

# Dynamic and Hierarchical Multi-Structure Geometric Model Fitting

Hoi Sim Wong, Tat-Jun Chin, Jin Yu and David Suter

The Australian Centre for Visual Technologies

School of Computer Science, The University of Adelaide, South Australia

{hwong, tjchin, jin.yu, dsuter}@cs.adelaide.edu.au

## Abstract

*The ability to generate good model hypotheses is instrumental to accurate and robust geometric model fitting. We present a novel dynamic hypothesis generation algorithm for robust fitting of multiple structures. Underpinning our method is a fast guided sampling scheme enabled by analysing correlation of preferences induced by data and hypothesis residuals. Our method progressively accumulates evidence in the search space, and uses the information to dynamically (1) identify outliers, (2) filter unpromising hypotheses, and (3) bias the sampling for active discovery of multiple structures in the data—All achieved without sacrificing the speed associated with sampling-based methods.*

*Our algorithm yields a disproportionately higher number of good hypotheses among the sampling outcomes, i.e., most hypotheses correspond to the genuine structures in the data. This directly supports a novel hierarchical model fitting algorithm that elicits the underlying stratified manner in which the structures are organized, allowing more meaningful results than traditional “flat” multi-structure fitting.*

## 1. Introduction

Hypothesis sampling is at the core of state-of-the-art robust geometric model fitting. The popular hypothesize-and-test strategy [7] depends on the ability to retrieve at least one hypothesis close to the true underlying instance (or structure) of the geometric model in the data. Mode seeking methods [5] require the presence of a large number of good hypotheses (relative to bad or spurious ones) that form dense clusters in the parameter space, where each cluster is indicative of the existence of a structure in the data.

An effective hypothesis sampling scheme must not only retrieve good hypotheses rapidly, but must produce them in sufficient quantities. Simply retrieving a single all-inlier minimal subset is no guarantee of success, as even different *all-inlier* minimal subsets (and the hypotheses they yield) vary greatly in quality [14]. To this end, much progress has

been made in the single structure case [4, 10, 18, 3, 8, 1, 13]. Most notable are methods that use keypoint matching scores to accelerate the search for all-inlier minimal subsets for the task of fitting various projective transformations.

Some efforts [2, 19] focus on guided sampling for multi-structure data. The presence of multiple structures requires a conditional sampling scheme to discourage inliers from different valid structures (and outliers) to be included in the same minimal subsets. This is often accomplished by building and propagating conditional inlier probabilities on the fly. As a result, multi-structure sampling methods are hampered by a relatively expensive sampling algorithm. Moreover, there is a lack of consensus on how to best exploit the resulting set of generated hypotheses.

In this paper we introduce a novel *dynamic* sampling scheme for multi-structure data. In contrast to [2, 19] which devote considerable computation on building and updating conditional inlier probabilities, we show that a simpler and cheaper sampling scheme based on preference analysis is possible without impacting sampling accuracy. The significant gain in efficiency permits an online routine to dynamically identify outliers and filter unpromising hypotheses which further refine the sampling. Fig. 1 provides an overview of our approach.

Unlike previous multi-structure sampling methods, our algorithm also consists of a step that actively explores the space of model parameters with the aim of reinforcing relatively unknown structures. This prevents a monopoly of the sampling effort by the larger structures in the data, and allows the smaller structures to be well represented in the total generated set of hypotheses. This aspect of our work differs from *single* structure sampling schemes [8, 1] that alternate between local refinement of promising hypotheses or global exploration to escape local minima.

The proposed sampling algorithm returns a set of (filtered) hypotheses consisting overwhelmingly of good parameters corresponding to the true structures in the data. Based on such results, we propose a *hierarchical* multi-structure model fitting approach based on top- $k$  preference analysis [6]. In our method, each remaining hypothesis is

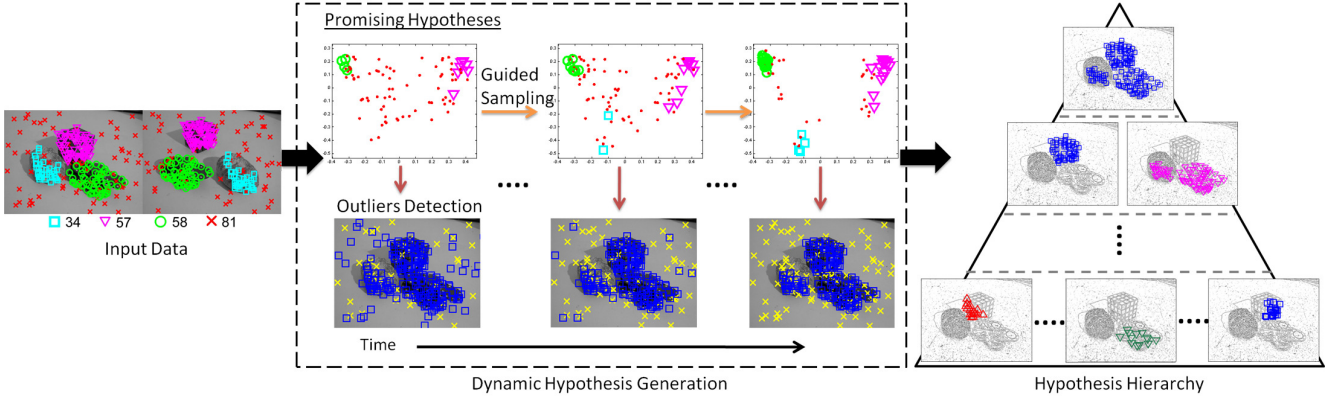


Figure 1. Overview of our proposed method on fundamental matrix estimation (best viewed in color). Input data: Red crosses indicate the gross outliers, and other colored markers indicate the inliers of the 3 structures in data. *Dynamic Hypothesis Generation*: Our method represents each hypothesis with a preference (permutation). To aid for presentation, we show the hypotheses in the preference space by projecting the preferences into a two dimensional space using Multi-Dimensional Scaling [11]. Red dots indicate the hypotheses generated from impure inliers samples, and other colored markers indicate the hypotheses generated from all-inlier samples of a structure marked by the same markers in input data. As sampling proceeds, our method uses the promising hypotheses after filtering (top row) to guide the sampling and identify gross outliers (marked by yellow crosses in bottom row). *Hypothesis Hierarchy*: we use the promising hypotheses after filtering and identified inliers to construct a hypothesis hierarchy, which allows multiple levels of details to be captured. For illustration purposes, we show the model fitting results in the hierarchy using the edge image throughout this paper.

represented as a permutation indicating the preference of the hypothesis to the data points. The implicit scale parameter underlying the hierarchical clustering of structures is directly related to the number  $k$  of the top preferences of hypotheses. Our formulation is more powerful than mean shift clustering [5] since it does not require an underlying Euclidean structure in the parameter space, thus obviating the need to adapt to more exotic geometric models [15, 16]. Fig. 1 illustrates an example result.

The rest of the paper is organized as follows: Sec. 2 describes our proposed dynamic hypothesis generation algorithm. Sec. 3 describes how the hypothesis hierarchy is constructed by top- $k$  preference of hypotheses. Sec. 4 presents the experimental result on real data and we finally draw conclusions in Sec. 5.

## 2. Dynamic Hypothesis Generation

This section describes the proposed dynamic sampling methods based on the preference analysis. By accumulating the information during sampling, we use such information to dynamically (1) filter the unpromising hypotheses, (2) guide the sampling for discovery of multiple structures in data, and (3) identify outliers. Once the sampling is done, a set of promising hypotheses and identified inliers are immediately available. The algorithm of our method is summarized in Alg. 1.

### 2.1. Hypothesis Goodness from Preference

We measure the distance between two hypotheses by comparing their preferences to data. We further analyse the

---

### Algorithm 1 Dynamic Hypothesis Generation

---

- 1: **Input** input data  $X$ , batch size  $b$ , Time budget  $T$ , size of a minimal subset  $p$
  - 2: **Output** a set  $\bar{\Theta}$  of promising hypotheses and a set  $\bar{X}$  of identified inliers
  - 3: **Let**  $\bar{\Theta}$  and  $\bar{X}$  be an empty set
  - 4: **for**  $t = 1, 2, \dots, T$  **do**
  - 5:   **if**  $t \leq b$  **then**
  - 6:     generate a hypothesis  $\theta_t$  by randomly sample  $p$  data
  - 7:   **else**
  - 8:     generate a hypothesis  $\theta_t$  by guided sampling scheme (Sec. 2.3)
  - 9:   **end if**
  - 10:    $\bar{\Theta} = \bar{\Theta} \cup \{\theta_t\}$
  - 11:   **if**  $t \geq b$  and  $\text{mod}(t, b) == 0$  **then**
  - 12:     filter unpromising hypotheses in  $\bar{\Theta}$  (Sec. 2.2)
  - 13:     identify outliers and update  $\bar{X}$  (Sec. 2.4)
  - 14:   **end if**
  - 15: **end for**
- 

distribution of hypotheses in the preference space to derive a hypothesis goodness measure.

Let  $X = \{x_i\}_{i=1}^N$  be a set of  $N$  input data and  $\Theta = \{\theta_j\}_{j=1}^M$  be a set of  $M$  hypotheses, where each hypothesis  $\theta_j$  is fitted from a minimal subset of  $p$  data (e.g.  $p=4$  for homography estimation). For each hypothesis  $\theta_j$ , we compute its absolute residual as measured to  $N$  data, i.e.  $r^{(j)} = [r_1^{(j)}, r_2^{(j)}, \dots, r_N^{(j)}]$ . We sort the elements in  $r^{(j)}$

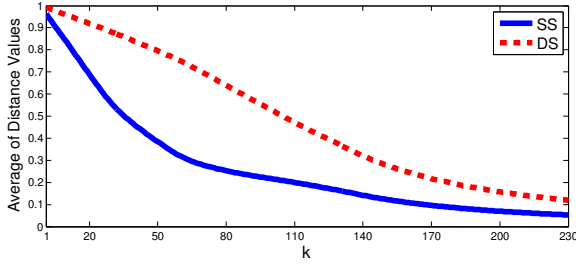


Figure 2. Under various  $k$ , the average of distance values between two hypotheses generated from inliers of the same structure (SS) and two hypotheses generated from different structures (DS) using the input data in Fig. 1.

to obtain a sorted residual list  $\tilde{r}^{(j)} = [r_{\lambda_1^{(j)}}^{(j)}, \dots, r_{\lambda_N^{(j)}}^{(j)}]$  such that  $r_{\lambda_1^{(j)}}^{(j)} \leq \dots \leq r_{\lambda_N^{(j)}}^{(j)}$ . The *top- $k$  data* of a hypothesis  $\theta_j$  is defined as the first  $k$  elements in the permutation  $[\lambda_1^{(j)}, \dots, \lambda_N^{(j)}]$ :

$$\tau_j = [\lambda_1^{(j)}, \dots, \lambda_k^{(j)}]. \quad (1)$$

Essentially,  $\tau_j$  identifies the  $k$  data that are more likely to be inliers to  $\theta_j$ .

We measure the distance between two hypotheses by comparing their preferences to data. Given two hypotheses  $\theta_j$  and  $\theta_{j'}$ , we compare their corresponding top- $k$  data  $\tau_j$  and  $\tau_{j'}$  using the Spearman Footrule distance [6]. Let  $D_{\tau_j}$  be a set of elements contained in  $\tau_j$ , and  $\tau_j(m)$  be the position of the element  $m \in D_{\tau_j}$  in  $\tau_j$ , similarly for  $\tau_{j'}$ . The distance between  $\theta_j$  and  $\theta_{j'}$  is defined as

$$d(\theta_j, \theta_{j'}) = \frac{1}{k \times (k+1)} \sum_{m \in D_{\tau_j} \cup D_{\tau_{j'}}} |\bar{\tau}_j(m) - \bar{\tau}_{j'}(m)|, \quad (2)$$

where  $\bar{\tau}_j(m) = \tau_j(m)$  if  $m \in D_{\tau_j}$ ; otherwise  $\bar{\tau}_j(m) = k+1$ , and  $\bar{\tau}_{j'}$  is similarly obtained from  $\tau_{j'}$ . Note that the distance is normalized such that  $d(\theta_j, \theta_{j'})$  is between 0 (identical) and 1 (completely different). Fig. 2 shows that the distance between two hypotheses generated from inliers of the same structure (SS) is lower than that from different structures (DS) across various values of  $k$ .

Given a set of hypotheses  $\hat{\Theta} \subseteq \Theta$ , we use the following Gaussian kernel to measure how good a hypothesis  $\theta_j$  is, compared to other hypotheses in  $\hat{\Theta}$ :

$$w_j^{(\hat{\Theta})} = \sum_{\theta_{j'} \in \hat{\Theta}} \exp\left(-d(\theta_j, \theta_{j'})^2 / 2\delta^2\right), \quad (3)$$

where  $d(\theta_j, \theta_{j'})$  is measured by Eq. 2 and  $\delta$  is the bandwidth. The higher the value of  $w_j^{(\hat{\Theta})}$ , the more promising the hypothesis  $\theta_j$  is.

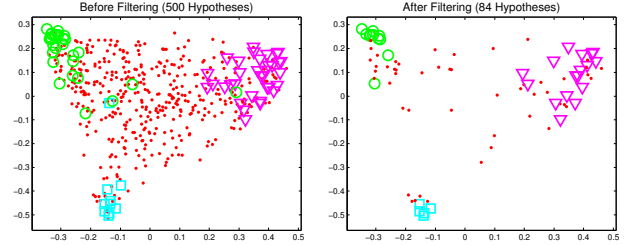


Figure 3. Example of hypothesis filtering using the input data in Fig. 1. To aid for presentation, we show the hypotheses by projecting the preferences into a two dimensional space using Multi-Dimensional Scaling [11]. Red dots indicate the hypotheses generated from impure inliers minimal subset. Other colored markers indicate the hypotheses generated from all-inlier minimal subset of the corresponding structure with the same markers in input data Fig. 1.

## 2.2. Hypothesis Filtering

Using the goodness measured on a set of hypotheses, we seek a promising hypothesis for each datum using its preference to the hypotheses.

Similar to the definition of top- $k$  data of a hypothesis (Eq. 1), we define the preference of data to hypotheses using residual sorting information. For each datum  $x_i$ , we sort its absolute residual as measured to  $M$  hypotheses  $r_i = [r_i^{(1)}, r_i^{(2)}, \dots, r_i^{(M)}]$ , and a sorted residual list  $\tilde{r}_i = [r_i^{(\lambda_i^{(1)})}, \dots, r_i^{(\lambda_i^{(M)})}]$  is obtained such that  $r_i^{(\lambda_i^{(1)})} \leq \dots \leq r_i^{(\lambda_i^{(M)})}$ . The *top- $h$  hypotheses* of a datum  $x_i$  is defined as the first  $h$  elements in the permutation  $[\lambda_i^{(1)}, \dots, \lambda_i^{(M)}]$ :

$$\sigma_i = [\lambda_i^{(1)}, \dots, \lambda_i^{(h)}]. \quad (4)$$

For each datum  $x_i$ , we seek the most promising hypothesis for  $x_i$  within its top- $h$  hypotheses  $\sigma_i$ , *i.e.*

$$e_i = \arg \max_{j \in D_{\sigma_i}} w_j^{(\Theta_{\sigma_i})}, \quad (5)$$

where  $\Theta_{\sigma_i} = \{\theta_j \in \Theta | j \in D_{\sigma_i}\}$  is a set of hypotheses corresponding to  $\sigma_i$ . We call this promising hypothesis  $e_i$  an *exemplar*. There are at most  $N$  unique exemplar associated to data in  $X$ . We prune out the hypotheses that fail to be an exemplar to any data in  $X$  as they are unlikely to be good hypotheses to explain the data. A set  $\bar{\Theta}$  of promising hypotheses collects the unique exemplar of all data:

$$\bar{\Theta} = \{\theta_j \in \Theta | j \in \{e_1, \dots, e_N\}\}. \quad (6)$$

$\bar{\Theta}$  is incrementally maintained as sampling proceeds. Fig. 3 illustrates the effect of hypothesis filtering. As can be seen, a large number of unpromising hypotheses are filtered, and the hypotheses generated from all-inliers samples of each structure are successfully retained after filtering.

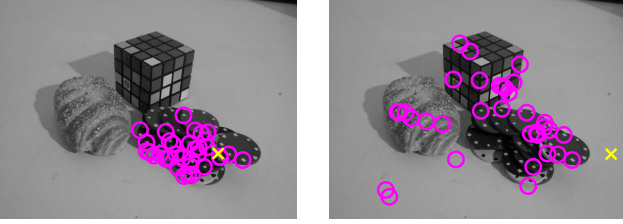


Figure 4. Left: the exemplar is associated to an inlier marked by a yellow cross (Exploitation). Right: the exemplar is associated to a gross outlier marked by a yellow cross (Exploration). Magenta circles indicate the top- $k$  data of an exemplar (best viewed in color).

### 2.3. Guided Sampling

By maintaining a set of promising hypotheses as sampling proceeds, we use the top- $k$  data of the promising hypotheses to guide the sampling.

Given a set  $\bar{\Theta}$  of promising hypotheses after filtering, we sample a hypothesis  $\theta_j$  from  $\bar{\Theta}$  with the following weight:

$$\max_{\theta_{j'} \in \bar{\Theta}} w_{j'}^{(\bar{\Theta})} - w_j^{(\bar{\Theta})}. \quad (7)$$

This ensures that the sampled hypothesis is less promising than other hypotheses in  $\bar{\Theta}$ , hence encouraging the discovery of other possible structures in data. Once a hypothesis  $\theta_j$  is selected from  $\bar{\Theta}$ , we construct a minimal subset by randomly selecting  $p$  data from its top- $k$  data  $\tau_j$ .

Intuitively, our guided sampling scheme is composed of exploration and exploitation steps. Each hypothesis  $\theta_j \in \bar{\Theta}$  is either an exemplar of an inlier or a gross outlier. If  $\theta_j$  is an exemplar to an inlier of a structure, its top- $k$  data  $\tau_j$  contains mostly inliers from the same structure, *i.e.* exploitation on this structure. Conversely, we achieve exploration on other possible structure if  $\theta_j$  is an exemplar to a gross outlier. We demonstrate this in Fig. 4.

Our method is different to Multi-GS [2] and ITKSF [19], since we do not explicitly compute the data similarity which is computationally expensive. In our method, the data similarity is captured in the correlation of the top- $k$  data of a hypothesis and the top- $h$  hypotheses of a datum. Moreover, we use the set of promising hypotheses to achieve exploration in the early stage of sampling but exploitation in the later stage. This is because the set of promising hypotheses is progressively refined as sampling proceeds. Fig. 5 shows the percentage of all-inlier sample found by our method, Multi-GS [2] and ITKSF [19]. As can be seen, the percentages of all-inlier sample found by Multi-GS and ITKSF do not increase with time, but our method shows an increasing trend over time.

### 2.4. Gross Outliers Identification

Having assigned an exemplar to data at different time, we use the information given by the exemplar to determine

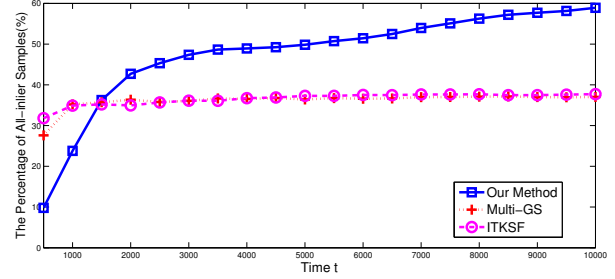


Figure 5. The average percentage of all-inlier samples found by our method, Multi-GS [2] and ITKSF [19] over time.

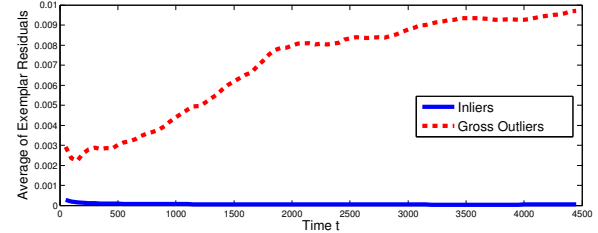


Figure 6. The average of exemplar residuals corresponds to inliers and outliers over time.

whether a datum is inlier or gross outlier.

For each datum  $x_i$ , a set of its exemplar is denoted by  $E_i = \{e_i^{(1)}, \dots, e_i^{(t)}\}$  where  $e_i^{(t)}$  is the exemplar assigned to  $x_i$  at time  $t$ . We average the residuals of  $x_i$  as measured to the hypotheses corresponding to the exemplars in  $E_i$ , *i.e.*

$$s_i^{(t)} = \frac{\sum_{e_i^{(t)} \in E_i} r(i, e_i^{(t)})}{|E_i|}, \quad (8)$$

where  $r(i, e_i^{(t)})$  is the absolute residual of  $x_i$  as measured to the hypothesis  $\theta_{e_i^{(t)}}$ . We call this averaged residual  $s_i^{(t)}$  an *exemplar residual* of  $x_i$  at time  $t$ .

To identify outliers at time  $t$ , we separate the data into two clusters by fitting a Gaussian mixture model with two components to the exemplar residuals of all data, *i.e.*  $\{s_1^{(t)}, \dots, s_N^{(t)}\}$ . The identified outliers are the data assigned to the component with larger mean. Conversely, a set  $\bar{X}$  of identified inliers is the data assigned to the component with the smaller mean.

Fig. 6 shows the average of the exemplar residuals corresponding to inliers and outliers over time. It can be seen that gross outliers have higher value of exemplar residual than inliers, and the exemplar residual of outliers is gradually increased as time progresses. This is because the set of promising hypotheses are progressively refined as sampling proceeds, hence, we have stronger evidence to identify outliers (as illustrated in Fig. 1).

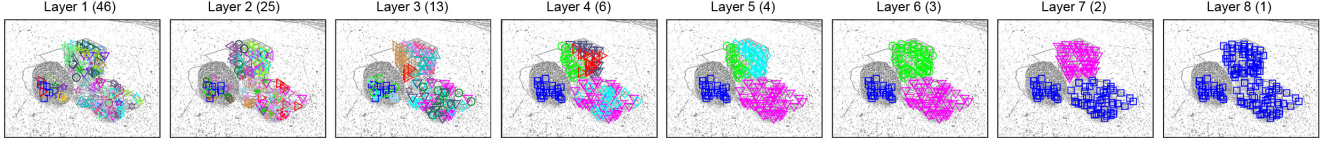


Figure 7. Hypothesis hierarchy for input data in Fig. 1: the model fitting result of each layer (Best viewed in color). The bottom most layer is shown in the first column, and the top most layer is shown in the last column. The number of hypotheses in each layer is bracketed. Note that the identified outliers are not shown.

### 3. Hypothesis Hierarchy

By leveraging a set  $\bar{\Theta}$  of promising hypotheses and a set  $\bar{X}$  of identified inliers returned by our proposed sampling method, we construct a hypothesis hierarchy in a bottom-up manner, such that it captures the fine to coarse interpretation of data. We cluster the hypotheses in each layer and only propagate a subset of hypotheses to the upper layer. The algorithm of constructing the hypothesis hierarchy is summarized in Alg. 2.

Similar to the top- $k$  data defined in Sec. 2.1, for each hypothesis  $\theta_j \in \bar{\Theta}$ , we obtain its top- $k$  data  $\tau_j$  by sorting the residual of data in  $\bar{X}$  as measured to  $\theta_j$ . Let  $H^{(\ell)}$  be the set of hypotheses in layer  $\ell$  of the hierarchy and  $H^{(1)} = \bar{\Theta}$ . To construct the layer  $\ell > 1$ , we cluster the hypotheses in  $H^{(\ell-1)}$  by measuring the intersection between their corresponding top- $k$  data. The intersection of two hypotheses  $\theta_j$  and  $\theta_{j'}$  in  $H^{(\ell-1)}$  is measured by

$$d^{(\ell-1)}(\theta_j, \theta_{j'}) = \frac{|D_{\tau_j} \cap D_{\tau_{j'}}|}{k}, \quad (9)$$

where  $k = p + (\ell - 1)$ . The value of  $k$  indicates the number of top ranked data to consider, *i.e.* for larger  $k$ ,  $d^{(\ell-1)}(\theta_j, \theta_{j'})$  is higher since they share more common elements in their top- $k$  data.

A hypothesis  $\theta_j$  is merged with hypothesis  $\theta_{j'}$  if the following conditions are satisfied:

$$d^{(\ell-1)}(\theta_j, \theta_{j'}) > 0.5 \text{ and } \max_{\theta_{j'} \in H^{(\ell-1)}/\theta_j} d^{(\ell-1)}(\theta_j, \theta_{j'}), \quad (10)$$

which means that  $\theta_j$  is merged to its closest  $\theta_{j'}$  if it shares more than 50% common data with  $\theta_{j'}$ . After merging all hypotheses in  $H^{(\ell-1)}$ , we obtain a set of clusters. For each cluster  $C$ , we only propagate the hypothesis  $\theta_{j'}$  to  $H^{(\ell)}$  if  $\theta_{j'}$  is the hypothesis in  $C$  with the minimum sum of squared residuals over its top- $k$  data, *i.e.*

$$\theta_{j'} = \arg \min_{\theta_j \in C} \sum_{i \in D_{\tau_j}} r(i, j)^2, \quad (11)$$

where  $r(i, j)$  is the absolute residual of  $x_i$  as measured to the hypothesis  $\theta_j$ . We keep constructing the layers of the hierarchy until there is only one hypothesis left.

---

#### Algorithm 2 Construct Hypothesis Hierarchy

---

- 1: **Input** a set  $\bar{\Theta}$  of promising hypotheses, a set  $\bar{X}$  of identified inliers, size of a minimal subset  $p$
  - 2: **Output** a hypothesis hierarchy  $H$
  - 3: Initialize the layer 1 as  $H^{(1)} = \bar{\Theta}$
  - 4: **Let**  $\ell = 2$  and  $k = p + (\ell - 1)$
  - 5: **while**  $|H^{(\ell)}| > 1$  **do**
  - 6:   obtain a set of cluster by merging the hypotheses in  $H^{(\ell-1)}$  based on the conditions in (10)
  - 7:   **for** each cluster  $C$  **do**
  - 8:      $H^{(\ell)} \leftarrow \theta_{j'} \in C$  which satisfies (11)
  - 9:   **end for**
  - 10: **if**  $H^{(\ell)} = H^{(\ell-1)}$  **then**
  - 11:    $k = k + 1$
  - 12: **else**
  - 13:    $\ell = \ell + 1$
  - 14:    $k = k + 1$
  - 15: **end if**
  - 16: **end while**
- 

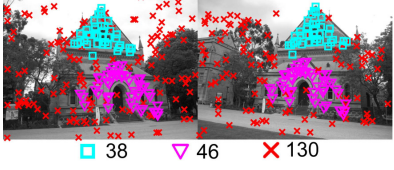
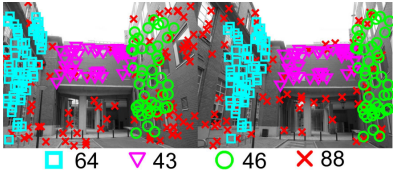
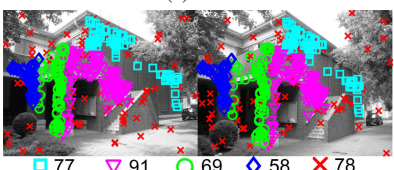
During the construction of hierarchy, we gradually increase  $k$  from the lower layer to the upper layer. In the lower layer, it contains a larger but less distinctive set of hypotheses, hence, we obtain a finer interpretation of data. Conversely, the upper layer contains a smaller but more distinctive set of hypotheses, hence, we obtain a coarser interpretation of data. Fig. 7 shows an example of our hypothesis hierarchy. As can be seen from the result in each layer, our hierarchical approach to model fitting allows multiple levels of details to be captured.

### 4. Experiments

We evaluate the performance of the proposed method (*DHF*) on our dataset, called *AdelaideRMF*<sup>1</sup>. We compare our guided sampling scheme against 7 sampling techniques: Random sampling as in RANSAC (Random) [7], proximity sampling (Proximity) [10], LO-RANSAC [4], Guided-MLESAC (GLMESAC) [18], PROSAC [3], Multi-GS [2] and ITKSF [19]. We also compare our hypothesis filtering

<sup>1</sup>Our dataset AdelaideRMF can be downloaded from <http://cs.adelaide.edu.au/~hwong/doku.php?id=data>

Table 1. Performance of sampling methods on multiple homographies estimation. We record the number of generated hypotheses (M), the number of all-inlier minimal subset found on each structure (Structures), the percentage of all-inlier samples found within the time budget (IS) and the parameter error (PE). We also record the CPU time to hit at least one all-inlier sample on each structure (Time) which is penalized by the time budget if a method fails. For ITKSF and DHF, the hypothesis filtering result is shown in bracket. The results represent the average over 50 runs with the best result boldfaced.

Data	Method	M	Time (s)	Structures	IS (%)	PE
 <p>(a) Hall</p> <p>□ 38   ▽ 46   × 130</p>	Random	<b>2427</b>	2.403	[2,5]	0.29	1.103
	Proximity	2376	1.001	[6,10]	0.67	1.095
	LO-RANSAC	2418	1.200	[8,27]	1.45	1.092
	GMLESAC	<b>2427</b>	0.765	[8,14]	0.91	1.066
	PROSAC	2422	<b>0.141</b>	[9,16]	1.03	1.058
	Multi-GS	849	0.142	[87,86]	20.38	0.946
	ITKSF	1032 (2)	0.234	[124(0),133(0)]	24.90	0.916
	DHF	1624 (64)	0.187	<b>[472(22),555(32)]</b>	<b>63.24</b>	<b>0.793</b>
 <p>(b) Building</p> <p>□ 64   ▽ 43   ○ 46   × 88</p>	Random	<b>2425</b>	2.453	[11,3,3]	0.7	0.979
	Proximity	2357	1.109	[35,5,9]	2.08	0.968
	LO-RANSAC	2394	1.757	[21,5,4]	1.25	0.965
	GMLESAC	2424	1.416	[22,6,5]	1.36	0.830
	PROSAC	2411	0.810	[40,13,9]	2.57	0.839
	Multi-GS	802	<b>0.263</b>	[115,28,66]	26.06	0.731
	ITKSF	983 (1)	0.320	[163(0),57(0),104(0)]	32.96	<b>0.728</b>
	DHF	1608 (52)	0.266	<b>[358(15),173(12),204(9)]</b>	<b>45.71</b>	0.740
 <p>(c) Lab</p> <p>□ 77   ▽ 91   ○ 69   ◇ 58   × 78</p>	Random	<b>2439</b>	3.43	[4,8,2,1]	0.64	0.717
	Proximity	2268	1.443	[10,14,5,6]	1.54	0.713
	LO-RANSAC	2319	2.23	[6,30,3,5]	1.9	0.706
	GMLESAC	2346	3.509	[6,13,2,1]	0.94	0.625
	PROSAC	2332	3.085	[7,20,2,3]	1.37	0.658
	Multi-GS	648	0.406	[62,76,16,37]	29.48	<b>0.547</b>
	ITKSF	803(1)	0.39	[95(0),119(0),34(0),57(0)]	37.98	0.555
	DHF	1423(75)	<b>0.298</b>	<b>[214(24),317(21),140(12),207(14)]</b>	<b>61.7</b>	0.56

against ITKSF. The model fitting performance is compared to RANSAC [7], J-Linkage [17] and Mean Shift [5].

In all sampling experiments, we fix  $b = 50$ ,  $k = \lceil 0.1 \times N \rceil$ ,  $h = \lceil 0.05 \times |\bar{\Theta}| \rceil$  and  $\delta = 0.3$ ,  $b$  being the batch size and  $|\bar{\Theta}|$  being the number of promising hypotheses (cf. Sec. 2.2). All experiments are run on a machine with 2.53 GHz Intel Core 2 Duo processor and 4GB RAM.

#### 4.1. Homography Estimation

We evaluate our method on estimation of multiple planar homographies on real images. The keypoint correspondences in each image pair are generated by SIFT matching<sup>2</sup> [12]. Table 1 shows the marked keypoint correspondences and false correspondences are marked as red crosses. The Direct Linear Transformation [9] is used to estimate a homography from 4 correspondences. Each method is given 50 runs, each for 5 CPU seconds.

Apart from the evaluation criteria used in previous work [2, 19] to compare sampling methods, we also compare the methods using the parameter error (PE). The parameter error is the minimum distance of a hypothesis  $\theta$  to the ground truth parameter  $\theta^*$ , i.e.  $\min_s \|\theta - \theta_s^*\|^2$  where  $\theta_s^*$  is the ground truth parameter of structure  $s$ . The ground

truth parameter of structure  $s$  is generated from all inliers of structure  $s$ , and both  $\theta$  and  $\theta^*$  is normalized to 1.

Table 1 summaries the experimental result of sampling methods. Multi-GS, ITKSF and DHF outperform all other methods in terms of the time to hit at least one all-inlier samples on each structure (Time), the average percentage of all-inlier samples found (IS) and the average of parameter error (PE). Our method obtains more all-inlier samples on each structure (Structures) than Multi-GS and ITKSF on all test data. For hypothesis filtering, as shown in Table 1, ITKSF fails to keep the all-inlier samples found on each structure and leads to the failure of the model fitting procedure. Our method is more stable than ITKSF, and guarantees that enough hypotheses and all-inlier samples on each structure remains for model fitting.

To conduct the model fitting experiment in a fair way, we use our guided sampling scheme for the competing methods. The model fitting result of our method is a hypothesis hierarchy. However, for the ease of comparison, the number of structure is given to each method and the performance is compared to the ground truth parameter using the model fitting error (Error). The model fitting error is taken as the minimum parameter error (PE) given by different permutation of the estimated models to the ground truth models. For RANSAC and J-Linkage, the inlier threshold is set to the average inlier threshold of all structures. For Mean Shift,

<sup>2</sup>Code from <http://www.vlfeat.org/~vedaldi/code/sift.html>

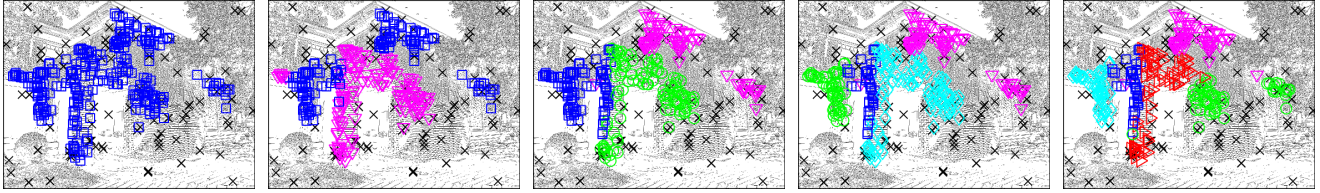


Figure 8. Top 5 layers in hypothesis hierarchy on Lab data (Table 1(c)). The top most layer in the hierarchy is shown in the first column, and black crosses indicate the gross outliers detected by DHF (Best viewed in color).

Table 2. Performance of model fitting methods on multiple homographies estimation. We report the average of model fitting error (Error) over 50 runs with best result boldfaced.

Error	RANSAC	J-Linkage	Mean Shift	DHF
Hall	0.105	0.068	0.089	<b>0.054</b>
Building	0.451	0.382	0.354	<b>0.228</b>
Lab	0.2	0.129	0.25	<b>0.094</b>

we tune the bandwidth to obtain the desirable number of models.

The model fitting results are shown in Table 2. It can be seen that our method obtains the smallest model fitting error on all test data due to the promising hypotheses and identified inliers returned by our proposed dynamic sampling. For Lab data, Fig. 8 shows the identified outliers (marked by black crosses) and the top 5 layers in the hypothesis hierarchy. As can be seen, the hierarchy provides different levels of details to interpret the data.

## 4.2. Fundamental Matrix Estimation

We also evaluate our proposed method on fundamental matrix estimation. The image pairs with marked keypoint correspondences (obtained from SIFT matching) are shown in Table 3. We use the 8-point algorithm [9] to estimate a fundamental matrix<sup>3</sup>. Each method is given 50 random runs and each for 10 CPU seconds.

The sampling results are shown in Table 3. For data with more than two structures, apart from Multi-GS, ITKSF and DHF, all other competing methods fail to find at least one all-inlier sample on each structure within the time budget. Multi-GS and ITKSF is more efficient in hitting all-inlier sample on each structure (Time). However, DHF obtains significantly more all-inlier samples on all test data than Multi-GS and ITKSF. DHF yields at least 50% IS on all test data. From Table 3, our hypothesis filtering scheme is very effective in filtering unpromising hypotheses, but also successfully remains the promising hypotheses on each structure. Our hypothesis filtering scheme is more reliable than ITKSF.

The model fitting results are summarized in Table 4. As expected, our proposed method achieves the smallest model

<sup>3</sup><http://www.robots.ox.ac.uk/~vgg/hzbook/code/>

Table 4. Performance of model fitting methods in fundamental matrix estimation. The average of model fitting error (Error) over 50 runs is reported with best result boldfaced.

Error	RANSAC	J-Linkage	Mean Shift	DHF
Cube-Toy	0.715	0.428	<b>0.33</b>	<b>0.33</b>
Book-Box	0.468	0.365	0.324	<b>0.318</b>
Toys	0.631	0.462	0.49	<b>0.366</b>

fitting error on all test data. Fig. 9 shows the top 5 layers in the hypothesis hierarchy on Book-Box data. It can be seen that our method successfully identifies the gross outliers (black crosses) in the data, and the model fitting result on each layer shows the advantage of our hierarchical approach, which captures multiple levels of details to interpret the data.

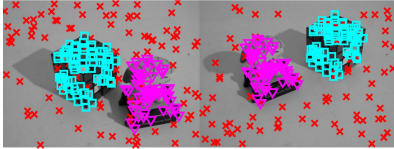
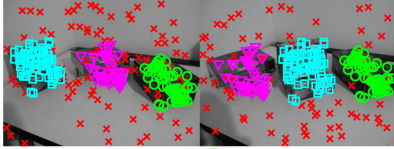
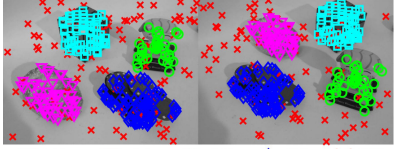
## 5. Conclusion

We proposed a novel dynamic hypothesis generation algorithm for multi-structure data based on analysing the correlation of preference induced by data and hypothesis residuals. By accumulating the evidence in the search space, our method uses the information to dynamically (1) identify outliers, (2) filter unpromising hypotheses, and (3) bias the sampling for discovery of multiple structures in data. The outcome of our sampling method is a set of promising hypotheses and identified inliers. This directly supports the proposed hierarchical approach to multi-structure model fitting, which captures the fine to coarse interpretation of data. Experimental results show that our proposed method outperforms previous methods in terms of both sampling accuracy and model fitting accuracy.

## References

- [1] A. Brahmachari and S. Sarkar. Blogs: Balanced local and global search for non-degenerate two view epipolar geometry. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1685–1692. IEEE, 2009.
- [2] T.-J. Chin, J. Yu, and D. Suter. Accelerated hypothesis generation for multi-structure robust fitting. In *European Conference on Computer Vision*, 2010.
- [3] O. Chum and J. Matas. Matching with PROSAC- progressive sample consensus. In *CVPR*, 2005.
- [4] O. Chum, J. Matas, and J. Kittler. Locally optimized RANSAC. In *DAGM*, 2003.

Table 3. Performance of sampling methods on fundamental matrix estimation. The notation is same as Table 1.

Data	Method	M	Time (s)	Structures	IS (%)	PE
(a) Cube-Toy  <span style="color: cyan;">□</span> 78 <span style="color: magenta;">▽</span> 72 <span style="color: red;">×</span> 99	Random	<b>6193</b>	9.39	[1,0]	0.02	1.148
	Proximity	5990	3.884	[4,3]	0.12	1.137
	LO-RANSAC	6112	1.83	[8,11]	0.31	1.147
	GMLESAC	6146	6.29	[3,1]	0.07	1.166
	PROSAC	6126	8.137	[1,1]	0.03	1.153
	Multi-GS	939	<b>0.149</b>	[179,172]	37.38	0.823
	ITKSF	1870 (17)	0.230	[380(2),369(9)]	40.05	<b>0.815</b>
	DHF	3466 (74)	0.258	<b>[1084(27),720(25)]</b>	<b>52.05</b>	0.837
(b) Book-Box  <span style="color: cyan;">□</span> 67 <span style="color: magenta;">▽</span> 41 <span style="color: green;">○</span> 54 <span style="color: red;">×</span> 97	Random	<b>6009</b>	10.0	[0,0,0]	0.0	0.949
	Proximity	5822	9.928	[2,0,1]	0.05	0.939
	LO-RANSAC	5950	8.66	[17,1,4]	0.37	0.945
	GMLESAC	5962	10.0	[0,0,0]	0.0	0.899
	PROSAC	5964	10.0	[0,0,1]	0.02	0.930
	Multi-GS	917	<b>0.246</b>	[151,58,103]	34.02	0.717
	ITKSF	1813(4)	0.329	[313(2),138(0),225(0)]	37.29	0.698
	DHF	3405(66)	0.636	<b>[1199(26),225(6),353(10)]</b>	<b>52.19</b>	<b>0.636</b>
(c) Toys  <span style="color: cyan;">□</span> 71 <span style="color: magenta;">▽</span> 49 <span style="color: green;">○</span> 38 <span style="color: blue;">◇</span> 81 <span style="color: red;">×</span> 88	Random	<b>4997</b>	10.0	[0,0,0,0]	0.0	1.091
	Proximity	4724	10.0	[1,0,0,1]	0.04	1.069
	LO-RANSAC	4985	10.0	[3,0,0,5]	0.16	1.087
	GMLESAC	4968	10.0	[0,0,0,0]	0.0	1.085
	PROSAC	4818	10.0	[0,0,0,1]	0.02	1.086
	Multi-GS	795	<b>0.592</b>	[105,54,30,119]	38.74	0.658
	ITKSF	1490 (1)	0.63	[222(1),129(0),84(0),249(0)]	45.91	<b>0.622</b>
	DHF	2836 (65)	1.596	<b>[375(15),230(12),120(6),815(25)]</b>	<b>54.3</b>	0.638

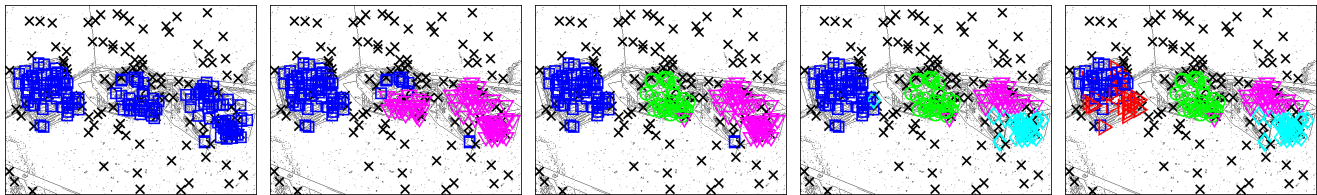


Figure 9. Top 5 layers in hypothesis hierarchy on Book-Box data (Table 3(b)). The top most layer in the hierarchy is shown in the first column, and black crosses indicate the gross outliers detected by our method (Best viewed in color).

- [5] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 24(5):603–619, 2002.
- [6] R. Fagin, R. Kumar, and D. Sivakumar. Comparing Top  $k$  Lists. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, page 36. Society for Industrial and Applied Mathematics, 2003.
- [7] M. A. Fischler and R. C. Bolles. RANSAC: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM*, 24:381–395, 1981.
- [8] L. Goshen and I. Shimshoni. Balanced exploration and exploitation model search for efficient epipolar geometry estimation. *IEEE transactions on pattern analysis and machine intelligence*, pages 1230–1242, 2007.
- [9] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.
- [10] Y. Kanazawa and H. Kawakami. Detection of planar regions with uncalibrated stereo using distributions of feature points. In *BMVC*, 2004.
- [11] P. Kidwell, G. Lebanon, and W. Cleveland. Visualizing incomplete and partially ranked data. *IEEE Transactions on Visualization and Computer Graphics*, 14:1356–1363, 2008.
- [12] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.
- [13] P. McIlroy, E. Rosten, S. Taylor, and T. Drummond. Deterministic sample consensus with multiple match hypotheses. In *BMVC*, 2010.
- [14] N. Scherer-Negenborn and R. Schaefer. Model fitting with sufficient random sample coverage. *IJCV*, 89:120–128, 2010.
- [15] R. Subbarao and P. Meer. Nonlinear mean shift for clustering over analytic manifolds. In *CVPR*, 2006.
- [16] R. Subbarao and P. Meer. Nonlinear mean shift over Riemannian manifolds. *IJCV*, 84(1):1–20, 2009.
- [17] R. Toldo and A. Fusiello. Robust multiple structures estimation with j-linkage. In *European Conference on Computer Vision*, pages 537–547. Springer, 2008.
- [18] B. J. Tordoff and D. W. Murray. Guided-MLESAC: Faster image transform estimation by using matching priors. *TPAMI*, 27(10):1523–1535, 2005.
- [19] H. S. Wong, T.-J. Chin, J. Yu, and D. Suter. Efficient multi-structure robust fitting with incremental top-k lists comparison. In *Asian Conference on Computer Vision (ACCV)*, 2010.