Energy based multi-model *fitting & matching* for 3D reconstruction

Hossam Isack
University of Western Ontario
isack.hossam@gmail.com

Yuri Boykov
University of Western Ontario
yboykov@csd.uwo.ca

Abstract

Standard geometric model fitting methods take as an input a fixed set of feature pairs greedily matched based only on their appearances. Inadvertently, many valid matches are discarded due to repetitive texture or large baseline between view points. To address this problem, matching should consider both feature appearances and geometric fitting errors. We jointly solve feature matching and multi-model fitting problems by optimizing one energy. The formulation is based on our generalization of the assignment problem and its efficient min-cost-max-flow solver. Our approach significantly increases the number of correctly matched features, improves the accuracy of fitted models, and is robust to larger baselines.

1. Introduction

Many existing methods for model fitting and 3D structure estimation use pre-matched image features as an input (bundle adjustment, homography fitting, rigid motion estimation). Vice versa, many matching methods (sparse/dense stereo) often use some pre-estimated structural constraints, e.g. epipolar geometry, to identify correct matches/inliers. This paper introduces a novel framework for simultaneous estimation of high-level structures (multi-model fitting) and low-level correspondences (feature matching). We discuss a regularized formulation of the proposed *fit & match* (FM) problem. That formulation uses a generalization of the *assignment problem* and we solve it using an efficient specialized *min-cost-max-flow* solver. This paper primarily focuses on jointly solving multi-homography fitting and sparse feature matching as a simple show case for the FM paradigm. Other applications would be rigid motions estimation, camera pose estimation, etc.

Related Work: An attempt to formulate an objective function for *fitting-&-matching* naturally leads to a version of the assignment problem. The majority of prior work could be divided into two major groups: matching techniques using quadratic assignment problems and FM techniques using linear assignment as sub-problems.

Quadratic assignment problem (QAP) normally appears in the context of non-parametric matching. For example, the methods in [16, 3, 12] estimate non-rigid motion correspondences as a sparse vector field. They rely on a quadratic term in the objective function to encourage geometric regularity between identified matched pairs. Such QAP formulations often appear in shape matching and object recognition. QAP is NP-hard and these methods use different techniques to approximate it. For example, [7] approximates QAP by iteratively minimizing its first-order Taylor expansion, which reduces to a *linear assignment problem*.

If correspondences are constrained by some parametric model(s), matching often simplifies to *linear assignment problem* when model parameters are fixed. In this case, the geometric regularity is enforced by a model fidelity term (linear w.r.t. matching variables) and pair-wise consistencies [16, 3, 12] are no longer needed. Typically for FM problems, feature matching as a *linear assign-*

ment problem and model parameter fitting are preformed in a coordinate descent fashion. For example, SoftPOSIT [5] matches 2D image features to 3D object points and estimate camera pose in such iterative fashion. Building on these ideas [15] fit a single homography using geometric and appearance priors with unknown correspondences.

Our work develops a generalization of *linear assignment problem* for solving FM problem when matching is constrained by an unknown number of geometric models. In contrast to [5, 15], we do not assume that matches/correspondences are constrained by a single parametric model. Note that in order to solve FM problem for multi-models, a regularization term is required to avoid over fitting. Unlike [15, 5, 16], our energy formulation includes label cost regularization as in [6].

Another related approach, *guided matching*, is a post-processing heuristic for increasing the number of matches in case of single model fitting [9]. Similar to our approach, guided matching iteratively re-estimates matches and refines the model. In contrast to our approach, guided matching pursues different objectives at refitting and re-matching steps[1] and does not guarantee convergence. Our method could be seen as an energy-based guided matching with guaranteed convergence. Moreover, unlike guided matching [9], our regularization approach is designed for significantly harder problems where data supports multiple models.

Contributions: We propose a *fit-&-match* energy functional (1)-(2) for jointly solving the matching and multi-model fitting problems. Our energy consists of *unary potentials* term that describes geometric model fitting errors and feature appearance matching costs, and *label cost* term that discourages over-fitting by penalizing the number of models/labels assigned to matched features.

Our *fit-&-match* framework is based on a novel generalization of *linear assignment problem* to multi-model case, namely *generalized assignment problem* (GAP), which jointly formulates feature-to-feature matching and match-to-model assignment. We propose a fast approach for solving the

regularized GAP (in which the number of models are penalized) by generalizing *min-cost-max-flow* techniques for bipartite weighted matching [8]. Our main technical contributions are summarized here:

- *fit-&-match* energy formulation (1)-(2)
- generalized assignment problem, GAP
- a fast solver for regularized GAP.

We compare our joint *fitting-&-matching* framework with a state-of-the-art multi-model fitting algorithm that uses pre-matched features [6] and a variant of it that uses guided matching [9]. Our approach finds more matches, estimates models more accurately, and is more robust to larger baselines between cameras.

2. *Fit-&-Match* Energy

We will use the following notations. $\mathcal{F}_l$ and $\mathcal{F}_r$ are the sets of observed SIFT features [13] in the left and right images, respectively. $\mathcal{L}$ is the set of indices to all models (labels). θ is the set of all models' parameters, $\theta = \{\theta_h | h \in \mathcal{L}\}$ where θ_h is the parameters of model $h \in \mathcal{L}$. In practice, θ could be a set of randomly sampled models, e.g. homographies. f is a *labelling* of all features in the left image, $f = \{f_p | p \in \mathcal{F}_l\}$ where f_p is the label assigned to feature p such that $f_p \in \mathcal{L}$. x_{pq} is a binary variable which is 1 if p and q are matched and 0 otherwise. A matching $\mathcal{M}$ is $\{x_{pq} \mid (p, q) \in \mathcal{F}_l \times \mathcal{F}_r\}$. $Q(p, q)$ is an appearance penalty for matching features $p \in \mathcal{F}_l$ and $q \in \mathcal{F}_r$ based on the similarity of their descriptors.

We define fitting and matching score between features $p \in \mathcal{F}_l$ and $q \in \mathcal{F}_r$ for a given model θ_h as

$$D_{pq}(\theta_h) = ||\theta_h \cdot p - q|| + Q(p, q)$$

combining geometric error and appearance penalty where $|| \ ||$ denotes geometric error e.g. symmetric transfer error. A similar matching score was used in computing the ground truth matching in [14]. We use symmetric appearance penalty[2] $Q(p, q)$, e.g. the angle between the features' descriptors of p and q.

---

[1]Geometric errors minimization vs. inliers maximization.

[2]In this work, $Q(p, q) = 0$ if the angle between two features' descriptors is less than $\pi/4$ and ∞ otherwise.

For simplification, we will introduce our energy functional under the assumption that there are no occlusions/outliers

$$E(f, \theta, \mathcal{M}) = \sum_{p \in \mathcal{F}_l, \, q \in \mathcal{F}_r} D_{pq}(\theta_{f_p}) \cdot x_{pq} + \beta \sum_{h \in \mathcal{L}} \delta_h(f) \quad (1)$$

$$\left. \begin{array}{ll} \sum_{p \in \mathcal{F}_l} x_{pq} = 1 & \forall q \in \mathcal{F}_r \\ \sum_{q \in \mathcal{F}_r} x_{pq} = 1 & \forall p \in \mathcal{F}_l \\ x_{pq} \in \{0, 1\} & \forall p \in \mathcal{F}_l, \, \forall q \in \mathcal{F}_r \end{array} \right\} \quad (2)$$

where $\delta_h(f) = [\exists p \in \mathcal{F}_l : f_p = h]$ and $[\cdot]$ are *Iverson brackets*.

Occlusions/Outliers: Due to occlusions $|\mathcal{F}_l| \neq |\mathcal{F}_r|$ and that renders (1)-(2) unfeasible since the one-to-one constraints could never be met. We add $||\mathcal{F}_l| - |\mathcal{F}_r||$ dummy features, with a fixed matching cost T, to the smaller set of features to ensure feasibility. This is equivalent to changing a rectangular assignment problem to a square one. Also, to make our approach robust to outliers we introduce an outlier model ϕ such that $D_{pq}(\phi) = T$ for any $p \in \mathcal{F}_l$ and $q \in \mathcal{F}_r$. The use of an outlier model with a uniformly distributed cost T is a common technique in Computer Vision [6, 10].

3. Overview of Our Approach

Energy (1)-(2) is NP-hard. We propose to find an approximate solution by minimizing this energy in a block coordinate descent fashion.

In general, to minimize a function via block coordinate descent its coordinates (variables) are partitioned into n blocks, not necessarily mutually exclusive. At each iteration the function is sequentially minimized with respect to the coordinates in each block while fixing other coordinates that are not in this block. Intuitively, fixing different sets of variables in (1)-(2) reduces the energy to special cases which are *easier* to minimize.

In our case, the set of all variables $\{f, \theta, \mathcal{M}\}$ for energy (1) are partitioned into two blocks $\{f, \theta\}$ and $\{f, \mathcal{M}\}$. Note that *labelling f* is in both blocks and, therefore, it is optimized at all steps.

The first block $\{f, \theta\}$ fixes matching $\mathcal{M}$ and en-

ergy (1)-(2) reduces to

$$E(f, \theta) = \sum_{p \in \mathcal{F}_l} D_p(\theta_{f_p}) + \beta \sum_{h \in \mathcal{L}} \delta_h(f) \quad (3)$$

where $D_p(\theta_h) = D_{pq}(\theta_h)$ for all $h \in \mathcal{L}$ provided that $\mathcal{M}$ assigns q to p, i.e. $x_{pq} = 1$. Energy (3) could be efficiently minimized over f and θ using standard multi-model fitting methods for fixed matching, e.g. PEARL [6].

The second block $\{f, \mathcal{M}\}$ fixes parameters θ. We separately consider two cases: $\beta = 0$ and $\beta > 0$. In case $\beta = 0$ the optimization problem could be optimally solved in polynomial time and the corresponding algorithm is used as a building block for the more general case $\beta > 0$. Thus, we first discuss the simpler case $\beta = 0$ when energy (1)-(2) does not penalize the number of models. It reduces to

$$E(f, \mathcal{M}) = \sum_{p \in \mathcal{F}_l, \, q \in \mathcal{F}_r} D_{pq}(\theta_{f_p}) \cdot x_{pq} \quad (4)$$

subject to constraints (2). We will refer to (4)-(2) as the *generalized assignment problem* (GAP). GAP is a weighted matching problem over a fixed set of models that match features and assigns each match to a model[3]. Furthermore, GAP is an *integral linear program*, see our proof in [11]. Section 4.1 describes a solver for GAP that finds its global minimum in polynomial time.

The more general case $\beta > 0$ of the block coordinate descent for (1)-(2) with respect to variables $\{f, \mathcal{M}\}$ reduces to optimization of energy

$$E(f, \mathcal{M}) = \sum_{p \in \mathcal{F}_l, \, q \in \mathcal{F}_r} D_{pq}(\theta_{f_p}) \cdot x_{pq} + \beta \sum_{h \in \mathcal{L}} \delta_h(f) \quad (5)$$

subject to constraints (2), which is NP-hard. We will refer to (5)-(2) as *regularized*-GAP. Section 4.2 introduces *Local Search-GAP* (LS-GAP) approximation algorithm for energy (5)-(2). It uses our GAP solver in a combinatorial local search fashion iteratively exploring different subsets of models and selecting solutions reducing energy (5) which requires solving a series of similar GAP instances efficiently.

---

[3]Our definition of GAP is different from some generalizations of the assignment problem in the optimization literature.

Our **E**nergy-based **F**itting & **M**atching (EFM) algorithm for energy (1)-(2) can be summarized as

---

Energy-based Fitting & Matching (EFM)

Initialization: Find an initial matching $\mathcal{M}$ using standard matching techniques
repeat
  1-Given matching $\mathcal{M}$, minimize (3) using PEARL [6] to find *labelling* f and models' parameters θ.
  2-Given models' parameters θ, minimize (5)-(2) using LS-GAP, see Sec. 4.2, to find matching $\mathcal{M}$ and *labelling* f.
until energy (1) converges

EFM finds an initial matching using standard matching techniques, e.g. standard SIFT matching [13]. Then, it iteratively optimizes energy (1)-(2) by alternatively minimizing energy (3) over *labelling* f and models' parameters θ for fixed matching $\mathcal{M}$, and minimizing energy (5)-(2) over f and $\mathcal{M}$ for fixed θ. Although EFM is guaranteed to converge since energy (1) is bounded below, i.e. (1) $\geq \beta$, it is not trivial to derive a theoretical bound on the convergence rate and approximation ratio for EFM. However, in Section 5, we empirically show that EFM converges in a few iterations to a near optimal solution.

4. Optimization

Section 4.1 shows how to optimally solve GAP (4)-(2) and a series of similar GAP instances efficiently. Section 4.2 covers Local Search-GAP (LS-GAP) algorithm for energy (5)-(2) that requires solving many similar GAP instances.

4.1. Solving GAP

There are alternative ways to solve GAP [11]. This section describes an approach that we find most efficient. We reduce GAP to standard *Linear Assignment Problem* (LAP) and propose an efficient solver for sequences of similar problems.

Reducing GAP to LAP: GAP (4)-(2) reduces to LAP since f and $\mathcal{M}$ are independent: any pair (p, q) has optimal label $f_p = \underset{h \in \mathcal{L}}{arg\,min}\, D_{pq}(\theta_h)$ independently from the value of x_{pq}. Then, optimal

$\mathcal{M}$ in (4)-(2) is found by solving the following LAP

$$E(\mathcal{M}) = \sum_{p \in \mathcal{F}_l, \, q \in \mathcal{F}_r} D_{pq} \cdot x_{pq} \qquad (6)$$

subject to (2) where $D_{pq} := \underset{h \in \mathcal{L}}{min}\, D_{pq}(\theta_h)$.

LAP as MCMF (overview): LAP (6)-(2) can be equivalently formulated as a standard *min-cost-max-flow* (MCMF) problem with known efficient solvers [1]. This problem is defined as follows. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a graph with vertices $\mathcal{V}$ and edges $\mathcal{E}$ where each edge $(v, w) \in \mathcal{E}$ has a capacity $u(v, w)$ and cost $c(v, w)$. Let F be a flow function such that $0 \leq F(v, w) \leq u(v, w)$ for over all edges in $\mathcal{E}$. The cost of an arbitrary flow function F is defined as $cost(F) = \sum_{(v,w) \in \mathcal{E}} c(v, w) \cdot F(v, w)$. MCMF is a valid maximum flow F from s to t in $\mathcal{V}$ that has minimum cost.

To formulate LAP (6)-(2) as MCMF problem we build graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with nodes

$$\mathcal{V} = \{s, t\} \cup \{p \mid p \in \mathcal{F}_l\} \cup \{q \mid q \in \mathcal{F}_r\},$$

edges

$$\mathcal{E} = \{(s, p), (q, t), (p, q) | p \in \mathcal{F}_l, q \in \mathcal{F}_r\},$$

capacity $u(v, w) = 1$ for all edges $(v, w) \in \mathcal{E}$, and cost $c(p, q) = D_{pq}$ for edges $(p, q) \in \mathcal{F}_l \times \mathcal{F}_r$ and 0 for other edges. The optimal $\mathcal{M}$ and f for GAP can be obtained from MCMF flow F^* for $\mathcal{G}$ as $x_{pq} = F^*(p, q)$ for all $(p, q) \in \mathcal{F}_l \times \mathcal{F}_r$ and $f_p = \underset{h \in \mathcal{L}}{arg\,min}\, D_{pq}(\theta_h)$ if p, q are matched, $x_{pq} = 1$.

Solving MCMF (overview): There are many algorithms for finding MCMF for a given graph [1]. We overview the *Successive Shortest Path* (SSP) algorithm [1] in order to introduce our *flow recycling* technique for efficiently solving similar GAP instances. SSP successively finds the shortest path w.r.t. edge costs from s to t and augments these paths until the network is saturated. For unit capacity graphs, augmentation of an edge reverses its direction and flips its cost sign. Finding the shortest path with negative costs is expensive. Instead of the original costs SSP uses *reduced costs* $c^\pi(v, w) := c(v, w) - \pi(v) + \pi(w) \geq 0$ where $\pi(v)$

is the *potential* of node v. Initially set to zero, node potentials are updated after each path augmentation to ensure that the reduced costs non-negativity constraints are satisfied, see [1] for more details. Let $n = |\mathcal{F}_r| = |\mathcal{F}_l|$ be the number of features. A shortest path w.r.t. c^π could be found in $O(n^2)$ using Dijkstra's algorithm. By construction, we need to find n paths. Thus, SSP is $O(n^3)$ when solving LAP.

Solving a Series of GAPs: We propose $O(n^2)$ method for solving MCMF corresponding to a modified LAP (6)-(2) after changing one or all edge costs associated with one feature in $\mathcal{F}_l$. Assume MCMF F for $\mathcal{G}$ and node potential function π that satisfy the reduced costs non-negativity constraints on the residual graph $\mathcal{G}_F$. Changing edge costs associated with feature p may violate reduced cost non-negativity constraints involving p. To regain feasibility after dropping the no longer needed artificial nodes s and t and their edges, we reverse the flow through (p, q) where p and q are matched by F and update $\pi(p)$

$$\pi(p) = \min c(p, v) + \pi(v) \quad \forall v \in \mathcal{F}_r.$$

Finally, we push one unit of flow from p to q, i.e. find the shortest path w.r.t. c^π, to maximize the flow. The reduced cost optimally theorem [1] grantees that the resulting flow is MCMF. In case m features in $\mathcal{F}_l$ had their associated costs changed, the new MCMF could be found in $O(mn^2)$ by applying the steps above sequentially to each feature. These steps could be used with any LAP [4] or MCMF solver not just SSP. Given an optimal solution for LAP (6)-(2), it is possible to compute optimal node potentials that satisfy reduced cost non-negativity constraints in polynomial time [1, 11].

4.2. Local Search-GAP (LS-GAP)

Now we introduce a local search algorithm that solves regularized GAP (5)-(2) using GAP algorithm introduced in Section 4.1 as a sub-procedure. Assume that $\mathcal{L}$ is the current set of possible models[4]. Let $\mathcal{L}_c$ be an arbitrary subset of $\mathcal{L}$ and $\mathcal{M}_f(\mathcal{L}_c)$ denote the GAP solution when the label

---

[4]In practice, we restrict $\mathcal{L}$ to be the set of models that were assigned to at least one matched pair of features in (3) solution.

space is restricted to $\mathcal{L}_c$. Note that GAP ignores the label cost term in (5) but we could easily evaluate energy (5) for $\mathcal{M}_f(\mathcal{L}_c)$. The proposed LS-GAP algorithm greedily searches over different subsets $\mathcal{L}_c \in \mathcal{L}$ for one such that $\mathcal{M}_f(\mathcal{L}_c)$ has the lowest value of energy (5). Our motivation to search for minima of (5)-(2) only among GAP solutions comes from an obvious observation that a global minima of (5)-(2) must also solve the GAP if the label space is restricted to the right subset of $\mathcal{L}$.

We define sets of all possible *add*, *delete* and *swap* combinatorial search moves as

$$
\begin{aligned}
\mathcal{N}^a(\mathcal{L}_c) &= \cup_{h \in \mathcal{L} \setminus \mathcal{L}_c} \{\mathcal{L}_c \cup h\} \\
\mathcal{N}^d(\mathcal{L}_c) &= \cup_{h \in \mathcal{L}_c} \{\mathcal{L}_c \setminus h\} \\
\mathcal{N}^s(\mathcal{L}_c) &= \cup_{\substack{h \in \mathcal{L}_c \\ \ell \in \mathcal{L} \setminus \mathcal{L}_c}} \{\mathcal{L}_c \cup \ell \setminus h\}.
\end{aligned}
$$

These are three different local neighbourhoods around $\mathcal{L}_c$. We also define a larger neighbourhood $\mathcal{N}^\star$ around $\mathcal{L}_c$ as the union of $\mathcal{N}^a(\mathcal{L}_c)$, $\mathcal{N}^d(\mathcal{L}_c)$ and $\mathcal{N}^s(\mathcal{L}_c)$. LS-GAP uses a combination of *add*, *delete* and *swap* moves, similar to the work in [2], to greedily find a set of labels near current set $\mathcal{L}_t$ that is better w.r.t. energy (5).

Local Search-GAP (LS-GAP)

$\quad \mathcal{L}_t \leftarrow \phi, \ \mathcal{N}_t \leftarrow \mathcal{N}^\star(\mathcal{L}_t)$
while $\exists \, \mathcal{L}_c \in \mathcal{N}_t$
$\quad$ **if** energy (5) of $\mathcal{M}_f(\mathcal{L}_c) <$
$\quad\quad\quad$ of energy (5) $\mathcal{M}_f(\mathcal{L}_t)$
$\quad\quad \mathcal{L}_t \leftarrow \mathcal{L}_c, \ \mathcal{N}_t \leftarrow \mathcal{N}^\star(\mathcal{L}_t)$
$\quad$ **else**
$\quad\quad \mathcal{N}_t \leftarrow \mathcal{N}_t \setminus \mathcal{L}_c$
return GAP solution $\mathcal{M}_f(\mathcal{L}_t)$

---

5. Evaluation

In this section, we discuss some of the EFM framework properties, e.g. convergence rate etc. Then we compare the matching quality of our proposed EFM framework to standard SIFT matching [13]. We also compare the matching quality and the accuracy of models estimated by the EFM framework, **E**nergy-based multi-model **F**itting (EF) algorithm [6, 10], and EF followed by guided matching [9] (EF+GM). Carrying out these experiments

requires knowing the ground truth of the dataset at hand. We computed ground truth; matching $\mathcal{M}_{GT}$, model estimates θ_{GT} and *labelling* f_{GT} as described in [11].

The effect of EFM iterations on energy (1) for different $|\mathcal{L}|$ is shown in Fig. 1(a). For each $|\mathcal{L}|$ the experiment is repeated 50 times. On the average each iteration took 1 min., and most of the energy was reduced in the first three iterations. EFM converged on the average after 5 iterations.

EFM is non-deterministic as it uses a set of randomly sampled models $\mathcal{L}$. Figure 1(b) shows final energy histograms to different sizes of $\mathcal{L}$. As shown the bigger $|\mathcal{L}|$ is the more likely the final energy is going to be small, i.e. better solutions. Using a large $|\mathcal{L}|$ helps EFM avoid local minima.

Our matching evaluation criterion is based on Receiver Operating Characteristics (ROC) of the True Positives Rate vs. the False Positives Rate. The ROC attributes for an estimated matching $\mathcal{M}$ and ground truth matching $\mathcal{M}_{GT}$ are defined as follows: *Positives* (P) number of matches in $\mathcal{M}_{GT}$, *Negatives* (N) number of potential matches that were rejected by $\mathcal{M}_{GT}$, *True Positives* (TP) number of matches in $\mathcal{M}$ and $\mathcal{M}_{GT}$ (intersection), *False Positives* (FP) number of matches in $\mathcal{M}$ but not in $\mathcal{M}_{GT}$, *True Positives Rate* (TPR) $\frac{TP}{P}$, and *False Positives Rate* (FPR) $\frac{FP}{N}$.

A basic comparison between the matching quality of EFM and standard SIFT matching is shown in Fig. 2. The ROC curve, in Fig. 2, of SIFT matching [13] is achieved by varying the threshold on the second best ratio (SBR) [5]. For EFM we show a scatter plot since it is non-deterministic. We also related the EFM scatter plot to the achieved final energy by colour coding it. As can be seen, for EFM the lower the final energy (blue is low energy) the better the matching quality. Also, EFM outperformed SIFT matching by reaching high TPR values.

---

[5]SBR is the ratio of the distance between a left feature descriptor and its closest right feature descriptor to the distance of its second closest

(a) Energy vs Time

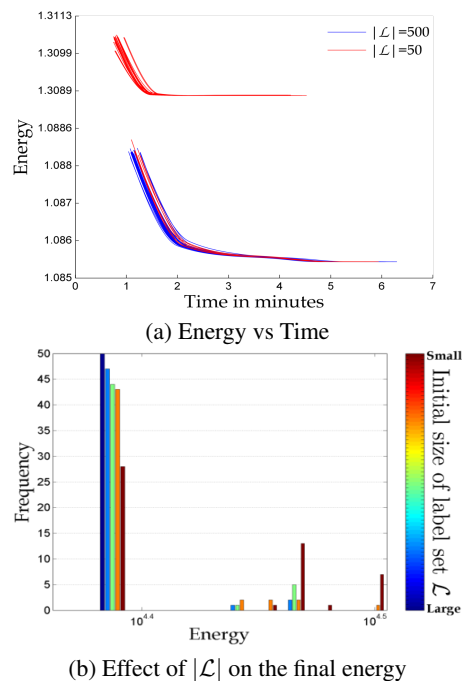(b) Effect of $|\mathcal{L}|$ on the final energy

Figure 1: Best seen in Colour, Fig. (a) shows the effect of EFM iterations on energy (1). EFM converged on the average after 5 iterations, and each iteration on the average took 1 min. Figure (b) shows multiple histograms of the final energies for different sizes of initial set of proposals—blue indicates a large set of proposals. The larger the set of initial proposals $\mathcal{L}$ the more likely that EFM will converge to a low energy.
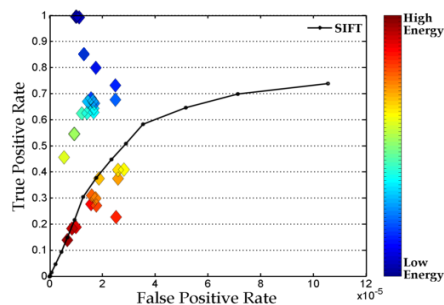

Figure 2: Best seen in Colour, Fig. (a) shows ROC curve for standard SIFT matches by varying SBR threshold, and the scatter plot represents EFM results for different sizes of initial set or proposals. As can be seen, the lower the final energy (blue) the better the matching.

		GQ			ROC	
		med.	mean	var.	TPR	FPR
small baseline	EFM	**1.01**	1.01	4E-6	**0.98**	**2E-6**
	EF	1.04	1.05	1E-3	0.78	5E-6
	EF + GM	1.01	1.02	3E-4	0.96	2E-5
medium baseline	EFM	**1.02**	1.02	1E-6	**0.97**	**3E-6**
	EF	1.20	1.30	9E-2	0.33	4E-6
	EF + GM	1.07	1.12	2E-2	0.94	4E-5
large baseline	EFM	**1.05**	1.07	2E-3	**0.96**	2E-6
	EF	1.90	2.24	1E+0	0.06	**6E-7**
	EF + GM	1.49	1.78	9E-1	0.90	7E-5

Table 1: Graphite VGG Oxford dataset, single model and increasing baseline. The table shows the averages of *GQ* and ROC attributes, over 50 runs, for EFM, EF, and EF+GM model estimates.

The plots in Fig. 1 and 2 are shown for Oxford's Merton College example of Fig. 3(b), to illustrate the characteristics/behaviour of the EFM algorithm. Note that it will be meaningless to average these plots over different examples as they would not share the same energy scale.

For measuring the accuracy of an estimated model θ_h, we used the following geometric error ratio $GQ(\theta_h) = \frac{STE(\theta_h, f_{GT}, \mathcal{M}_{GT})}{STE(\theta_{GT}, f_{GT}, \mathcal{M}_{GT})}$ where $STE(\theta_h, f, \mathcal{M})$ is the Symmetric Transfer Error of θ_h computed for labelling f and matching $\mathcal{M}$. A perfect model estimate will have $GQ = 1$.

Table 1 shows the effect of increasing the cameras' baseline on the quality of estimated models and matching for EFM, EF and EF+GM methods. For small baseline, EFM and EF+GM results were comparable. For larger baselines, unlike the EF and EF+GM, the EFM model estimates' accuracy and matching quality did not deteriorate. In general, EF+GM was prone to higher false positive rates compared to EFM.

As a multi-model show case example, Fig. 3(a) and (b) show the *labelling f* result of EF and EFM on one of the stereo images, respectively. EFM, on the average of 50 runs, found double the number of matches compared to EF which takes SIFT matching as an input. Figures 3(c-f) are the enlargements of the segments shown in (a) and (b) as white rectangles. Figures 3(g) and (h) show part of the feature

(a) EF left image result

(b) EFM left image result

(c) Segment 1 in (a)

(d) Segment 1 in (b)

(e) Segment 2 in (a)

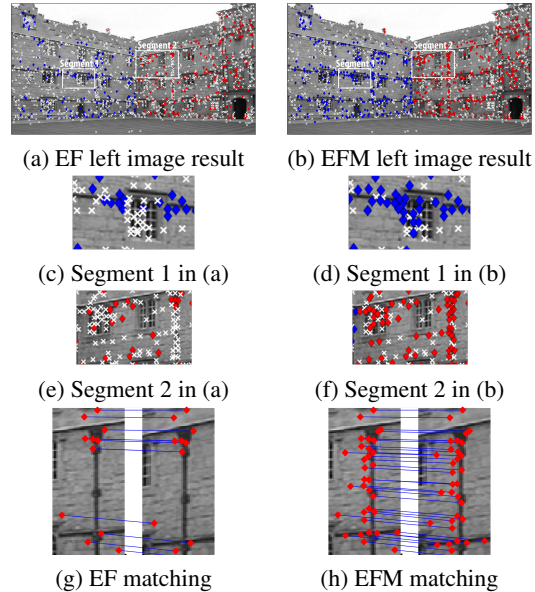(f) Segment 2 in (b)

(g) EF matching

(h) EFM matching

Figure 3: Best seen in Colour, Fig. (a) shows EF *labelling* result (average TPR=0.51 and FPR=1.6E-05) and (b) shows EFM *labelling* result (average TPR=0.98 and FPR=9.1E-06). Features assigned to the same model/label are drawn in the same colour and unmatched features are shown as white x. Figures (c-f) show the enlargement of segments 1 and 2 in (a) and (b). Figures (g-h) show the matching, between two small regions in the stereo images, of the EF and EFM results, respectively.

matching between the left and right images of the EF and EFM results, respectively.

More results are shown in Fig. 4. In general, EFM found more matches, but in particular, EFM outperformed EF in two examples: graphite example (second row) with a the large baseline between camera positions, and redbrick house example (third row) with repetitive texture reduced the discriminative power of SIFT. EFM found approximately 5% to 8% more matches than EF+GM. The EF+GM results are not shown as they where visually similar to EFM.

6. Conclusions

We introduced regularized energy functional that jointly formulates multi-model fitting and
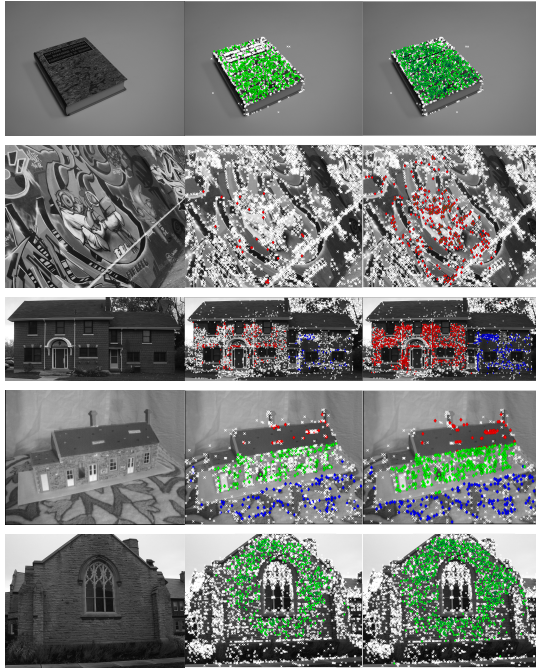
Figure 4: Best in Colour, the first column shows one of the stereo images for each example, second and third columns show the EF and EFM *labelling* results superimposed on the images shown in the first column, respectively. On average of 50 runs, EFM found 0.75, 10.53, 3.31, 0.44, and 0.68 times more inliers than EF. EFM and EF+GM results where comparable, EFM found approximately between 0.05 and 0.08 times more matches.

matching problems, and a framework to optimize it. Our results show that the framework finds near optimal matching, and when compared to state-of-the-art multi-model fitting algorithm our framework finds better models' estimates and more robust to large baselines. We also showed how to efficiently find optimal feature-to-feature matching and match-to-model assignment for a given set of models with label cost. Furthermore, our framework can be used with more complex models, e.g. fundamental matrices, without affecting the framework's complexity, unlike [15]. Currently, our framework requires initial matching for future work we aim to alleviate the need for it.

References

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, Inc., 1993.

[2] V. Arya, N. Garg, R. Khandekar, A. Meyerson, K. Munagala, and V. Pandit. Local search heuristic for k-median and facility location problems. In *Theory of computing*, pages 21–29. ACM, 2001.

[3] A. Berg, T. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondences. In *CVPR*, pages 26–33, 2005.

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

[5] P. David, D. Dementhon, R. Duraiswami, and H. Samet. Softposit: Simultaneous pose and correspondence determination. *IJCV*, 59(3):259–284, 2004.

[6] A. Delong, A. Osokin, H. Isack, and Y. Boykov. Fast Approximate Energy Minization with Label Costs. *IJCV*, 96(1):1–27, 2012.

[7] S. Gold and A. Rangarajan. A graduated assignment algorithm for graph matching. *PAMI*, 18(4):377–388, 1996.

[8] A. V. Goldberg. An efficient implementation of a scaling minimum-cost flow algorithm. *Journal of Algorithms*, 22:1–29, 1992.

[9] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[10] H. Isack and Y. Boykov. Energy-based geometric multi-model fitting. *IJCV*, 97(2), April 2012.

[11] H. Isack and Y. Boykov. Joint optimization of fitting & matching in multi-view reconstruction. *arXiv:1303.2607v2*, April 2014.

[12] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. In *ICCV*, pages 1482–1489, 2005.

[13] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 2004.

[14] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. In *CVPR*, volume 2, pages 257–263, 2003.

[15] E. Serradell, M. Özuysal, V. Lepetit, P. Fua, and F. Moreno-Noguer. Combining geometric and appearance priors for robust homography estimation. *ECCV*, pages 58–72, 2010.

[16] L. Torresani, V. Kolmogorov, and C. Rother. A dual decomposition approach to feature correspondence. *PAMI*, 99(PrePrint), 2012.