang@sjtu.edu.cn, schu@cn.ibm.com
[1]Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China
[2]IBM Research – China, Shanghai, 201203, China
[3]Software Engineering Institute, East China Normal University, Shanghai, 200062, China
[4]College of Computing, Georgia Institute of Technology, Atlanta, Georgia, 30332, USA
[5]Institue of Data Science and Technology, Alibaba Group, Seattle, WA, 98101, USA

Abstract

Multi-view point registration is a relatively less studied problem compared with two-view point registration. Directly applying pairwise registration often leads to matching discrepancy as the mapping between two point sets can be determined either by direct correspondences or by any intermediate point set. Also, the local two-view registration tends to be sensitive to noises. We propose a novel multi-view registration method, where the optimal registration is achieved via an efficient and effective alternating concave minimization process. We further extend our solution to a general case in practice of registration among point sets with different cardinalities. Extensive empirical evaluations of peer methods on both synthetic data and real images suggest our method is robust to large disturbance. In particular, it is shown that our method outperforms peer point matching methods and performs competitively against graph matching approaches. The latter approaches utilize the additional second-order information at the cost of exponentially increased run-time, thus usually being less efficient.

Introduction

Point matching or registration is a fundamental research topic in computer vision and pattern recognition since many applications rely on accurate geometric model acquisition, including object recognition, localization, tracking, appearance and texture analysis, virtual reality, among others (Lian and Zhang 2012; Tsin and Kanade 2004; Chui and Rangarajan 2003; Zhang 1994; Myronenko and Song 2010). As a basic form, two-view point registration generally consists of two subproblems: i) finding correspondence from two views; ii) estimating the pose transformation. However, the key challenge is a typical chicken-and-egg problem: two subproblems are interlocked, and the overall optimization is highly non-convex. In practice, data acquisition often involves obtaining either intensity or depth data of an object from more than two view points (Dorai et al. 1998). Hence, multi-view point matching or registration has attracted more attention recently because it involves more matching parameters and correspondences among a set of views with potential performance improvements. Fig.1 gives some exemplar results of the proposed multi-view registration method which is applied on real images, where the conventional two-view registration methods often produce unsatisfactory results due to various reasons, such as local occlusions and large transformations, among others.

Despite of the empirical success of two-view point registration methods, the extension to multi-view registration remains as an open problem (Sharp, Lee, and Wehe 2004; Pooja and Govindu 2010). As a straightway matching strategy based on two-view solution, performing pairwise registration sequentially over all views hardly works well (Pulli 1999). Another potential scheme to reuse the two-view registration techniques is repeating the two-view registration on each pair of views and then aligning the pairwise alignment results to derive the final solution. However, this strategy suffers two major limitations: i) different orders of pairwise registration may result in non-unique correspondences; ii) the registration error by the local noise of a pairwise alignment, in an extreme case, may propagate along the pairwise registration sequence and would be further accumulated if no extra information is available to dismiss the noise.

It is generally recognized that additional information from multi-view can help improve the alignment than the case when only a pair of views is given. To address the above issues and leverage multi-view data, we design a novel multi-view registration algorithm via an alternative optimization procedure. We transform the objective function to a concave one over iteration efficiently, which can be minimized by a global optimizer, preventing a local optimum in the sequential cycle of registration process. Extensive empirical studies corroborate the effectiveness of our method that generates more accurate and consistent aligning results.

Related Work

In this section, we provide a structured view on previous work concerning point registration and a more broad topic i.e. finding point correspondence.

Application scenario: multi-view vs. two-view There is extensive work for two-view registration. For instance, given a rough initial registration, the iterative closest point (ICP) method iterates between finding the point correspondence via nearest neighbor method and updating the transforma-

*The work is partially supported by NSF DMS-1317424, NSFC 61129001/F010403 and 61025005/F010403.

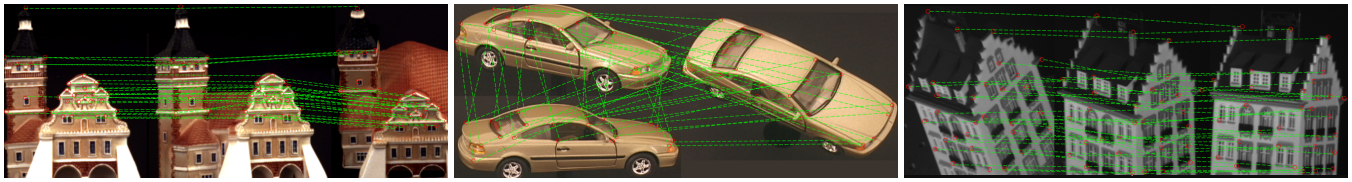

Figure 1: Examples of registration results. From left to right: house and car from the Pose dataset, hotel from the CMU sequence.

tion with the least square error (Chen and Medioni 1992; Zhang 1994). Many other methods, such as the Robust Point Matching (RPM) (Chui and Rangarajan 2003) and the Coherent Point Drift (CPD) method (Myronenko and Song 2010), are designed to perform the pairwise registration. On the contrary, the multi-view alignment problem is less studied and existing work mainly focuses on developing the registration strategies for aligning multiple pairwise results (Bergevin et al. 1996; Sharp, Lee, and Wehe 2004; Pooja and Govindu 2010).

Alignment strategy: simultaneous vs. sequential One intuitive strategy is to perform pair-view registration at a time, and then repeat the pairwise registration in a sequential manner (Chen and Medioni 1992; Turk and Levoy 1994). Such a strategy enjoys the simplicity of both computational overhead and implementation efforts. However, this simple strategy is known to be sensitive to noises. Another alternative registration method is first presented by Bergevin et. al. (Bergevin et al. 1996), and then further refined in (Benjemaa and Schmitt 1997). More recently, (Pooja and Govindu 2010) proposes an extension to the standard ICP algorithm to simultaneously average the redundant information in multi-view data. (A.Torsello, Rodola, and Albarelli 2011) proposes to use dual quaternion and adopts a new distortion measure derived from the screw motion. Note that most of these methods are specifically designed for the setting of the rigid transformation, neglecting non-rigid transformations.

Optimization technique: global vs. heuristic Most of the well-known point registration methods are heuristic. Briefly speaking, one starts from an initial estimation and solves the overall problem by alternately solving the two sub-problems, i.e., pose estimation and point correspondence. The above mentioned ICP based approaches fall into this category. In addition, Robust Point Matching (RPM) (Chui and Rangarajan 2003) relaxes point correspondence to be continuously valued and achieves an optimized solution in the continuous space. The coherent point drift (CPD) method (Myronenko and Song 2010) models point matching under a probabilistic framework and uses expectation maximization (EM) for obtaining optimal solutions. (Lian and Zhang 2012) forms a concave objective and obtains a globally optimal solution. Similarly, (Li and Hartley 2007) achieves a global optimality grounded by the Lipschitz global optimization process. However, the global optimized results are achieved only for the two-view case.

Transformation: similarity vs. affine transformation Given a point $\mathbf{x} \in \mathbb{R}^2$, its *similarity* transformation is calculated as $s\mathbf{R}\mathbf{x} + \mathbf{t}$, where s is the scaling factor, and $\mathbf{R}$, $\mathbf{t}$ denotes rotation matrix and translation vector respectively. The *affine* transformation is computed as $\mathbf{V}\mathbf{x} + \mathbf{t}$ where $\mathbf{V} \in \mathbb{R}^{2\times 2}$ denotes the affine matrix and $\mathbf{t}$ for translation. In particular, for the similarity transformation, if the scalar factor equals 1, then it becomes rigid transformation. In general, both two types are non-rigid transformations. There are also other types of parameterized transformations, such as RBF non-rigid transformation (Chui and Rangarajan 2003; Myronenko and Song 2010), which is parameterized by a set of basis points. etc. In our experiment, we would focus on the two most popular types as described above.

Point registration *vs.* graph matching Graph matching (Tian et al. 2012; Yan et al. 2013; 2014; Zhou and la Torre 2012; Cho, Lee, and Lee 2010) and point registration are closely related formulations, both of which are used for solving the point set correspondence problem. However, there are two important differences between these two approaches. First, graph matching typically explicitly infuses pairwise or higher-order point tuple affinity, which is beyond unary similarity, to derive the objective function. While point registration does not explicitly use such point-tuple affinities between two point sets. Its optimization can be casted as a linear programming problem given known transformation parameters. The other major difference is graph matching does not impose parametric prior or constraints to the transformation between two point sets, yet point registration involves various parameterized transformation models such as affinity or similarity transform, as we discussed above. Therefore, both formulations are mathematically challenging: graph matching is casted to a constrained quadratic (due to including higher-order affinity terms) assignment programming problem which is well known to be NP-hard; In contrast, point registration involves additional unknown transformation parameters which need be estimated jointly with the point correspondence. Note simultaneously optimizing with respect to these two variables are non-convex without a closed form solution.

From Two-view to Multi-view Registration

Before diving into main body, first several notations are defined. Without loss of generality, this paper mainly presents the proposed multi-view registration formulation and algorithm in the context of three views. Let $\mathcal{X} = \{\mathbf{x}_i\}, \mathcal{Y} = \{\mathbf{y}_i\}, \mathcal{Z} = \{\mathbf{z}_i\}$ denote three point sets, whose cardinality equals m, n, l respectively. For point set $\mathcal{X}$, denote each point by $\mathbf{x}_i = [x_i^1, ..., x_i^d] \in \mathbb{R}^d$, so for $\mathcal{Y}$, $\mathcal{Z}$. $\mathbf{P} \in \{0,1\}^{m\times n}$, denotes the point matching matrix with $P_{ij} = 1$ indicating that $\mathbf{x}_i$ corresponds to $\mathbf{y}_j$ and $P_{ij} = 0$ otherwise. The pairwise transform parameters over the three point sets are $\boldsymbol{\theta}_{xy} \in \mathbb{R}^k, \boldsymbol{\theta}_{xz} \in \mathbb{R}^k, \boldsymbol{\theta}_{yz} \in \mathbb{R}^k$, where k is the de-

gree of freedom of the transformation. Let $\mathbf{J}_{xi}$, $\mathbf{J}_{yi}$, $\mathbf{J}_{zi}$ denote the Jacobian matrices of the i-th point in point sets $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, respectively. To make the problem mathematically tractable, we assume the transformation $T(\mathbf{x}_i|\boldsymbol{\theta})$ is linear with respect to its parameters $\boldsymbol{\theta}$. Their matrix forms are $\mathbf{J}_x \in \mathbb{R}^{2m\times d}$, $\mathbf{J}_y \in \mathbb{R}^{2n\times d}$, $\mathbf{J}_z \in \mathbb{R}^{2l\times d}$. In addition, we define $\bar{\mathbf{Y}} \triangleq [\|\mathbf{y}_1\|_2^2, \ldots, \|\mathbf{y}_n\|_2^2]^T$, $\bar{\mathbf{Z}} \triangleq [\|\mathbf{z}_1\|_2^2, \ldots, \|\mathbf{z}_l\|_2^2]^T$. $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ denotes the identity matrix, and $\mathbf{1}_m \in \mathbb{R}^{m\times 1}$ denotes the column vector whose elements being all ones. $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker product of matrices $\mathbf{A}$ and $\mathbf{B}$.

We introduce one important formulation for two-view point registration. Being a baseline approach, this pairwise method is also the starting point of the proposed method.

Baseline: Pairwise Robust Point Matching

For pairwise form of robust point matching (RPM) model (Chui and Rangarajan 2003), the problem is expressed via minimizing an objective w.r.t. $\mathbf{P}_{xy}$ (xy is omitted in Eq.1), $\boldsymbol{\theta}_{xy}$ from $\mathcal{X}$ to $\mathcal{Y}$, which consists of a point-to-point distance term and a regularizer $g(\boldsymbol{\theta}_{xy})$ related to a prior transform $\boldsymbol{\theta}_0$ and a weight matrix $\mathbf{H} \in \mathbb{R}^{k\times k}$:

$$E(\mathbf{P}, \boldsymbol{\theta}_{xy}) = \sum_{i,j} P_{ij}\|\mathbf{y}_j - T(\mathbf{x}_i|\boldsymbol{\theta})\| + g(\boldsymbol{\theta}_{xy}) \tag{1}$$

$$\mathbf{P}\mathbf{1}_n = \mathbf{1}_m, \mathbf{1}_m^T\mathbf{P} \le \mathbf{1}_n, g(\boldsymbol{\theta}_{xy}) = (\boldsymbol{\theta}_{xy} - \boldsymbol{\theta}_0)^T\mathbf{H}(\boldsymbol{\theta}_{xy} - \boldsymbol{\theta}_0)$$

The optimal $\hat{\boldsymbol{\theta}}_{xy}$ is found by zeroing its derivative[1].

$$\hat{\boldsymbol{\theta}}_{xy} = (\mathbf{J}_x^T\mathbf{J}_x + \mathbf{H})^{-1}[\mathbf{J}_x^T(\mathbf{P}_{xy} \otimes \mathbf{I}_d)\mathbf{Y} + \mathbf{H}\boldsymbol{\theta}_0] \tag{2}$$

Proposed Multiple-View RPM Formulation

Without loss of generality, assume $|\mathcal{X}| \le |\mathcal{Y}| \le |\mathcal{Z}|$ and the mapping for $\mathcal{X}\to\mathcal{Y}$ is injection, so for $\mathcal{Y}\to\mathcal{Z}$. Based on these assumptions, in line with footnote (1), we propose the following three-view registration formulation:

$$\begin{aligned} E_{xyz} =& E_{xy} + E_{xz} + E_{yz} \\ =& \sum_{i,j}\|y_j - \mathbf{J}_{xi}\boldsymbol{\theta}_{xy}\|^2 + \sum_{i,k}\|z_k - \mathbf{J}_{xi}\boldsymbol{\theta}_{xz}\|^2 \\ &+ \sum_{j,k}\|z_k - \mathbf{J}_{yj}\boldsymbol{\theta}_{yz}\|^2 + g(\boldsymbol{\theta}_{xy}) + g(\boldsymbol{\theta}_{xz}) + g(\boldsymbol{\theta}_{yz}) \end{aligned} \tag{3}$$

In the explicit form w.r.t. correspondence matching matrix $\mathbf{P}_{xy}$, E_{xy} (similar for E_{xz}, E_{yz}) can be written as:

$$\boldsymbol{\theta}_{xy}^T\mathbf{J}_x^T\mathbf{J}_x\boldsymbol{\theta}_{xy} - 2\boldsymbol{\theta}_{xy}^T\mathbf{J}_x^T(\mathbf{P}_{xy} \otimes \mathbf{I}_d)\mathbf{Y} + \mathbf{1}_m^T\mathbf{P}_{xy}\bar{\mathbf{Y}} + g(\boldsymbol{\theta}_{xy})$$

where $g(\boldsymbol{\theta}) = (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T\mathbf{H}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$. To make the problem mathematically tractable, we assume $\boldsymbol{\theta}_{xy}$, $\boldsymbol{\theta}_{xz}$, $\boldsymbol{\theta}_{yz}$ are independent to each other[2], thus can be decided by the matching matrix $\mathbf{P}_{xy}$, $\mathbf{P}_{xz}$, $\mathbf{P}_{yz}$ independently. By letting $\frac{\partial E}{\partial \boldsymbol{\theta}_{xy}} = \frac{\partial E}{\partial \boldsymbol{\theta}_{xz}} = \frac{\partial E}{\partial \boldsymbol{\theta}_{yz}} = 0$ the optimal $\boldsymbol{\theta}_{xy}$, $\boldsymbol{\theta}_{xz}$, $\boldsymbol{\theta}_{yz}$ can be calculated by Eq.2. As a result, they are eliminated in the objective function and we arrive at a new nonlinear objective function w.r.t. $\mathbf{P}_{xy}$, $\mathbf{P}_{xz}$, $\mathbf{P}_{yz}$ where E_{xy} can be written as follows (similar for E_{xz}, E_{yz})

$$\begin{aligned} E_{xy} =& \mathbf{1}_m^T\mathbf{P}_{xy}\bar{Y} - \left(\mathbf{Y}^T(\mathbf{P}_{xy} \otimes \mathbf{I}_d)^T\mathbf{J}_x + \boldsymbol{\theta}_0^T\mathbf{H}\right)(\mathbf{J}_x^T\mathbf{J}_x + \mathbf{H})^{-1} \\ &\cdot \left(\mathbf{J}_x^T(\mathbf{P}_{xy} \otimes \mathbf{I}_d)^T\mathbf{Y} + \mathbf{H}\boldsymbol{\theta}_0\right) + \boldsymbol{\theta}_0\mathbf{H}\boldsymbol{\theta}_0 \end{aligned}$$

By eliminating the constant terms, the expansion of E_{xy} (similar for E_{xz}, E_{yz}) w.r.t. matching matrix $\mathbf{P}_{xy}$ becomes:

$$\begin{aligned} E_{xy} =& \mathbf{1}_m^T\mathbf{P}_{xy}\bar{\mathbf{Y}} - 2\boldsymbol{\theta}_0^T\mathbf{H}(\mathbf{J}_x^T\mathbf{J}_x + \mathbf{H})^{-1}\mathbf{J}_x^T(\mathbf{P}_{xy} \otimes \mathbf{I}_d)\mathbf{Y} \\ &+ \mathbf{Y}^T(\mathbf{P}_{xy} \otimes \mathbf{I}_d)^T\mathbf{J}_x(\mathbf{J}_x^T\mathbf{J}_x + \mathbf{H})^{-1}\mathbf{J}_x^T(\mathbf{P}_{xy} \otimes \mathbf{I}_d)\mathbf{Y} \end{aligned} \tag{4}$$

To obtain a more clear expression, we vectorize the matrix version of $\mathbf{P}$ to $\mathbf{p}$=vec($\mathbf{P}$) by stacking its rows. By separating the linear E_{xyz}^{lin} and the quadratic parts E_{xyz}^{qd} we reach:

$$\begin{aligned} E_{xyz} =& E_{xyz}^{lin}(\mathbf{p}_{xy}, \mathbf{p}_{xz}, \mathbf{p}_{yz}) - E_{xyz}^{qd}(\mathbf{p}_{xy}, \mathbf{p}_{xz}, \mathbf{p}_{yz}) \\ E_{xyz}^{lin} =& (\mathbf{1}_m^T \otimes \bar{\mathbf{Y}}^T + (\mathbf{K}_x \otimes \mathbf{Y})\mathbf{W}_{xy})\mathbf{p}_{xy} \\ &+ (\mathbf{1}_m^T \otimes \bar{\mathbf{Z}}^T + (\mathbf{K}_x \otimes \mathbf{Z})\mathbf{W}_{xz})\mathbf{p}_{xz} \\ &+ (\mathbf{1}_n^T \otimes \bar{\mathbf{Z}}^T + (\mathbf{K}_y \otimes \mathbf{Z})\mathbf{W}_{yz})\mathbf{p}_{yz} \\ E_{xyz}^{qd} =& \|(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Y}^T\mathbf{W}_{xy}\mathbf{p}_{xy}\|^2 + \|(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Z}^T\mathbf{W}_{xz}\mathbf{p}_{xz}\|^2 \\ &+ \|(\mathbf{U}_y\mathbf{J}_y^T) \otimes \mathbf{Z}^T\mathbf{W}_{yz}\mathbf{p}_{yz}\|^2 \end{aligned} \tag{5}$$

where $\mathbf{K}_x = 2(\boldsymbol{\theta}_0^T\mathbf{H}(\mathbf{J}_x^T\mathbf{J}_x + \mathbf{H})^{-1}\mathbf{J}_x^T)$ and similar for $\mathbf{K}_y$, $\mathbf{K}_z$. In the rest of the paper, we further assume $|\mathcal{Y}|$=$|\mathcal{Z}|$ and the correspondence is a bijection, then it is mathematically sound (refer to footnote 1) to rewrite the objective as E_{xyz}=E_{xy}+E_{xz}+E_{zy} by replacing E_{yz} with E_{zy}. The reason for introducing this alternative writing of the objective function would be clear in the next section.

$$\begin{aligned} E_{xyz} =& E_{xyz}^{lin}(\mathbf{p}_{xy}, \mathbf{p}_{xz}, \mathbf{p}_{zy}) - E_{xyz}^{qd}(\mathbf{p}_{xy}, \mathbf{p}_{xz}, \mathbf{p}_{zy}) \\ E_{xyz}^{lin} =& (\mathbf{1}_m^T \otimes \bar{\mathbf{Y}}^T + (\mathbf{K}_x \otimes \mathbf{Y})\mathbf{W}_{xy})\mathbf{p}_{xy} \\ &+ (\mathbf{1}_m^T \otimes \bar{\mathbf{Z}}^T + (\mathbf{K}_x \otimes \mathbf{Z})\mathbf{W}_{xz})\mathbf{p}_{xz} \\ &+ (\mathbf{1}_l^T \otimes \bar{\mathbf{Y}}^T + (\mathbf{K}_z \otimes \mathbf{Y})\mathbf{W}_{zy})\mathbf{p}_{zy} \\ E_{xyz}^{qd} =& \|(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Y}^T\mathbf{W}_{xy}\mathbf{p}_{xy}\|^2 + \|(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Z}^T\mathbf{W}_{xz}\mathbf{p}_{xz}\|^2 \\ &+ \|(\mathbf{U}_z\mathbf{J}_z^T) \otimes \mathbf{Y}^T\mathbf{W}_{zy}\mathbf{p}_{zy}\|^2 \end{aligned} \tag{6}$$

where $\mathbf{U}_x^T\mathbf{U}_x = (\mathbf{J}_x^T\mathbf{J}_x + \mathbf{H})^{-1}$, $\mathbf{U}_y^T\mathbf{U}_y = (\mathbf{J}_y^T\mathbf{J}_y + \mathbf{H})^{-1}$, $\mathbf{U}_z^T\mathbf{U}_z = (\mathbf{J}_z^T\mathbf{J}_z + \mathbf{H})^{-1}$, $\mathbf{W}_{xy} \triangleq \mathbf{I}_m \otimes [\mathbf{I}_n \otimes e_d^1, \ldots, \mathbf{I}_n \otimes e_d^d]^T$, $\mathbf{W}_{xz} \triangleq \mathbf{I}_m \otimes [\mathbf{I}_l \otimes e_d^1, \ldots, \mathbf{I}_l \otimes e_d^d]^T$, $\mathbf{W}_{yz} \triangleq \mathbf{I}_n \otimes [\mathbf{I}_l \otimes e_d^1, \ldots, \mathbf{I}_l \otimes e_d^d]^T$, $\mathbf{W}_{zy} \triangleq \mathbf{I}_l \otimes [\mathbf{I}_n \otimes e_d^1, \ldots, \mathbf{I}_n \otimes e_d^d]^T$, such that $\mathbf{W}$ satisfies vec($\mathbf{P} \otimes \mathbf{I}_d$)=$\mathbf{W}$ vec($\mathbf{P}$) and e_d^i is a d-dimensional column vector with the i-th element being 1 and the rest 0s.

Alternating Optimization

The rough idea for alternating optimization is to fix one pairwise alignment variable and then update the other in a rotating manner. For three-view scenario as we discussed so far,

[1] We assume each point in $\mathcal{X}$ has its counterpart in $\mathcal{Y}$: $\mathbf{P}_{xy}\mathbf{1}_n$=$\mathbf{1}_m$. Therefore, we can simplify the mathematics and reach Eq.2 which is the foundation of later derivation of this paper. This constraint is widely used in related literature such as (Chui and Rangarajan 2003; Lian and Zhang 2012) and reference therein.

[2] The reader may have the concern that $\boldsymbol{\theta}_{xy}$, $\boldsymbol{\theta}_{xz}$, $\boldsymbol{\theta}_{yz}$ should be compatible with each other while computing them independently as simplified in the paper in general cannot ensure this compatibility. In fact, from our experiments, we empirically found the estimated $\boldsymbol{\theta}_{xy}$, $\boldsymbol{\theta}_{xz}$, $\boldsymbol{\theta}_{yz}$ are roughly consistent. Note one way of mitigating the consistency issue is that to some extent we bring this compatibility back since the point correspondence matching consistency is preserved in our formulation i.e. $\mathbf{P}_{xy} = \mathbf{P}_{xz}\mathbf{P}_{zy}$.

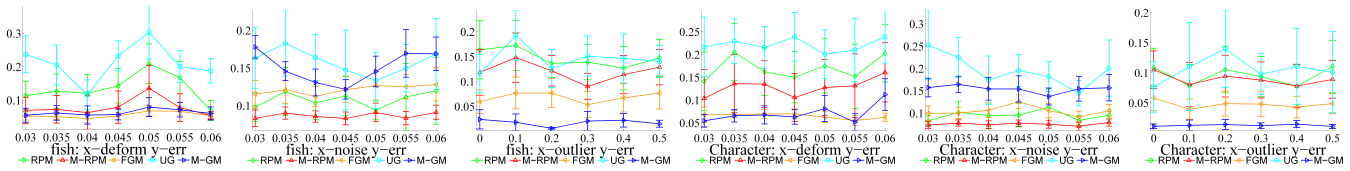

Figure 2: Similarity transform test on Chui-Rangarajan data: Point matching distance error by 5 methods (RPM, M-RPM, FGM, UG, M-GM) against deformation, noise and outlier. The bars indicate the standard deviation of the error over 20 random trials.

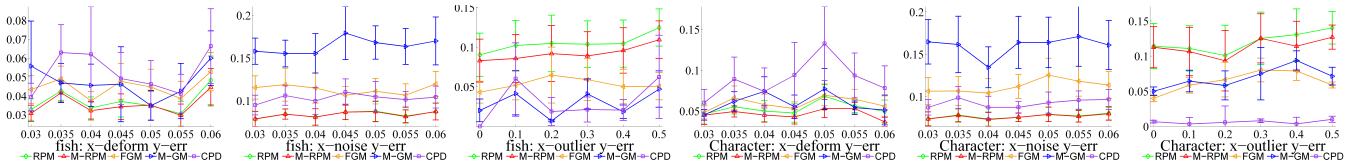

Figure 3: Affine transform test on Chui-Rangarajan data: Point matching distance errors by 5 methods (RPM, M-RPM, FGM, M-GM, CPD) against deformation, noise and outlier. The bars indicate the standard deviation of the error over 20 random trials.

there are three variables $\mathbf{p}_{xy}$, $\mathbf{p}_{xz}$, $\mathbf{p}_{yz}$. Remember it is still under the assumption as stated in Sec.3 that $|\mathcal{X}| \leq |\mathcal{Y}| = |\mathcal{Z}|$ and $\mathbf{P}_{xz}$, $\mathbf{P}_{xy}$ are injection and $\mathbf{P}_{zy}$ bijection. Therefore, the two equations $\mathbf{P}_{xz} = \mathbf{P}_{xy}\mathbf{P}_{yz}$, $\mathbf{P}_{xy} = \mathbf{P}_{xz}\mathbf{P}_{zy}$ both hold. However, it would cause node-mapping information loss if one moves $\mathbf{P}_{yz}$ ($\mathbf{P}_{zy}$) to the left side of the equation: $\mathbf{P}_{yz} = \mathbf{P}_{yx}\mathbf{P}_{xz}$ ($\mathbf{P}_{zy} = \mathbf{P}_{zx}\mathbf{P}_{xy}$). In fact these two equations do not hold because $\mathcal{X}$ only corresponds to a subset of $\mathcal{Y}$ and $\mathcal{Z}$. In fact, the derived $\mathbf{P}^d_{yz} = \mathbf{P}_{yx}\mathbf{P}_{xz}$ cannot satisfy the injection requirement for E_{yz} thus Eq.2 is broken.

Based on the above analysis, we will chose to fix one of $\mathbf{p}_{xy}$ and $\mathbf{p}_{xz}$ alternatively and optimize with respect to $\mathbf{p}_{yz}$ ($\mathbf{p}_{zy}$) and the other. Specifically, given known $\mathbf{p}_{xy}$, we have $\mathbf{p}_{xz} = (\mathbf{I}_n \otimes \mathbf{P}_{xy})\mathbf{p}_{yz}$; when $\mathbf{p}_{xz}$ is fixed, it becomes $\mathbf{p}_{xy} = (\mathbf{I}_n \otimes \mathbf{P}_{xz})\mathbf{p}_{zy}$. Attention shall be taken here to the new form $\mathbf{p}_{zy}$ instead of $\mathbf{p}_{yz}$. One may argue to avoid $\mathbf{p}_{zy}$ (accordingly E_{zy}) by using $\mathbf{p}_{yx} = (\mathbf{I}_n \otimes \mathbf{P}_{xz})\mathbf{p}_{yz}$. However $\mathbf{p}_{yx}$ reverses the term E_{yx} which breaks the assumption in footnote (1) for $\mathcal{A} \to \mathcal{B}$ is an injection in the objective function E_{AB}. In contrast, as $\mathcal{Y} \to \mathcal{Z}$ is a bijection, reverse E_{yz} to E_{zy} is not harmful. Therefore, it becomes clear why we replace the term E_{yz} with E_{zy} as shown in Eq.6, in addition with Eq.5.

Note we impose no assumption about the property of the point distribution or the topology of the point sets. Hence, our approach and analysis are generic. In what follows, we will give a detailed description of the proposed method under the context of three-view registration, which will be extended to a more general setting with multiple views.

Three-view Alternating Optimization Mechanism

By fixing $\mathbf{p}_{xy}$ we can discard E_{xy} in Eq.5, and the objective function that is only with respect to $\mathbf{p}_{xz}$, $\mathbf{p}_{yz}$ becomes

$$\begin{aligned} E_{xzyz} =&(\mathbf{1}_m^T \otimes \bar{\mathbf{Z}}^T + (\mathbf{K}_x \otimes \mathbf{Z})\mathbf{W}_{xz})\mathbf{p}_{xz} - \|[(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Z}^T]\mathbf{W}_{xz}\mathbf{p}_{xz}\|^2 \\ &+(\mathbf{1}_n^T \otimes \bar{\mathbf{Y}}^T + (\mathbf{K}_y \otimes \mathbf{Z})\mathbf{W}_{yz})\mathbf{p}_{yz} - \|[(\mathbf{U}_y\mathbf{J}_y^T) \otimes \mathbf{Z}^T]\mathbf{W}_{yz}\mathbf{p}_{yz}\|^2 \end{aligned}$$

The two quadratic terms in E_{xzyz} can be written as:

$$\begin{aligned} E^{qd}_{xzyz} =&\mathbf{p}_{xz}^T[(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Z}^T\mathbf{W}_{xz}]^T[(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Z}^T\mathbf{W}_{xz}]\mathbf{p}_{xz} \\ &+\mathbf{p}_{yz}^T[(\mathbf{U}_y\mathbf{J}_y^T) \otimes \mathbf{Z}^T\mathbf{W}_{yz}]^T[(\mathbf{U}_y\mathbf{J}_y^T) \otimes \mathbf{Z}^T\mathbf{W}_{yz}]\mathbf{p}_{yz} \end{aligned}$$

Alternatively, one can fix $\mathbf{p}_{xz}$ and remove the constant term E_{xz} from Eq.6 to obtain the new objective function:

$$\begin{aligned} E_{xyzy} =&(\mathbf{1}_m^T \otimes \bar{\mathbf{Y}}^T + (\mathbf{K}_x \otimes \mathbf{Y})\mathbf{W}_{xy})\mathbf{p}_{xy} - \|[(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Y}^T]\mathbf{W}_{xy}\mathbf{p}_{xy}\|^2 \\ &+(\mathbf{1}_l^T \otimes \bar{\mathbf{Y}}^T + (\mathbf{K}_z \otimes \mathbf{Y})\mathbf{W}_{zy})\mathbf{p}_{zy} - \|[(\mathbf{U}_z\mathbf{J}_z^T) \otimes \mathbf{Y}^T]\mathbf{W}_{zy}\mathbf{p}_{zy}\|^2 \end{aligned}$$

Similarly, its quadratic term can be written as:

$$\begin{aligned} E^{qd}_{xyzy} =&\mathbf{p}_{xy}^T[(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Y}^T\mathbf{W}_{xy}]^T[(\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Y}^T\mathbf{W}_{xy}]\mathbf{p}_{xy} \\ &+\mathbf{p}_{zy}^T[(\mathbf{U}_z\mathbf{J}_z^T) \otimes \mathbf{Y}^T\mathbf{W}_{zy}]^T[(\mathbf{U}_z\mathbf{J}_z^T) \otimes \mathbf{Y}^T\mathbf{W}_{zy}]\mathbf{p}_{zy} \end{aligned}$$

To simplify the exposition of the objective function over iterations that would be shown later, we introduce $\mathbf{B}_{xz} = (\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Z}^T\mathbf{W}_{xz}$, $\mathbf{B}_{yz} = (\mathbf{U}_{xy}\mathbf{J}_y^T) \otimes \mathbf{Z}^T\mathbf{W}_{yz}$, $\mathbf{B}_{xy} = (\mathbf{U}_x\mathbf{J}_x^T) \otimes \mathbf{Y}^T\mathbf{W}_{xy}$, and $\mathbf{B}_{zy} = (\mathbf{U}_z\mathbf{J}_z^T) \otimes \mathbf{Y}^T\mathbf{W}_{zy}$, $\mathbf{F}_{xy} = \mathbf{I}_n \otimes \mathbf{P}_{xy}$, $\mathbf{F}_{xz} = \mathbf{I}_n \otimes \mathbf{P}_{xz}$. Based on these notations, we further define: $\mathbf{C}_{yzxz} = \begin{bmatrix} \mathbf{B}_{yz} \\ \mathbf{B}_{xz}\mathbf{F}_{xy} \end{bmatrix}$, $\mathbf{C}_{zyxy} = \begin{bmatrix} \mathbf{B}_{zy} \\ \mathbf{B}_{xy}\mathbf{F}_{xz} \end{bmatrix}$. Using the two objective function formula E_{xzyz}, E_{xyzy} and the above notations, the iteration can be compactly described as follows in a rotating manner.

Step 1) Fix $\mathbf{p}_{xy}$ update $\mathbf{p}_{yz}$ for minimizing partial objective function E_{xz}+E_{yz}. Replace $\mathbf{p}_{xz}$ by $\mathbf{p}_{xz}$=$\mathbf{F}_{xy}\mathbf{p}_{yz}$, the quadratic term in E_{xzyz} can be rewritten compactly as:

$$\begin{aligned} E^{qd}_{xzyz} &= \mathbf{p}_{yz}^T\mathbf{B}_{yz}^T\mathbf{B}_{yz}\mathbf{p}_{yz} + \mathbf{p}_{yz}^T(\mathbf{B}_{xz}\mathbf{F}_{xy})^T(\mathbf{B}_{xz}\mathbf{F}_{xy})\mathbf{p}_{yz} \\ &= \mathbf{p}_{yz}^T\mathbf{C}_{yzxz}^T\mathbf{C}_{yzxz}\mathbf{p}_{yz} \end{aligned}$$

Now we reach the objective function with respect to $\mathbf{p}_{yz}$:

$$E(\mathbf{p}_{yz}) = \mathbf{b}_{yz}^T\mathbf{p}_{yz} - \|\mathbf{C}_{yzxz}\mathbf{p}_{yz}\|^2 \tag{7}$$

Step 2) Fix $\mathbf{p}_{xz}$ update $\mathbf{p}_{zy}$ for minimizing partial objective function E_{xy}+E_{zy}. Replace $\mathbf{p}_{xy}$ by $\mathbf{p}_{xy}$=$\mathbf{F}_{xz}\mathbf{p}_{zy}$, which leads to E_{xyzy}=$\mathbf{p}_{zy}^T\mathbf{C}_{zyxy}^T\mathbf{C}_{zyxy}\mathbf{p}_{zy}$. Then we have:

$$E(\mathbf{p}_{zy}) = \mathbf{b}_{zy}^T\mathbf{p}_{zy} - \|\mathbf{C}_{zyxy}\mathbf{p}_{zy}\|^2 \tag{8}$$

where $\mathbf{b}_{yz}$ and $\mathbf{b}_{zy}$ have the following forms:

$$\begin{aligned} \mathbf{b}_{yz} =&\mathbf{1}_n^T \otimes \bar{\mathbf{Z}}^T + (\mathbf{K}_y \otimes \mathbf{Z}^T)\mathbf{W}_{yz} + \left(\mathbf{1}_m^T \otimes \bar{\mathbf{Z}}^T + (\mathbf{K}_x \otimes \mathbf{Z}^T)\mathbf{W}_{xz}\right)\mathbf{F}_{xy} \\ \mathbf{b}_{zy} =&\mathbf{1}_l^T \otimes \bar{\mathbf{Y}}^T + (\mathbf{K}_z \otimes \mathbf{Y}^T)\mathbf{W}_{zy} + \left(\mathbf{1}_l^T \otimes \bar{\mathbf{Y}}^T + (\mathbf{K}_x \otimes \mathbf{Y}^T)\mathbf{W}_{xy}\right)\mathbf{F}_{xz} \end{aligned}$$

Algorithm 1 Alternating concave optimization for three-view point registration

1: **Input:** $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ ($\mathcal{Y}$, $\mathcal{Z}$ are 'complete' point sets), θ_0, T_{max};
2: **Output:** Consistent point matching solution $\mathbf{P}_{xy}, \mathbf{P}_{xz}, \mathbf{P}_{yz}$;
3: **Procedure:**
4: Obtain $\mathbf{P}_{xy}$ via minimizing Eq.1 for initialization;
5: **for** $t = 1 : T_{max}$ **do**
6: Fix $\mathbf{P}_{xy}$, update $\mathbf{P}_{yz}$ ($\mathbf{P}_{xz}$ by $\mathbf{P}_{xz}$=$\mathbf{P}_{xy}\mathbf{P}_{yz}$) via minimizing Eq.7 by concave optimization (Lian and Zhang 2012);
7: Fix $\mathbf{P}_{xz}$, update $\mathbf{P}_{zy}$ ($\mathbf{P}_{xy}$ by $\mathbf{P}_{xy}$=$\mathbf{P}_{xz}\mathbf{P}_{zy}$)via minimizing Eq.8 by concave optimization (Lian and Zhang 2012);
8: **end for**

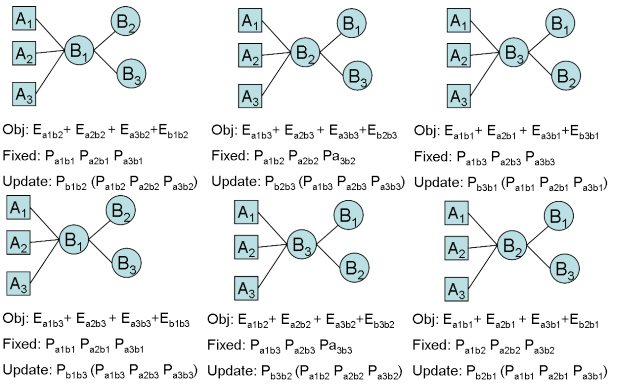

Figure 4: Illustration for the alternating optimization order, given complete point sets $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$ and incomplete sets $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$. Two rows denote two possible looping alternating optimization iteration paths, respectively. Top row: $(\mathbf{P}_{a1b1}, \mathbf{P}_{a2b1}, \mathbf{P}_{a3b1}) \rightarrow (\mathbf{P}_{a1b2}, \mathbf{P}_{a2b2}, \mathbf{P}_{a3b2}) \rightarrow (\mathbf{P}_{a1b3}, \mathbf{P}_{a2b3}, \mathbf{P}_{a3b3}) \rightarrow (\mathbf{P}_{a1b1}, \mathbf{P}_{a2b1}, \mathbf{P}_{a3b1}) \rightarrow \cdots$; Bottom row: $(\mathbf{P}_{a1b1}, \mathbf{P}_{a2b1}, \mathbf{P}_{a3b1}) \rightarrow (\mathbf{P}_{a1b3}, \mathbf{P}_{a2b3}, \mathbf{P}_{a3b3}) \rightarrow (\mathbf{P}_{a1b2}, \mathbf{P}_{a2b2}, \mathbf{P}_{a3b2}) \rightarrow (\mathbf{P}_{a1b1}, \mathbf{P}_{a2b1}, \mathbf{P}_{a3b1}) \rightarrow \cdots$.

After a series of derivation and transformation, the above two formulations enable us to employ the method used in (Lian and Zhang 2012) for solving the minimization problems of (7) and (8). Specifically, the objective function (7) or (8), in its concave quadratic program form, is first transformed into a separable form via eigen decomposition, and then derives the convex envelope of the resulting function over a rectangular region. Then, the normal rectangular algorithm (Horst and Tuy 1996) is used, as a tailored Branch-and-bound approach. Note that after the Eigen decomposition based linear transformation, the number of quadratic terms shrinks to be the number of transformation parameters, thus the method becomes more efficient. Readers are referred to (Lian and Zhang 2012) for more details. The overall algorithmic chart of our method specifically for three-view point registration is described in Alg.1.

Multiple-View Extension

Now we discuss how to generalize our method to the N-view registration problem with unequal cardinality of point sets. Denote the "complete" point sets as $\{\mathcal{B}\}_{i=1}^{N} = \{\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_N\}$, and those "incomplete" point sets $\{\mathcal{A}\}_{i=1}^{M} = \{\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_M\}$ such that the node mapping from $\mathcal{A}_i (i = 1, \ldots, M)$ to $\mathcal{B}_j (j = 1, \ldots, N)$ is an injection which satisfies the assumption in footnote (1), and the correspondences in any pair of $\mathcal{B}_i, \mathcal{B}_j$ is bijection. Now we use a_i to denote the index of $\mathcal{A}_i$ and b_i for $\mathcal{B}_i$. In general, the objective function suited to our specific alternating optimization method is tailored to the following form[3]:

$$E = \sum_{b_j=1}^{N} \sum_{a_i=1}^{M} E_{a_i b_j} + \sum_{b_i, b_j = 1, b_i \neq b_j}^{N} E_{b_i b_j} \qquad (9)$$

One can choose an anchor point set from $\{\mathcal{B}\}_{i=1}^{N}$ in rotation. The anchor view serves as the bridge to diffuse the information from $\{\mathcal{A}\}_{i=1}^{M}$ to $\{\mathcal{B}\}_{i=1}^{N}$ and within $\{\mathcal{B}\}_{i=1}^{N}$. Without loss of generality, in iteration k, view $\mathcal{B}_k$ is chosen as the anchor view. By using $\mathbf{p}_{a_u b_k}^{(k-1)}$ which is updated in the previous iteration k-1, for the bijection $\mathbf{p}_{b_k b_v}^{(k)}$ between $\mathcal{B}_k$ and $\mathcal{B}_v$ ($v \neq k$), one can optimize its associated partial objective function $E(\mathbf{p}_{b_k b_v}^{(k)}) = E_{b_k b_v} + \sum_{u=1}^{M} E_{a_u b_v}$ by rewriting $\mathbf{p}_{a_u b_v}^{(k)} = (\mathbf{I}_{a_u} \otimes \mathbf{P}_{a_u b_k}^{(k-1)}) \mathbf{p}_{b_k b_v}^{(k)}$ to replace $\mathbf{p}_{a_u b_v}^{(k)}$ with $\mathbf{p}_{b_k b_v}^{(k)}$. In consequence, one can update $\mathbf{p}_{b_k b_v}^{(k)}$ by applying the concave optimization algorithm similar to solving the problem of (7) and (8). Accordingly, $\mathbf{p}_{a_u b_{k+1}}^{(k)}$ is also updated which would be used in the next iteration. To further concretize our idea, Fig.4 illustrates two possible iteration rotation orders given three "complete" views $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ and three "incomplete" $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$. In practice, in order to obtain such "complete" point sets given a large number of views, one can divide the whole batch into subsets, and performs our method in each of the subset where the "complete" views exist.

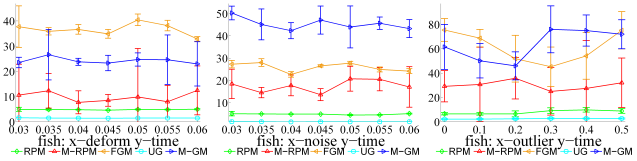

Figure 5: Mean run-time (in seconds) for similarity transformation test on Chui-Rangarajan dataset by 5 methods (RPM, M-RPM, FGM, UG, M-GM). Three types of disturbance, i.e. deformation, noise, outlier are measured, respectively.

Experiments

We implement all the competing methods in Matlab R2009 on a desktop with dual 2.53GHz CPU and 3G memory. Since all methods output point correspondences, we use the correspondences computed by a method to find the best transformation between the two point sets, and define the error as the mean of the Euclidean distances between the transformed model points and their ground truth.

[3]For $E_{a_i a_j}$ where the mapping from $\mathcal{A}_i$ to $\mathcal{A}_j$ is an injection, they cannot be incorporated under the current alternating optimization framework. This is because it is possible that the derived $\mathbf{P}_{a_i a_j}$=$\mathbf{P}_{a_i b_k} \mathbf{P}_{b_k a_j}$ used in $E_{a_i a_j}$ would break the assumption $\mathbf{P}_{a_i a_j}\mathbf{1}$=$\mathbf{1}$ if $\mathcal{A}_i$, $\mathcal{A}_j$ corresponds to different subsets.

The synthetic data used in this paper is generated by the template from Chui-Rangarajan data sets (Chui and Rangarajan 2003). The tested real image data is from CMU hotel and house sequences. The other two sequences (volvo and house) from the pose database (Vikstn et al. 2009).

In line with (Lian and Zhang 2012)'s experimental protocol, rotation-invariant and rotation non-invariant experiments are performed by setting a regularized and non-regularized transformation parameter θ, respectively. For rotation-invariant tests, we use 2-D *similarity* transformation which has been shown a good tradeoff between transformation flexibility and matching robustness. For rotation sensitive tests, we use *affine* transformation in our method as it has small number of parameters while has a certain flexibility to handle non-rigid transformation.

For comparison, we use the two-view registration RPM formulation (Lian and Zhang 2012) as the baseline, and the recent factorized graph matching method (FGM) (Zhou and la Torre 2012). Note these two-view methods both cannot ensure the matching consistency across different view-pairs. The recent graph matching method (M-GM) (Yan et al. 2013) is also tested by following the authors' settings. We further evaluate UG (Caetano and Caelli 2006) and CPD (Myronenko and Song 2010) for similarity and affine transformation tests, respectively. *We term the proposed multi-view robust point matching as M-RPM (red curves in plots)*. Note only the M-GM method and our method can ensure the matching consistency among all the compared methods.

Experiments on Synthetic Data

Similarity transform test For similarity transform test which is rotation invariant (non-regularized $\boldsymbol{\theta}$ in our method), the RPM (Lian and Zhang 2012), FGM (Zhou and la Torre 2012), M-GM (Yan et al. 2013) and UG (Caetano and Caelli 2006) are evaluated by non-rigid deformation, noise and outliers respectively. For each test, in addition with these disturbances, random rotations are imposed on the other two of the three point sets respectively. The matching errors, which follow the protocol of (Lian and Zhang 2012), are shown in Fig.2. It can be seen that for all deformation, noise and outlier tests, our method often outperforms other point-matching methods. The average run time shown in Fig.5. While Fig.7 shows a visual example of how accuracy is boosted during two iterations by our method, meanwhile, *still being more efficient than the higher-order graph matching methods*: FGM and M-GM, although outperform in the outlier test, yet being the slowest across tests.

Affine transform test For rotation sensitive affine transformation (regularized $\boldsymbol{\theta}$ in our method), the matching errors are displayed in Fig.3, where the deformation setting is similar to the similarity transformation test with no random rotation imposed. It can be seen that our method outperforms for deformation and noise tests significantly while being less effective on the outlier tests. Compared with the RPM method, the performance gain is consistent and notable.

Unequal size test We perform clutter test where one of the three point sets has less points than the other two. Fig.8 shows the improved alignment between the smaller-size object against the whole template, for the similarity transformation deform test on the used synthetic data. Compared with the baseline RPM, our method's improvement is gained by incorporating an additional, but also deformed whole template. We further perform this test over different completion level as shown in the third column of Fig.8.

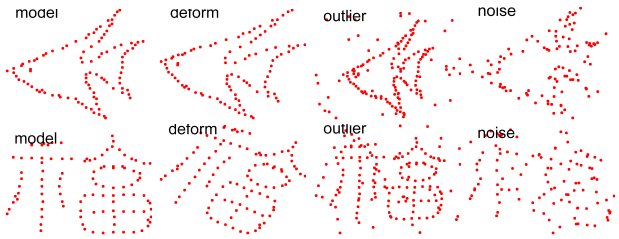

Figure 6: Examples of synthetic point sets of fish & Chinese character from (Chui and Rangarajan 2003). From left to right: reference template, deformation, outlier and noise.

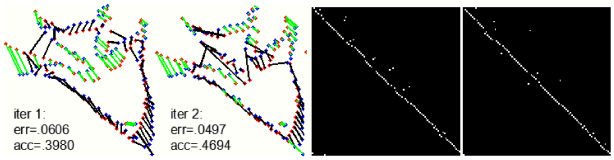

Figure 7: Iteration illustration: two iterations on similarity testing. Left two: registration results; right two: resulting estimated binary correspondence matrix, whose ground truth is the identity matrix.

Real Image Sequences

The matching accuracies (fraction of correct correspondences) against different frame intervals and view rotation angles by using different methods are shown in Fig.9. It can be seen that our method also achieves the best accuracy for most of the test cases. Note other methods like UG and the baseline RPM suffer from the performance fluctuation due to the periodic appearance variation of the testing image sequence, while our method tends to be more robust.

Discussion and Summary

The above experimental results suggest that the proposed method mostly outperform peer point matching methods, and performs competitively against graph matching methods, which utilize the second-order information with an exponential exploration time costs (Zhou and la Torre 2012; Yan et al. 2013). As a result, our method is more efficient than the comparing graph matching approaches.

Conclusion

We have proposed a novel formulation and an efficient optimization solution for multi-view registration. It ensures consistent matching solutions, and is able to handle point sets of unequal sizes. Extensive tests on both synthetic and real

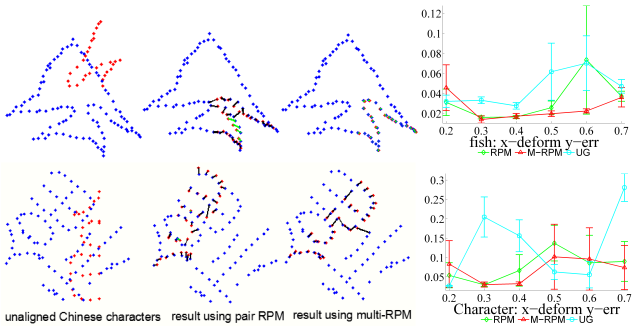

Figure 8: Matching errors of synthetic clutter test by similarity transformation (three point sets are used). From left to right: i) clutter before matching; ii) matching using the two-view RPM baseline (Lian and Zhang 2012); iii) our method; iv) average matching errors by varying the ratio of sub-set to whole-set over 10 tests for 3 methods: RPM, M-RPM, UG.

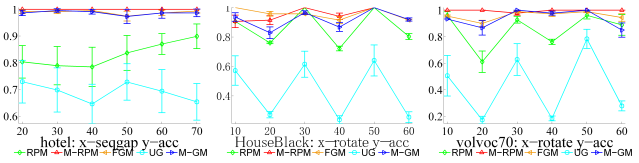

Figure 9: Matching accuracy on CMU and Pose datasets by 5 methods (RPM, M-RPM, FGM, UG, M-GM). From left to right: CMU hotel, Pose house, Pose volvo car.

datasets are performed with more promising results compared with other point registration methods, while being more efficient than higher-order graph matching methods.

Appendix

Why Eq.2 depends on $\mathbf{P}_{xy}\mathbf{1}_n = \mathbf{1}_m$

Here we give the derivation details of Eq.2 to show why it depends on $\mathbf{P}_{xy}\mathbf{1}_n = \mathbf{1}_m$. The matrix form of Eq.1 is:

$$\sum_{i,j} P_{ij}(\mathbf{y}_j^T\mathbf{y}_j + \boldsymbol{\theta}^T\mathbf{J}_{xi}^T\mathbf{J}_{xi}\boldsymbol{\theta} - 2\boldsymbol{\theta}^T\mathbf{J}_{xi}^T\mathbf{y}_j) + (\boldsymbol{\theta}-\boldsymbol{\theta}_0)^T\mathbf{H}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)$$

$$=\boldsymbol{\theta}^T\mathbf{J}_x^T(\mathrm{diag}(\mathbf{P}\mathbf{1}_n)\otimes\mathbf{I}_d)\mathbf{J}_x\boldsymbol{\theta} - 2\boldsymbol{\theta}^T\mathbf{J}_x^T(\mathbf{P}\otimes\mathbf{I}_d)\mathbf{Y} + \mathbf{1}_m^T\mathbf{P}\bar{\mathbf{Y}} + g(\boldsymbol{\theta})$$

$$=\boldsymbol{\theta}^T\mathbf{J}_x^T\mathbf{J}_x\boldsymbol{\theta} - 2\boldsymbol{\theta}^T\mathbf{J}_x^T(\mathbf{P}\otimes\mathbf{I}_d)\mathbf{Y} + \mathbf{1}_m^T\mathbf{P}\bar{\mathbf{Y}} + (\boldsymbol{\theta}-\boldsymbol{\theta}_0)^T\mathbf{H}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)$$

Zeroing the following partial derivative, we obtain Eq.2.

$$2(\mathbf{J}_x^T\mathbf{J}_x + \mathbf{H})\boldsymbol{\theta} - 2\mathbf{J}_x^T(\mathbf{P}\otimes\mathbf{I}_d)\mathbf{Y} - 2\mathbf{H}\boldsymbol{\theta}_0$$

References

A.Torsello; Rodola, E.; and Albarelli, A. 2011. Multi-view registration via graph diffusion of dual quaternions. In *CVPR*.

Benjemaa, R., and Schmitt, F. 1997. Fast global registration of 3d sampled surfaces using a multi-z-buffer technique. In *Int. conf. on 3D Digital Imaging and Modeling*.

Bergevin, R.; Soucy, M.; Gagnon, H.; and Laurendeau, D. 1996. Towards a general multi-view registration technique. *IEEE Trans. PAMI*.

Caetano, T., and Caelli, T. 2006. A unified formulation of invariant point pattern matching. In *ICPR*.

Chen, Y., and Medioni, G. 1992. Object modeling by registration of multiple range images. *IVC*.

Cho, M.; Lee, J.; and Lee, K. M. 2010. Reweighted random walks for graph matching. In *ECCV*.

Chui, H., and Rangarajan, A. 2003. A new point matching algorithm for non-rigid registration. *CVIU*.

Dorai, C.; Wang, G.; Jain, A. K.; and Mercer, C. 1998. Registration and integration of multiple object views for 3d model construction. *IEEE Trans. PAMI*.

Horst, R., and Tuy, H. 1996. *Global Optimization, Deterministic Approaches*. SpringerVerlag.

Li, H., and Hartley, R. 2007. The 3d-3d registration problem revisited. In *ICCV*.

Lian, W., and Zhang, L. 2012. Robust point matching revisited: A concave optimization approach. In *ECCV*.

Myronenko, A., and Song, X. 2010. Point set registration: Coherent point drift. *IEEE Trans. PAMI*.

Pooja, A., and Govindu, V. M. 2010. A multi-view extension of the icp algorithm. In *Indian Conference on Computer Vision, Graphics and Image Processing*.

Pulli, K. 1999. Multiview registration for large data sets. In *Int. conf. on 3D Digital Imaging and Modeling*.

Sharp, G.; Lee, S.; and Wehe, D. 2004. Multiview registration of 3d scenes minimizing error between coordinate frames. *IEEE Trans. PAMI*.

Tian, Y.; Yan, J.; Zhang, H.; Zhang, Y.; Yang, X.; and Zha, H. 2012. On the convergence of graph matching: Graduated assignment revisited. In *ECCV*.

Tsin, Y., and Kanade, T. 2004. A correlation-based approach to robust point set registration. In *ECCV*.

Turk, G., and Levoy, M. 1994. Zippered polygon meshes from range images. In *SIGGRAPH*.

Vikstn, F.; Forssn, P.; Johansson, B.; and Moe, A. 2009. Comparison of local image descriptors for full 6 degree-of-freedom pose estimation. In *ICRA*.

Yan, J.; Tian, Y.; Zha, H.; Yang, X.; and Zhang, Y. 2013. Joint optimization for consistent multiple graph matching. In *ICCV*.

Yan, J. C.; Li, Y.; Liu, W.; Zha, H. Y.; Yang, X. K.; and Chu, S. M. 2014. Graduated consistency-regularized optimization for multi-graph matching. In *ECCV*.

Zhang, Z. 1994. Iterative point matching for registration of free-form curves and surfaces. *IJCV*.

Zhou, F., and la Torre, F. D. 2012. Factorized graph matching. In *CVPR*.