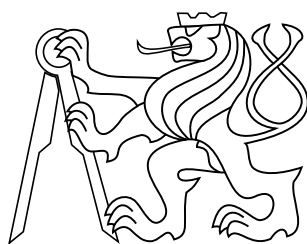


bachelor's thesis

Efficient Many-to-many RANSAC

Jonáš Šerých



May 2015

Supervisor: doc. Mgr. Ondřej Chum, Ph.D.

Czech Technical University in Prague
Faculty of Electrical Engineering, Katedra kybernetiky

BACHELOR PROJECT ASSIGNMENT

Student: Jonáš Šerých
Study programme: Open Informatics
Specialisation: Computer and Information Science
Title of Bachelor Project: Efficient Many-to-Many RANSAC

Guidelines:

The project addresses the problem of using groups of mutually exclusive correspondences (multiple correspondence).

1. Familiarize yourself with the literature.
2. Assemble the relevant set of pairs of images.
3. Study statistical properties of inliers in multiple correspondences.
4. Suggest algorithm(s) of RANSAC that uses multiple correspondence, either at some stage verification or even in generating hypotheses from the minimum samples in PROSAC.
5. Consider the possibility that tentative correspondences are given by vector quantization of the descriptor space (correspondence of visual words).
6. Check the properties of the proposed algorithms.

Bibliography/Sources:

- [1] Karel Lebeda, Ondřej Chum, Jiří Matas: Fixing LO RANSAC, BMVC 2012
- [2] Wei Zhang, Jana Košecká: Generalized RANSAC framework for relaxed correspondence problems. WACV06
- [3] Ondřej Chum, Jiří Matas: PROSAC, CVPR 2006
- [4] James Philbin et al. : Object retrieval with large vocabularies and fast spatial matching, CVPR 2007

Bachelor Project Supervisor: doc. Mgr. Ondřej Chum, Ph.D.

Valid until: the end of the summer semester of academic year 2015/2016

L.S.

doc. Dr. Ing. Jan Kybic
Head of Department

prof. Ing. Pavel Ripka, CSc.
Dean

Prague, January 21, 2015

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: Jonáš Š e r ý c h

Studijní program: Otevřená informatika (bakalářský)

Obor: Informatika a počítačové vědy

Název tématu: Efektivní RANSAC z nejednoznačných korespondencí

Pokyny pro vypracování:

Projekt se zabývá problémem využití skupin vzájemně se vylučujících korespondencí (násobných korespondencí).

1. Seznamte se s literaturou.
2. Sestavte relevantní datovou sadu párů obrázků.
3. Studujte statistické vlastnosti inlierů v násobných korespondencích.
4. Navrhněte algoritmus (algoritmy), které v RANSACu násobné korespondence využijí, a to buď v některé fázi verifikace nebo i při generování hypotéz z minimálních vzorků, i ve variantě PROSAC.
5. Uvažte možnost tentativních korespondencí daných vektorovou kvantizací descriptorů (korespondencí vizuálních slov).
6. Ověřte vlastnosti navržených algoritmů.

Seznam odborné literatury:

- [1] Karel Lebeda, Ondřej Chum, Jiří Matas: Fixing LO RANSAC, BMVC 2012
- [2] Wei Zhang, Jana Košecká: Generalized RANSAC framework for relaxed correspondence problems. WACV06
- [3] Ondřej Chum, Jiří Matas: PROSAC, CVPR 2006
- [4] James Philbin et al. : Object retrieval with large vocabularies and fast spatial matching, CVPR 2007

Vedoucí bakalářské práce: doc. Mgr. Ondřej Chum, Ph.D.

Platnost zadání: do konce letního semestru 2015/2016

L.S.

doc. Dr. Ing. Jan Kybic
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 21. 1. 2015

Acknowledgement

I would like to express my greatest gratitude to my supervisor Ondra Chum for his guidance and the amount of time he spent on the consultations.

I would also like to thank Jiří Matas for providing us with invaluable suggestions and ideas.

Last, but not least, I am deeply grateful to my girlfriend Anička, my family, friends and to everybody else who have supported and encouraged me.

Prohlášení autora práce

Prohlašuji, že jsem předloženou práci vypracoval samostatně, a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne

.....
podpis autora práce

Abstract

V této bakalářské práci se věnujeme problému robustního odhadu dvoupohledové geometrie z mnohonásobných korespondencí. Způsoby, kterými se standardně získávají tentativní korespondence, zajišťují, že se každý zájmový bod uplatní pouze v jediné potenciální korespondenci. Tyto konstrukce množiny potenciálních korespondencí, obvykle založené na testu poměru vzdáleností [1], nebo na vzájemné blízkosti [2] příznaků, přirozeně vyřazují potenciálně víceznačné korespondence.

Ukážeme, že v některých typech scén, například ve scénách obsahujících opakované struktury, nesou mnohonásobné korespondence cenné informace, které mohou být využity pro vylepšení odhadu geometrie.

Navrhujeme čtyři nové varianty algoritmu LO-RANSAC, z nichž všechny využívají mnohonásobných korespondencí k tomu, aby v určitých situacích poskytly lepší výsledky než standardní LO-RANSAC.

Shromáždili jsme přes 50 dvojic fotografií ze standardních testovacích datových sad obohacených o naše vlastní snímky a využili je k otestování navrhovaných algoritmů a jejich porovnání se standardním algoritmem LO-RANSAC. Na základě výsledků těchto testů jsme potvrdili, že jeden z námi navrhovaných algoritmů překonává standardní LO-RANSAC a přitom v případech, které zvládá původní algoritmus řešit dobře, není výpočetně náročnější.

Klíčová slova

LO-RANSAC; mnohonásobné korespondence; opakující se struktury

Abstract

In this bachelor thesis, we investigate the problem of robust two-view geometry estimation from many-to-many correspondences. In standard approaches, the construction of tentative correspondences ensures that each feature point participates in at most one potential correspondence. Such constructions, typically based on a distance ratio [1] test or mutually nearest property [2], naturally drops potentially ambiguous correspondences.

We show, that for certain types of scenes, such as those containing repeated structures, many-to-many correspondences contain valuable information that can be utilized in order to improve the geometry estimation.

Four new variants of the LO-RANSAC algorithm are proposed, each of them using the additional many-to-many correspondences in order to get better results than the standard algorithm in some scenarios.

We have collected more than 50 image pairs from standard benchmark datasets and our own photos and used them to test all of our proposed algorithms against the state-of-the-art LO-RANSAC. Based on the experimental results, we have concluded, that one of our proposed algorithms outperforms the standard LO-RANSAC, while not introducing any additional computational cost in the cases, when the original algorithm works well.

Keywords

LO-RANSAC; many-to-many correspondences; repeated structures

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Thesis contribution	1
1.3	Structure of the thesis	1
2	Prior work	2
2.1	Geometric relations between pairs of images	2
2.1.1	Homography	2
2.1.2	Epipolar geometry	3
2.2	Local features	4
2.2.1	Feature detection	4
2.2.2	Feature description	5
2.3	Feature matching	5
2.4	RANSAC	6
2.4.1	RANSAC variants	8
2.4.2	Sampling	8
	Degensac	8
	PROSAC	9
2.4.3	Hypothesis verification	9
	WaldSAC	9
	LO-RANSAC	9
3	Utilizing many-to-many correspondences	10
3.1	Matching	11
3.2	Proposed LO-RANSAC variants	12
3.2.1	Many-to-many TCs only inside LO	12
3.2.2	Many-to-many TCs in verification and LO	13
3.2.3	Full many-to-many TCs use in all steps	13
3.2.4	Switching the versions on the run	13
3.3	Determining the support set	14
3.4	Implementation details	14
3.4.1	Matching	14
3.4.2	Sampling	15
3.4.3	Verification	15
4	Experiments	16
4.1	Scenes handled well by LO-RANSAC	16
4.2	Scenes hard for LO-RANSAC	17
4.3	Very hard scenes	18
4.3.1	Multiple good models	18
4.3.2	Not enough inliers	20
4.4	Experiments summary	20
5	Conclusions	21
	Appendix A Implementation	22
A.1	Installation	22
A.2	The C library API	22

A.3 Using the library inside Matlab	23
A.3.1 Mass testing script	23
A.3.2 The dataset	24
A.3.3 Mex functions	24
Appendix B Complete experimental results	25
Bibliography	38

1 Introduction

1.1 Motivation

The problem of estimating a two-view geometry and is an essential part of several computer vision tasks such as wide-baseline stereo matching [3, 2], structure and motion estimation [4], image retrieval [5] and more.

Contemporary state-of-the-art algorithms achieve good results at reasonable speed on usual image pairs, but certain scene configurations can effect their accuracy, stability or speed in a bad way. Moreover, some scenes are even not solvable at all by those algorithms.

The standard algorithms use one-to-one feature correspondences and usually achieve good results, but there are certain scenarios that cause them to fail.

We discuss the use of many-to-many correspondences, because they help in some of these situations and their utilization has not been covered by the literature yet.

1.2 Thesis contribution

This bachelor thesis contributes to the state of the art in two-view geometry estimation by RANSAC. It provides a discussion about the effective utilization of many-to-many correspondences. New feature matching strategy is proposed and the newly created set of tentative correspondences is then utilized in proposed LO-RANSAC algorithm modifications, outperforming the standard state-of-the-art LO-RANSAC.

1.3 Structure of the thesis

In chapter 2, we start by an introduction to the prior work related to our topic. The geometry of image matching is briefly described, as well as the standard phases of image matching, including local feature detection and description and establishing of the tentative correspondence set. We finish the first introductory part by describing the RANSAC algorithm and some of its numerous variants.

In chapter 3, utilization of many-to-many correspondences is discussed and four different new RANSAC variants are proposed.

Experimental results are analyzed in chapter 4 and chapter 5 contains the conclusions.

Appendices contain manual for the implementation and all the experimental results.

2 Prior work

This chapter reviews some of the background of the image matching necessary for understanding the thesis.

The image matching is done by estimating the two-view geometry relating given images. The goal is to estimate the geometric model and to find a set of point-to-point correspondences between the images.

In most of the use cases of image matching, the images are heavily distorted by various effects, as described later in this chapter, and approaches based on some form of direct comparison of the image intensities are not suitable. Instead, local feature based methods are widely used.

The standard approach of image matching consists of three distinct steps. First, local regions of interest are found in each of the images. They are then represented by an appropriate descriptor. Second, tentative correspondences are generated by a search in the descriptor space.

Final step is then to estimate the geometric model from these correspondences by RANSAC [6] or some other robust estimator.

2.1 Geometric relations between pairs of images

We assume that we have two images taken by a projective camera (i.e. pinhole camera). Matched photos are usually related by one of the two geometric relations introduced in the text below.

2.1.1 Homography

A *homography* is a transformation often relating images in two special situations. The first one is that of two photos being taken with two cameras sharing one projection center. This often arises when the images for a panoramic photo are taken with one camera rotating around its projection center.

The second possible configuration where homography can be used are images of flat scenes. The camera poses and intrinsic parameters are not restricted in any way and the only condition that has to be met is that the scene 3D points have to be coplanar. Planar structures are often found in architecture, so homography can relate for example two views of one building facade.

Homography can be represented as rank 3 matrix $\mathbf{H} \in \mathbb{R}^{3 \times 3}$, bounding the pairs of image points in the following way:

$$\lambda_i \vec{x}'_i = \mathbf{H} \vec{x}_i \quad (1)$$

$$\lambda_i \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (2)$$

Here x_i, y_i and x'_i, y'_i are coordinates of the related points in the first and the second image respectively with respect to the bases of the images. The vectors augmented with a new row of 1 are called *homogeneous coordinates* of \vec{x}'_i and \vec{x}_i

After representing the matrix \mathbf{H} by its rows $\vec{h}_1^\top, \vec{h}_2^\top$ and \vec{h}_3^\top and rewriting the equation 2 we get:

$$\lambda_i x'_i = \vec{h}_1^\top \vec{x}_i \quad (3)$$

$$\lambda_i y'_i = \vec{h}_2^\top \vec{x}_i \quad (4)$$

$$\lambda_i = \vec{h}_3^\top \vec{x}_i \quad (5)$$

then it is possible to eliminate the lambdas from the first two equations

$$\left(\vec{h}_3^\top \vec{x}_i\right) x'_i = \vec{h}_1^\top \vec{x}_i \quad (6)$$

$$\left(\vec{h}_3^\top \vec{x}_i\right) y'_i = \vec{h}_2^\top \vec{x}_i \quad (7)$$

$$(8)$$

and after putting everything to one side

$$\vec{x}^\top \vec{h}_1 - (x'_i \vec{x}^\top) \vec{h}_3 = 0 \quad (9)$$

$$\vec{x}^\top \vec{h}_2 - (y'_i \vec{x}^\top) \vec{h}_3 = 0 \quad (10)$$

Then represent these two equations in a matrix form and we get

$$\begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -x'x & -x'y & -x' \\ 0 & 0 & 0 & x & y & 1 & -y'x & -y'y & -y' \end{bmatrix} \begin{bmatrix} \vec{h}_1 \\ \vec{h}_2 \\ \vec{h}_3 \end{bmatrix} = \vec{0} \quad (11)$$

$$\mathbf{Z} \vec{h} = \vec{0} \quad (12)$$

In order to compute the homography \mathbf{H} , the matrix \mathbf{Z} must be of rank 8. This is so because of the fact that if $\mu \mathbf{H}' = \mathbf{H}$ and $\mu \neq 0$, then both matrices represent the same homography and so we want one-dimensional subspace of rank 3 real 3×3 matrices.

Therefore we need 4 correspondences as each provide us with two rows of \mathbf{Z} . Moreover no three points can be collinear, since such configuration would leave us with $\text{rank}(\mathbf{Z}) < 8$

2.1.2 Epipolar geometry

An epipolar geometry is a relation between two images of the same 3D scene. An illustration of the scene configuration can be seen in Fig. 1. We denote X the scene 3D point and x_1 and x_2 its projections to first and second camera respectively. The camera's projection centers are called C_1 and C_2 . Their projections to the other camera images - e_1 and e_2 are called *epipoles* and the plane formed by the camera centers and X is a *epipolar plane*.

The epipolar geometry is then represented by a *fundamental matrix* $\mathbf{F} \in \mathbb{R}^{3 \times 3}$, with $\text{rank}(\mathbf{F}) = 2$, relating the image points in the following *epipolar constraint*

$$\vec{x}_2^\top \mathbf{F} \vec{x}_1 = 0 \quad (13)$$

with x_1 and x_2 represented in the homogeneous coordinates.

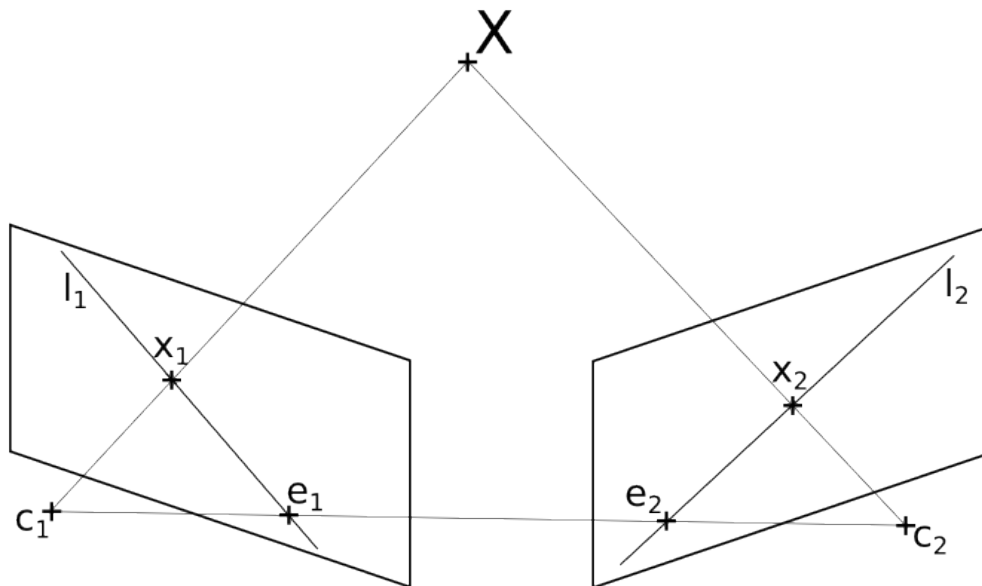


Figure 1 Epipolar geometry of two cameras

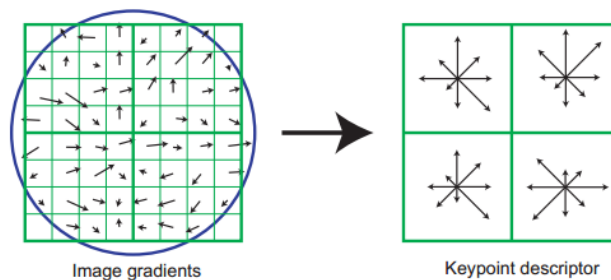


Figure 2 SIFT histograms of gradients (4 spacial bins are used in this illustration, but the standard is to use 16 of them.) (taken from [1])

This equation can be understood as a relation between the image points x_1 , x_2 and the corresponding lines l_2 and l_1 , respectively, in the second image (also represented in homogeneous coordinates).

$$\mathbf{F}x_1 = \vec{l}_2 \tag{14}$$

$$\mathbf{F}^T x_2 = \vec{l}_1 \tag{15}$$

The equations 14 and 15 show us, that by transforming an image point (with exception of the epipole) by a fundamental matrix, we get a line in the other image, on which the corresponding image point should lie.

A fundamental matrix can be estimated from 7 point correspondences as described in [7].

2.2 Local features

2.2.1 Feature detection

In the first step, the regions of interest are found in the images. These regions should satisfy the condition of repeatability i.e. the probability of detecting the same feature

in the other image should be high.

Hence, they should be local - only relatively small patch of pixels, because global features are prone to occlusion and viewpoint changes.

Also, the regions have to be information rich, i.e. have significant texture. Regions that are too small lead to too similar features which then in turn badly affects their discriminative power.

There are several image feature detection algorithms. Standard examples include Hessian affine region detector (HAF) by Mikolajczyk [8], Maximally stable extremal regions (MSER) by Matas et al. [2] and DoG (Difference of Gaussians) used in Scale-invariant feature transform (SIFT) detector by Lowe [1].

The detected regions are represented as affine frames.

2.2.2 Feature description

In order to be able to find corresponding features across the images, the detected pixel patches have to be represented in a way that is invariant to affine transformations and illumination changes and is both stable with respect to detection phase and highly discriminative.

The following commonly used procedure was described by Lowe [1]. First, the pixel patches defined by the detected region of interest affine frames are transformed to canonical patches. This is done by transforming the elliptical patch to a circle with given diameter and rotating it then by the features dominant orientation.

The patch orientation can be determined by computing a histogram of gradients orientations. The peaks are then detected in the histogram and the biggest one is considered to be the dominant orientation assigned to the feature. However, multiple peaks can be detected and if they are within 80% of the highest one, they are used to create new features.

Representing the features in this canonical way ensures affine invariance of all the measures performed on the patches.

Second, the canonical patches are described using histograms of gradients. The reasoning behind this choice lies in the fact that using the gradients ensures invariance to affine illumination changes and histogram introduces higher robustness.

The canonical patch is divided into 4×4 grid of spatial cells. A weighted histogram of gradients is computed in each of these cells. To eliminate the effect of any errors of the feature detection phase, the gradients on the boundaries of the spatial cells vote in each of the neighboring cells.

The histograms used have 8 bins each, so this whole description can be represented by $4 * 4 * 8 = 128$ -dimensional vector.

Last, the descriptor vectors are normalized to unit length, its values are thresholded at 0.2 in order to suppress the effects of non-linear illumination changes and then the descriptors are renormalized to unit length again.

2.3 Feature matching

When the features are detected and described, the next step is to generate tentative correspondences by matching the features based on a nearest neighbor search in the descriptor space. The standard approach described below was introduced by Lowe [1]. He discovered that the absolute distance of the descriptors is not a good measure of goodness of the match as not all the descriptors are equal. Some of them are much

more discriminative than others and there is no good distance threshold to be used in the search.

A *first-to-second* (FTS) ratio test was proposed in order to generate a set of so called *tentative correspondences* (TCs). The tentative correspondences are all of the point correspondences proposed to be used later in a model estimation algorithm. These correspondences have to be verified as not all of them are correct, e.g. the TC set may contain a correspondence relating a corner of a window in one image and a door knob in the other one.

If the ratio of the distances to the first and the second nearest neighbor is too high (usually thresholded at 0.8, but can be adjusted to get better results in some cases), the correspondence is discarded, because we cannot be sure which of the two similar possible matches is the good one and the probability of choosing the wrong one is too high. On the other hand, if the test is passed, the correspondence of the feature and its first nearest neighbor in the second image is added to the tentative correspondences set.

Setting the first-to-second ratio threshold to 0.8 eliminates 90% of false matches, while less than 5% of good ones are discarded on Lowe's synthetic set of randomly altered images (rotated, scaled, and affected by noise). While his threshold is widely adapted, better thresholding may be possible on different datasets.

Lowering the amount of false matches this dramatically increases the inlier ratio of the TC set, allowing big speedups of the robust estimator applied in the next phase of the image matching procedure.

2.4 RANSAC

Random sample and consensus - RANSAC [6] is a robust fitting algorithm. Classical techniques used for parameter estimation, such as least squares, aim to fit all the noisy data on input by given model, but they depend on an assumption that with enough input data, correct fitting model can be found despite the noise. These techniques deal well with small errors on the data points, but even one piece of data can completely destroy the fit if it is affected by a gross error as shown in Fig. 3.

In the image matching problem we need to fit tentative correspondences (TCs) of visually similar local image features by an appropriate geometric model - usually either homography or epipolar geometry.

The correspondences are affected by two kinds of errors. First, by some small noise coming from inexact feature detection, insufficient photo resolution and other factors.

The second type of error comes from bad local feature matching. This can be caused by illumination changes between the images, big difference in angles under which the photos were taken, repetition, occlusion and other factors.

Incorrectly matched features, called outliers (as opposed to correct correspondences - inliers), can take much more than 50% of all the data. See Fig. 4 for example of typical TCs.

The idea behind RANSAC lies in significant change in the optimization goal. Instead of minimizing the sum of squared errors it aims for maximizing the size of hypothesis support set (the number of data consistent with the hypothesis model under given threshold of some cost function, e.g. Sampson error for fundamental matrix estimation).

The RANSAC algorithm works as follows. Minimal size data samples (the size is the minimal number of data needed to produce finite number of hypotheses fitting the minimal sample exactly) are taken repeatedly and the model(s) estimated from

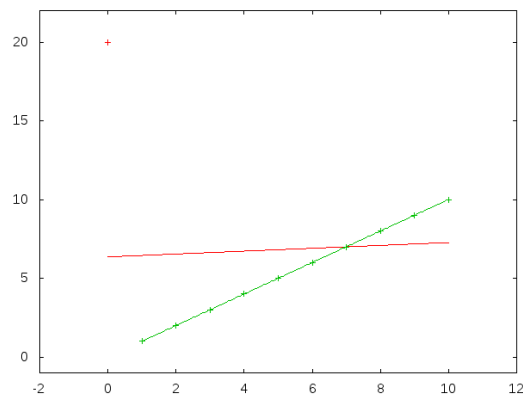


Figure 3 Even a single bad datapoint can completely ruin fitting with least squares. (red line is least squares fit to all the data, green line is the fit unaffected by gross error of the point (0, 20))

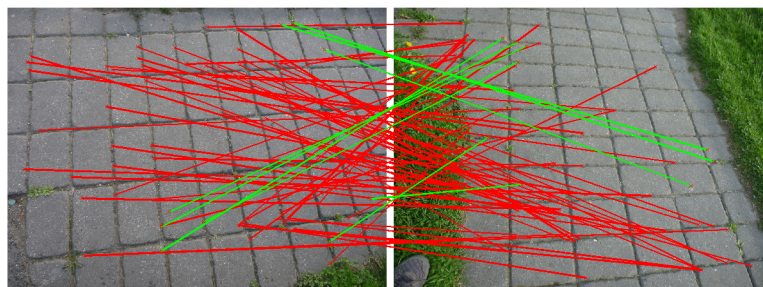


Figure 4 An example of a tentative correspondences set. Notice the high outlier ratio. (inliers are green, outliers red)

the sample is computed. Then, in the second phase, the quality of the estimation is evaluated on all the input data. Standard RANSAC uses number of inliers, that is data points consistent with the model (defined by a threshold on some loss function), as the cost function. The best-so-far model is remembered and the whole process is repeated until the probability of finding better model falls under some user-given threshold (usually 0.05 or 0.01).

The procedure stands on an assumption that fitting an all-inlier minimal sample produces a correct model.

The probability of picking such sample is:

$$P(I) = \frac{\binom{I}{m}}{\binom{N}{m}}$$

Where I is number of inliers, m is minimal sample size and N is total number of TCs.

We can then derive this upper bound:

$$P(I) = \prod_{j=0}^{m-1} \frac{I-j}{N-j} \leq \left(\frac{I}{N}\right)^m$$

2 Prior work

The probability of missing an all-inlier sample after k samples is then

$$\eta = (1 - P_I)^k \quad (16)$$

We want to repeat sampling until probability of a fail falls under predefined threshold η_0 . In order to find out the number of samples needed, we transform the equality 16 to an inequality and solve it for k

$$\begin{aligned} (1 - P_I)^k &\leq \eta_0 \\ k \log(1 - P_I) &\leq \log \eta_0 \\ k &\geq \frac{\log \eta_0}{\log(1 - P_I)} \end{aligned} \quad (17)$$

Equation 17 gives us a stopping criterion for the RANSAC procedure. See algorithm 1 for pseudocode description of the standard RANSAC procedure.

Algorithm 1 RANSAC

```
1:  $k \leftarrow 0, k^* \leftarrow \infty$ 
2: while  $k \leq k^*$  do
3:    $k \leftarrow k + 1$ 
4:    $S_k \leftarrow$  randomly drawn minimal sample
5:    $M_k \leftarrow$  model estimated from  $S_k$ 
6:    $I_k \leftarrow$  support set of  $M_k$ 
7:   if  $|I_k| > |I^*|$  then
8:      $I^* \leftarrow I_k$ 
9:      $M^* \leftarrow M_k$ 
10:     $k^* \leftarrow$  number of samples needed not to miss  $|I^*| + 1$  inliers (Eq. 17)
11:  end if
12: end while
```

2.4.1 RANSAC variants

Some of RANSAC's assumptions are not always valid, so there are many variants of the algorithm, each dealing with some of the original algorithm's shortcomings. Some of them are described in the following section.

2.4.2 Sampling

Degensac

Degensac [9] is a modification for EG estimation dealing with dominant plane in scene. If the scene contains such plane (as usual in architecture), there is a big probability of drawing a sample with 5 or more coplanar correspondences. These samples are degenerate and we cannot compute the correct epipolar geometry from them.

The Degensac algorithm checks for this kind of degeneracy by performing test on the minimal sample. It can then reject these samples early and save lot of time by skipping the resource-demanding verification phase of RANSAC.

PROSAC

The main idea of PROSAC [10] is to utilize the a priori knowledge of the TCs quality.

It has been proposed to sort the TCs according to their first-to-second ratio. PROSAC is then sampling from the set of the presumably best correspondences. This sampling pool set is then progressively enlarged according to a grow function described in [10]. After predefined number of samples PROSAC converges to the standard RANSAC's uniform sampling.

While the sampling phase is altered, the verification phase does not change - i.e. the hypothesis is verified on all of the data.

The PROSAC algorithm is designed in such a way that it essentially draws the same samples as the standard RANSAC, but in other order. This means it should not effect the quality of the estimation, only speed the process up.

2.4.3 Hypothesis verification

WaldSAC

There were numerous attempts to speed the verification step up. Notably, WaldSAC [11] uses sequential probability ratio test to minimize the amount of TC evaluations.

Imagine verifying a hypothesis and not getting a single inlier after evaluating 600 out of 750 correspondences. It seems like the probability of this hypothesis being the best one is very low and we can reject it without evaluating all the data. The idea of WaldSAC is similar, but it addresses the problem in an exact statistical language of Wald test - it is minimizing the number of the TCs being evaluated subject to probability of rejecting good hypothesis staying under user-given limit.

This approach leads to rejecting some of the good samples thus leading to more samples needed, but the speedup caused by discarding bad samples early is bigger.

LO-RANSAC

One of the important RANSAC's assumptions - that every all-inlier sample generate a good model - is not entirely true as discussed in [12].

Locally optimized RANSAC [12, 13] uses local optimization after finding the support set for the hypothesis in each RANSAC cycle. The inliers are then given to inner RANSAC routine that does not use minimal samples, because the data should be outlier-free at that point. It also performs a classic least-squares optimization in each cycle in order to get better model with more inliers.

This local optimization is performed only on the best-so-far samples, because it is computationally demanding. Moreover, the LO step is skipped for the first 30 samples.

3 Utilizing many-to-many correspondences

The FTS test discussed in previous chapter leads to higher inlier ratio by sacrificing absolute number of inliers in the TC set. This approach often pays off, but there is a risk of discarding too many of them.

Moreover, a good correspondence is not always the first nearest neighbor in the SIFT space.

We have analyzed the properties of inliers on our dataset and it turned out, that the number of second nearest neighbor inliers usually stays roughly between 20 and 75% of the number of first nearest neighbors. See Fig. 5 for the plot of the number of Kth nearest neighbor inliers relative to the number of the first nearest neighbor. In some cases, we have found even more second NN inliers than the first NN ones. Moreover, after performing the FTS ratio test, the number of first neighbor inliers lowers and the next nearest neighbors relative count becomes even higher.

This phenomenon can occur for example in a scene with lot of repetitive patterns, very commonly found in man-made structures, especially in architecture, where there are many similar features. For example all the windows on a facade may look very alike and there may be even more than one mutually exclusive correct geometric model, if there are not any unique discriminative features in the images. See Fig. 6 for an example of such situation.

Zhang and Kosecka [14] proposed a *generalized RANSAC* algorithm in order to utilize more tentative correspondences. They do not discard correspondences based on first-to-second ratio and they do not keep just the first nearest neighbor. Instead they use all the nearest neighbors as tentative correspondences and propose several sampling strategies in order to both utilize the additional good correspondences and keep the number of samples needed reasonably low.

They propose two stage sampling with a goal of not to getting more than one tentative correspondence of the same feature in one image. First, they sample the features in first image. After sampling the first points, the corresponding second image features are sampled.

They then introduced a proportional sampling, where the probability of the first

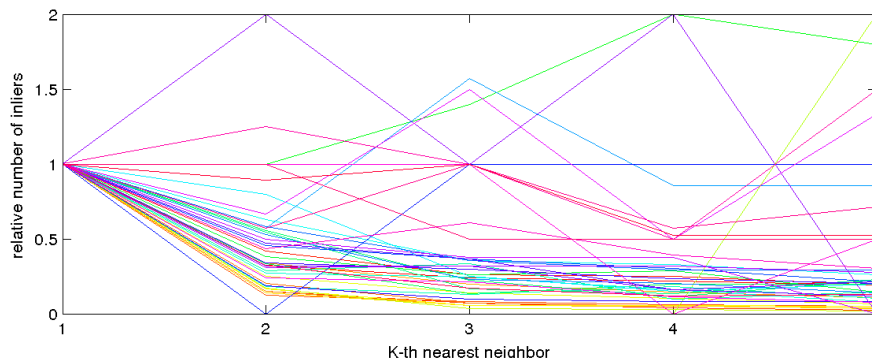


Figure 5 Relative numbers of inliers for KNN, each line represents one image pair.

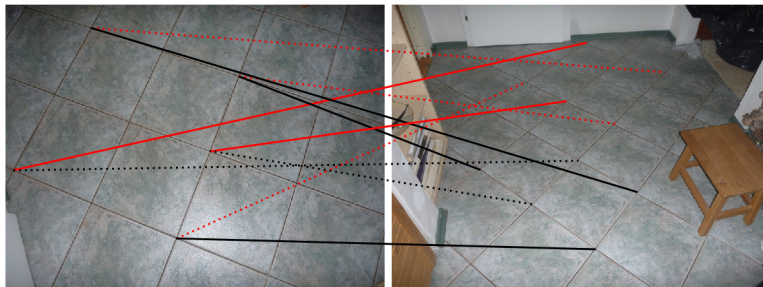


Figure 6 repetitive structures causing the standard matching to discard useful tentative correspondences. Red lines represent the correct TCs, the solid lines are the first nearest neighbors, the dotted lines are second nearest neighbors (standardly not used). Moreover, note that most of these correspondences would not pass the FTS test.

image feature being sampled is proportional to number of the feature’s correspondences. This approach prefers sampling of one-to-one correspondences and therefore is faster.

We propose a new approach to dealing with many-to-many correspondences.

3.1 Matching

Our method uses the standard best (in a sense of SIFT distance) matches passing the first-to-second test. We denote this set of tentative correspondences as \mathcal{T}_s (standard).

We then create another set \mathcal{T}_a (additional), containing many-to-many correspondences.

\mathcal{T}_a contains up to K nearest neighbors for each image feature. The i -th nearest neighbor is included in \mathcal{T}_a if three conditions hold.

1. the first-to-second nearest neighbor ratio is bigger than threshold θ_{fts} (i.e. the feature is not included in standard \mathcal{T}_s)
2. the first-to- i -th NN ratio is bigger than θ_{fts} (i.e. the i -th neighbor is still quite similar to the first one)
3. the absolute SIFT distance of the i -th nearest neighbor is lower than some threshold θ_{abs} (i.e. the two features are not too dissimilar).

We set the relative threshold θ_{fts} the standard 0.8 and we use 230 as the absolute threshold θ_{abs} . Our choice is based on a statistics of a good correspondences SIFT distances done by Mikulik et al. [15].

The last step is to create the union of these two sets, $\mathcal{T}_m = \mathcal{T}_s \cup \mathcal{T}_a$

This set of correspondences imposes a new constraint on the RANSAC support set. Any combination of the TCs can be a support set in the standard scenario with only one-to-one correspondences. However, this does not hold with the many-to-many \mathcal{T}_m set, since the TCs relating one point in the first image with multiple points in the second image can never be present in a support set at the same time.

3.2 Proposed LO-RANSAC variants

Standard LO-RANSAC performs very well on most of the image pairs. In such case the standard TC set \mathcal{T}_s contains so many good tentative correspondences, that it is sufficient to estimate a good model. Therefore, the best application for the additional set of correspondences \mathcal{T}_a is to increase the number of inliers of the correct model.

We assume that the noise on the good TCs from \mathcal{T}_a has the same distribution as the noise on the standard \mathcal{T}_s . If this assumption holds, getting more inliers is beneficial for the accuracy of the estimated geometry.

In the case that there are enough good correspondences in the \mathcal{T}_s set, it is not reasonable to sample from the whole \mathcal{T}_m as the *generalized RANSAC* would do on these images, because there are enough inliers in \mathcal{T}_s to find a good model and the inlier ratio of \mathcal{T}_s is much bigger than the one of \mathcal{T}_m . Therefore, the number of samples needed and the overall running time would be very high if we were sampling from the bigger set \mathcal{T}_m .

However when the number of inliers on \mathcal{T}_s is very small, it may become reasonable to sample from \mathcal{T}_m .

Tentative correspondences from \mathcal{T}_m can be employed at various phases of the LO-RANSAC - in sampling, verification or local optimization and in several combinations of these RANSAC phases.

We propose three LO-RANSAC-based algorithms ranging from minimal to maximal amount of use of the many-to-many TCs. The first one works on \mathcal{T}_s and only the LO is done on the \mathcal{T}_m set, the second one computes the hypothesis verification on \mathcal{T}_m on top of that. The last one works on the whole \mathcal{T}_m in all of the phases of the LO-RANSAC. We describe all of them and their purposes in the following sections.

3.2.1 Many-to-many TCs only inside LO

The first of our proposed algorithms is the most lightweight one. It is purposed for situations that the standard LO-RANSAC can solve, but its solution has low number of inliers.

The aim of introducing the \mathcal{T}_a correspondences is to enhance the good model by supporting it by more inliers and to stabilize the solution.

The problem can be solved by LO-RANSAC and therefore it would not be reasonable to sample from the bigger \mathcal{T}_m set as discussed above. Hence, the first proposed algorithm samples and verifies only on \mathcal{T}_s and it uses the whole \mathcal{T}_m only inside the local optimization.

The LO step with \mathcal{T}_m involved is entered after getting a best-so-far sample on \mathcal{T}_s . We assume that there are enough good \mathcal{T}_s TCs for the good hypothesis to be recognized.

The stopping sample count is computed in a standard way from the number of inliers on \mathcal{T}_s in accordance with the inequality 17, with

$$P_I = \frac{|\mathcal{S} \cap \mathcal{T}_s|}{|\mathcal{T}_s|}$$

where \mathcal{S} is the best model's support set.

This algorithm version is efficient, because the number of samples needed should be the same as the original LO-RANSAC. However, the use of additional inliers from \mathcal{T}_a introduces a slowdown as there are more TCs entering the computation demanding local optimization step. This loss of the speed should not be very significant, as the number of LO steps performed is usually very low (usually staying under 5).

3.2.2 Many-to-many TCs in verification and LO

The first algorithm relies on the assumption that the \mathcal{T}_s set contains so many good TCs that it is not only possible to take an all-inlier sample, but that it is also possible to distinguish a hypothesis generated by such sample from a hypothesis generated by contaminated minimal sample (i.e. containing at least one outlier) based only on the information provided by \mathcal{T}_s .

Our second proposed algorithm is intended for situations where there is at least one good minimal sample present in \mathcal{T}_s , but the above assumption does not hold.

It still samples only from the \mathcal{T}_s , because it is possible to take an all-inlier sample from it. Nevertheless, in order to find the good model, the hypotheses have to be verified on the whole \mathcal{T}_m . Because of this, it is slower than the first algorithm, but it enters the LO step at presumably better samples and it should be able to find a good solution even when there are almost no good TCs present in \mathcal{T}_s .

The stopping criterion is computed in the same way as in algorithm 1, i.e. only from inliers on \mathcal{T}_s .

3.2.3 Full many-to-many TCs use in all steps

The last proposed algorithm uses correspondences from the whole \mathcal{T}_m in all of the phases of LO-RANSAC. It samples from all of the tentative correspondences in \mathcal{T}_m , verifies each hypothesis on the whole \mathcal{T}_m and performs the LO on it too.

It is much slower than the previous two versions because it not only has to process more TCs in verification and LO phase, but mainly because the inlier ratio on \mathcal{T}_m is usually much smaller, than the inlier ratio on just \mathcal{T}_s , therefore the number of samples needed is much higher than in the previous two versions and LO-RANSAC.

There are two scenarios when it pays off to use this algorithm. The first one is when there is no all-inlier sample present in standard TC set \mathcal{T}_s at all. A one-to-one RANSAC would have no chance of succeeding under such conditions.

The other use may be in the case when we cannot generate the one-to-one correspondences in a first place. One example of such situation is the use of quantized descriptors - the bag of visual words (BoW), because when we do not have the original descriptors and only their quantization is available, creating one-to-one set of tentative correspondences in a standard way is not possible as the first-to-second ratios are unknown.

The proposed algorithm samples uniformly from the \mathcal{T}_m , i.e. each TC is equally likely to be sampled. Proportional sampling from *generalized RANSAC* could be used when the K (number of nearest neighbors used in matching) is higher than 2 as not to prefer big clusters of unlikely correspondences.

3.2.4 Switching the versions on the run

The proposed algorithm versions are each suitable for different situations as described above. However, one of our goals is to design an universal algorithm, that would be as fast as the standard LO-RANSAC in 'easy' situations when it performs well and still utilize the many-to-many correspondences in order to get better results on 'hard' scenes.

We therefore propose a switching version, that works only on \mathcal{T}_s and switches to using the \mathcal{T}_m set only when necessary. This change of the strategy is based on the stopping criterion from equation 17.

3 Utilizing many-to-many correspondences

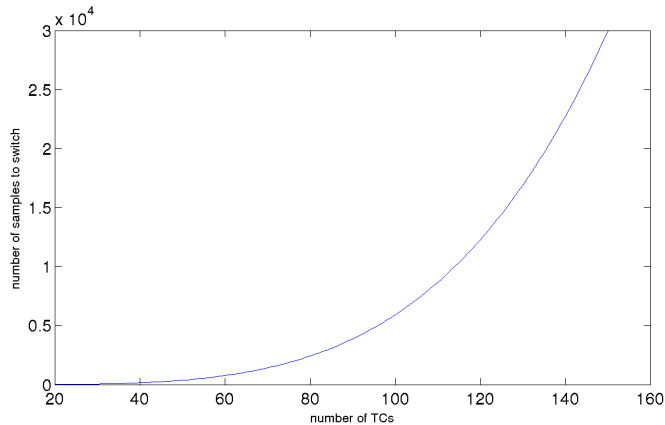


Figure 7 Plot of the switching criterion with respect to $|\mathcal{T}_s|$

The stopping criterion for the hypothetical case that the standard TC set \mathcal{T}_s contains only θ_{min} (we empirically set $\theta_{min} = 15$) inliers is computed.

If this stopping criterion is reached and if the support set of the best model found so far is smaller than θ_{min} , the algorithm switches to our proposed version with many-to-many correspondences in verification and LO phase, since in such case, the probability of a success of the plain LO-RANSAC working on the \mathcal{T}_s is very low.

We can see from Fig. 7, that this approach can be useful only with relatively small TC sets, since the amount of samples needed for a strategy switch to occur is over 10000 for a \mathcal{T}_s set with more than 100 TCs. However it can be concluded from the results of our experiments, that there are scenes that both satisfy this condition and benefit from switching the strategy.

3.3 Determining the support set

Introducing many-to-many correspondences leads to a problem of determining the support set while not violating the constraints of mutual exclusiveness of the many-to-many correspondences.

Not checking the consistence of the support set can easily lead to impossible inlier sets, which in turn can ruin the results of the whole procedure. The effect is well visible in epipolar geometry estimation of urban scenes. When some of the repetitive structures in the images are positioned on a line, there is chance that all tentative correspondences containing that structure will be included in the support set since their reprojection error fits under the error threshold.

3.4 Implementation details

3.4.1 Matching

We assign two numbers to each tentative correspondence, id_L and id_R , representing the identification numbers of left and right tentative correspondence's features. We also store the information about which set each TC belongs to (\mathcal{T}_s or \mathcal{T}_a).

3.4.2 Sampling

When we sample from the whole \mathcal{T}_m , we do so uniformly, storing the *ids* of the sampled correspondence. When a TC is picked and at least one of its features is already in the sample, it is moved aside and the random picking of the correspondence is repeated.

3.4.3 Verification

The problem of finding the hypothesis support set on many-to-many correspondences set is a bipartite graph matching problem.

We have implemented a greedy algorithm that first sorts the conflicting inliers (inliers in the naïve sense - passing the error threshold) by their reprojection errors and then, starting from the lowest one, traverses the sorted list and allow each TC to be inserted to the support set only if there is no TC with colliding feature *id* in there.

4 Experiments

In this chapter we present the experimental results of our proposed algorithms used for homography estimation in comparison with the Lebeda’s LO-RANSAC [13].

The results are summarized in tables containing statistics of the overall number of inliers found (I), the number of inliers found on the standard TC set \mathcal{T}_s ($Isamp$), the inlier ratio on the used TC set (\mathcal{T}_s for LO-RANSAC, \mathcal{T}_m for our proposed algorithms) noted $I(\%)$, the number of samples needed ($Samp$), CPU time ($time$), mean error of the ground truth correspondences ($Error$) and number of local optimizations (LO). All the bold values are mean over 50 runs. The tables also contain standard deviations of all of the quantities except for CPU time.

The error was computed by reprojecting hand-made ground truth correspondences (about 8 of them for each image pair) by the model found by the algorithm used. The smaller the error, the better the precision of the algorithms, but this holds only to some extent, as the ground truth correspondences were obtained manually and thus their precision is not perfect.

The \mathcal{T}_m set of many-to-many correspondences was generated from first two nearest neighbors (both directions), the error threshold was set by using $\sigma = 0.5$ and the RANSAC confidence was set to 0.95.

All of the results can be found in Appendix B. We divide the problems into three types. First, scenes that can be solved well by the standard LO-RANSAC, second, scenes hard for LO-RANSAC on which our proposed algorithms provide better results and third, scenes on which all of the tested algorithms failed to provide good results.

Each of the types is discussed in detail on running examples in the following section.

4.1 Scenes handled well by LO-RANSAC

We have selected the *Boston* images (see Fig. 8), for the purpose of demonstrating the properties of our proposed algorithms on images in which the standard LO-RANSAC algorithm achieves stable results with low error and number of samples needed.

Solver→		LO-RANSAC	m2m v1	m2m v2	m2m full	m2m switch
Quantity ↓						
Boston	I	280.8 ±0.4	580.7 ±1.5	580.9 ±1.4	581.2 ±1.6	280.8 ±0.4
	Isamp	280.8 ±0.4	280.6 ±0.6	280.7 ±0.5	281.0 ±0.2	280.8 ±0.4
	I(%)	87.5 ±0.1	74.0 ±0.2	74.0 ±0.2	74.0 ±0.2	35.8 ±0.1
	Samp	9.3 ±4.7	9.3 ±4.7	9.3 ±4.7	181.1 ±0.6	9.3 ±4.7
	Time	18.2	43.6	45.2	118.6	16.2
	Error	0.53 ±0.04	0.55 ±0.04	0.55 ±0.04	0.55 ±0.04	0.53 ±0.04
	LO	1.0 ±0.0	1.0 ±0.0	1.0 ±0.0	2.3 ±1.1	1.0 ±0.0

Table 1 results for the Boston pair (Fig. 8)

The achieved results are summarized in Tab. 1.



Figure 8 Boston

Our proposed algorithms have found more than two times the amount of inliers that the standard LO-RANSAC found as was expected. Also the number of inliers on the \mathcal{T}_s set stays the same. The inlier ratio is slightly lower on our proposed TC set \mathcal{T}_m in agreement with our expectations.

The number of samples needed is consistent across the versions sampling only from \mathcal{T}_s and is much higher in the full many-to-many version, which samples from the whole \mathcal{T}_m . Because so much samples are needed, the running time is almost ten times longer in the full version. Our other proposed versions are about two times slower, because more TCs are entering the demanding LO step than in the plain LO-RANSAC.

Note that the last proposed algorithm (strategy switching) did not switch the strategy and therefore worked exactly like the standard LO-RANSAC. The measured time is lower, but this is only due to matlab timing imperfections.

4.2 Scenes hard for LO-RANSAC



Figure 9 Notredame14

The *sym_notredame14* Fig. 9 image pair is difficult to match mainly because of the big difference of the illumination of the photos and because of the symmetry and presence of repetitive structures.

The high difficulty of matching of the *sym_notredame14* pair is visible on the small number of inliers found by the standard LO-RANSAC algorithm as shown in Tab. 2. Although the mean LO-RANSAC error is high, its median is 3.4 meaning that the standard algorithm can sometimes deal with this situation, but its results are not stable.

Solver→		LO-RANSAC	m2m v1	m2m v2	m2m full	m2m switch
Quantity ↓						
Notredame	I	11.1 ±1.3	45.7 ±6.4	47.2 ±1.7	48.6 ±0.6	47.4 ±1.7
	Isamp	11.1 ±1.3	12.4 ±1.2	12.7 ±0.7	13.0 ±0.0	12.8 ±0.6
	I(%)	18.1 ±2.1	14.6 ±2.1	15.1 ±0.6	15.6 ±0.2	15.2 ±0.5
	Samp	4421.3 ±2384.4	3460.5 ±2542.8	1918.7 ±599.2	500000.0 ±0.0	3028.8 ±551.7
	Time	64.2	93.2	68.0	5602.8	98.0
	Error	24.84 ±55.39	7.80 ±32.52	1.16 ±0.50	0.95 ±0.05	1.12 ±0.44
	LO	3.5 ±1.0	3.4 ±0.9	3.9 ±1.2	8.2 ±2.5	7.6 ±1.5

Table 2 results for the sym_notredame14 pair (Fig. 9)

This example also captures one of the cases described in the previous chapter - we have enough inliers to find the correct model, but not enough to consistently recognize that we have found it. We can see that the mean error of proposed algorithm version 2 is much smaller than that of the version 1, as we have assumed (although the medians of both errors are equal to 1), the number of samples needed is also greatly reduced by using the \mathcal{T}_m set in the verification phase as version 2 does.

The full many-to-many version achieves even smaller mean error for the price of needing many samples (it had reached the maximal number of samples allowed in our experiments) and entering more LO steps and thus being very slow.

In terms of the mean error, the last proposed algorithm (switching) achieved similar results as algorithm 2, while needing smaller number of samples than the standard LO-RANSAC.

4.3 Very hard scenes

The results on very hard image pairs are discussed in this section. We have found two types of such scenarios, first, scenes with lot of repetition and multiple good models and second, scenes with so small number of good tentative correspondences available, that the correct model has support set smaller than random bad models.

4.3.1 Multiple good models

The first scenario is represented by the *dlazdice* pair that contains a lot of repetitive patterns and the first image contains very little discriminative features. This causes the pair to be difficult to match correctly even for a human.

Both the original LO-RANSAC and our proposed algorithms usually find a good model, but it is often not the correct one.

The errors of all of the algorithms in Tab. 3 indicate that the correct model was not found, however, the number of inliers is consistently high and the models found were not random, as shown in Fig. 11. The images at the top of the figure contain manually acquired ground truth. The reprojection of the mesh in the left image by the homography found by LO-RANSAC is shown at the bottom. When we remove the inliers of this homography from the TC set and run the LO-RANSAC again, we get more models that are matching some of the tiles well, but the correct model is only found after many iterations of this procedure and its support set is small.



Figure 10 Dlazdice

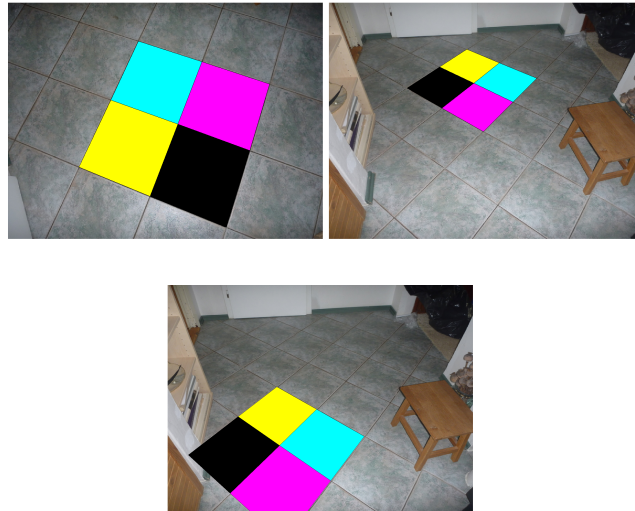


Figure 11 dlazdice scene ground truth (top) and best model found by LO-RANSAC (bottom)

	LO-RANSAC	m2m v1	m2m v2	m2m full	m2m switch
I	48.2 ±2.4	171.6 ±9.5	170.9 ±10.5	173.0 ±7.1	48.2 ±2.4
Isamp	48.2 ±2.4	48.3 ±2.2	48.2 ±2.4	48.6 ±1.6	48.2 ±2.4
I(%)	9.4 ±0.5	6.8 ±0.4	6.7 ±0.4	6.8 ±0.3	1.9 ±0.1
Samp(10 ³)	42.1 ±13.0	41.4 ±12.0	41.6 ±12.1	500.0 ±0.0	42.1 ±13.0
Time	681.4	1025.0	1300.2	13666.6	460.2
Error	227.46 ±76.20	218.08 ±52.17	221.86 ±57.56	210.46 ±37.43	227.46 ±76.20
LO	5.5 ±1.5	5.5 ±1.5	5.8 ±1.8	8.8 ±2.1	5.5 ±1.5

Table 3 results for the dlazdice pair (Fig. 10)

Introducing the \mathcal{T}_m and our proposed algorithms did not solve this issue. The algorithm that would be able to find the correct homography relating this pair of images would have to somehow emphasize the influence of the small part of doors visible in the bottom left corner of the left image and the tiny part of the leg of the stool visible near the upper right corner.

Despite the fact that our algorithms were not able to find the correct model, the additional correspondences in the \mathcal{T}_a can still be utilized as they allow more good models to be found, which, in turn, can serve to recognize that there are more than one possible solutions.



Figure 12 LePoint3

4.3.2 Not enough inliers

The scenario with too small number of inliers, where no good model can be found, are represented by the *LePoint3* pair (Fig. 12).

The viewpoint change is so massive, that the local feature detection, description and matching phases fail to provide good set of tentative correspondences.

The results are summarized in Tab. 4. All of the algorithms have found models with similar errors, but unlike the *dlazdice* scene, the found model was completely wrong every time.

	LO-RANSAC	m2m v1	m2m v2	m2m full	m2m switch
I	9.2 ± 0.4	15.6 ± 2.6	17.3 ± 0.9	18.1 ± 1.0	17.4 ± 1.0
Isamp	9.2 ± 0.4	8.3 ± 1.4	9.1 ± 0.6	9.6 ± 0.5	9.2 ± 0.6
I(%)	32.7 ± 1.3	25.5 ± 4.3	28.3 ± 1.5	29.7 ± 1.6	28.5 ± 1.6
Samp	279.1 ± 43.5	293.6 ± 61.6	298.5 ± 80.9	5872.9 ± 1349.6	323.8 ± 70.6
Time	8.8	22.2	18.2	99.2	28.4
Error	72.26 ± 7.68	71.67 ± 9.41	71.03 ± 8.60	71.48 ± 10.27	71.35 ± 7.54
LO	1.8 ± 0.7	1.8 ± 0.7	2.0 ± 0.9	3.1 ± 1.2	3.6 ± 1.2

Table 4 results for the LePoint3 pair (Fig. 12)

4.4 Experiments summary

We have found three types of scenes, first solvable well by LO-RANSAC, second, where our proposed algorithm achieves better results, and last, not solved well by either of the algorithms used. We did the experiments on over 50 image pairs from various datasets and our own photos and we came to the conclusion that most of the scenes fall into the first and the last category and only very little of the image pairs could not be solved by the standard LO-RANSAC and still were solved successfully by some of our proposed algorithms.

5 Conclusions

In this thesis, we have investigated the possibility of utilizing many-to-many correspondences inside the RANSAC robust estimator.

We have justified the use of many-to-many tentative correspondences and proposed a new algorithm for generating the tentative correspondences.

We have then proposed three LO-RANSAC-based algorithms each focused on different scenario and an 'universal' algorithm, that keeps the properties of the state-of-the-art LO-RANSAC while still being able to utilize many-to-many correspondences and outperform the standard version on some of the difficult scenes.

We have then tested all of our proposed algorithms on a dataset consisting of more than 50 image pairs collected from various public datasets and enriched by our own images. The proposed algorithms were compared to the state-of-the-art LO-RANSAC and their assumed properties were confirmed.

The last proposed strategy switching algorithm has proved to perform at least as good as the standard, while providing better results in situations, where LO-RANSAC fails to find the correct model and/or shows a big instability. It does so while not introducing any additional cost for the image pairs already well solvable by the standard LO-RANSAC.

We have found only few image pairs on which our proposed algorithms could gain an advantage over the standard one-to-one approach, but it is possible, that the algorithms can be utilized for major improvement of the image matching in some specialized field.

A Implementation

The implementation is based on Lebeda’s LO-RANSAC [13].

The proposed algorithms are written in C, but Matlab mex functions are provided too. The installation, the library API and the user manual are provided in the following sections.

A.1 Installation

The ransac library depends on LAPACK [16], which should be available via packaging systems on most Linux distributions (package `liblapack`).

The `src` directory contains a Makefile for an easy building. Running it compiles all the source codes and creates a file `libransac.a` which can then be linked with other software.

Please change the `MEXEXT` variable inside the Makefile to the string returned by running Matlab command `mexext`. It is also necessary to set the `LAPACK_INCL` variable to point to your Matlab `extern/include` directory.

Running `make` in the `src` directory builds the library, compiles the matlab mex functions and installs everything into the `build` directory.

A.2 The C library API

The `libransac.a` library provides both the standard LO-RANSAC (`ransacH`) and our proposed many-to-many RANSAC variants (`m2m_ransacH`). Their headers are listed in Tab. 5.

All of their parameters are summarized in Tab. 6. There are four modes that the `m2m_ransac` can be working in, as described in Tab. 7.

All of the matrices are stored column-wise as to keep compatibility with Matlab.

Each tentative correspondence is represented by 6-dimensional vector obtained by concatenating the homogeneous coordinates of the corresponding points. The ids have to be positive and the maximal id should be as low as possible (sparse numbering would cause unnecessary memory and CPU usage).

Score	<code>ransacH</code>	<code>(double *u, int len, double th, double conf, int max_sam, double *H, unsigned char *inl, int *data_out, int do_lo, int inlLimit);</code>
Score	<code>m2m_ransacH</code>	<code>(double *u, int len, int len_first, int *ids, int mode, double th, double conf, int max_sam, double *H, unsigned char *inl, int *data_out, int do_lo, int inlLimit);</code>

Table 5 ransac functions headers

Name	I/O	type	size	content
H	out	double	3×3	the estimated homography
inl	out	unsigned char	len	inliers (1) and outliers (0)
data_out	out	int	3	# of samples, # of LOs and # of rejected models
u	in	double	$6 \times len$	the tentative correspondences set
len	in	int	scalar	number of all tentative correspondences used
th	in	double	scalar	squared inlier/outlier error threshold
conf	in	double	scalar	confidence ($1 - \eta_0$)
max_sam	in	int	scalar	maximal number of samples drawn
do_lo	in	int	scalar	LO switch (1/0)
inLimit	in	int	scalar	maximal number of inliers used in iterative least squares (0 = no limit)
len_first	in	int	scalar	number of standard tentative correspondences
ids	in	int	$2 \times len$	left and right ids of the tentative correspondences
mode	in	int	scalar	m2m algorithm variants (See Tab. 7)

Table 6 *ransacH* and *m2m_ransacH* parameters

mode	description
0	m2m version 1, the whole \mathcal{T}_m set is used only in LO
1	m2m version 2, the whole \mathcal{T}_m set is used in verification and LO
2	full m2m version, \mathcal{T}_m is used in sampling, verification and LO
4	switching m2m version, \mathcal{T}_m used only when \mathcal{T}_s contains too few inliers

Table 7 Description of *m2m_ransac* modes

A.3 Using the library inside Matlab

We provide a Matlab mex functions to use the `libransac.a` library inside Matlab.

A.3.1 Mass testing script

First, the paths have to be set up by running `setpaths.m` inside the `matlab` directory. It is then possible to navigate to the `data` directory and run `masstest.m` to execute the mass testing (run it once for each of the algorithms, configured at the beginning of `masstest.m`). Its results are stored in the `results` directory.

Run `$../generate_pdf ../data` inside the `results` directory to create a `results.mrt.pdf` report afterwards.

A Implementation

GT.h	3 by 3 double	ground truth homography
GT.tcs	6 by n double	ground truth correspondences
TC.std	6 by m double	standard \mathcal{T}_s set
TC.m2m.tcs	6 by k double	many-to-many \mathcal{T}_m set
TC.m2m.ids	2 by k integer	the \mathcal{T}_m correspondences ids
TC.m2m.std_len	scalar integer	number of \mathcal{T}_s correspondences at the beginning of TC.m2m.tcs
DATA.imgs(1,2).data	w by h by 3 integer	the image data
DATA.dataset	char string	the name of the dataset

Table 8 The structure of `mat` files in the `data` directory

A.3.2 The dataset

The dataset is stored inside `data` directory in form of pairs of image files and correspondent Matlab `mat` files (described in Tab. 8).

A.3.3 Mex functions

The provided mex functions have an API similar to the C `libransac.a` library and their usage is straightforward. Run `m2m_ransacH()` and `ransacH()` to get the usage help.

B Complete experimental results

The following pages contain the results of all the discussed LO-RANSAC variants on our dataset that consists of image pairs taken from the following sources:






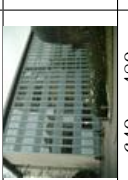
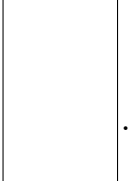
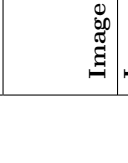






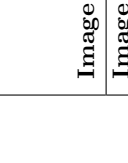
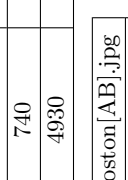
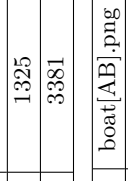
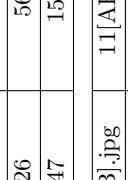
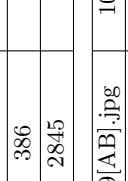
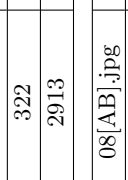
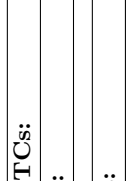
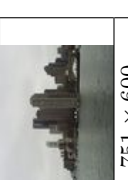

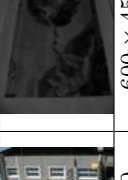




- *own* - our own images
- *lebeda* - images from Lebeda’s homog dataset [13]
- *sym* - images from symbench dataset [17]
- *zubud* - images from ZuBuD dataset [18]
- *sattler* - unpublished images by Torsten Sattler
- *chal* - images used in [19]
- *jav* - images collected from the internet by Javier Aldana-Iuit














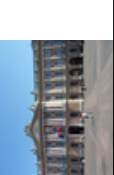
The experiments were conducted with the settings summarized in Tab. 9.

detector	Hessian Affine (HAF)
descriptor	SIFT
σ	0.5
$1 - \eta_0$	0.95
number of runs	100
sample limit	500000















Table 9 Experiments settings




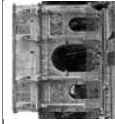










Mass RANSAC Test



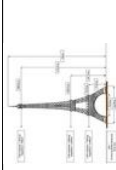









Filenames:	01[AB].jpg	02[AB].jpg	03[AB].jpg	04[AB].jpg	05[AB].jpg	06[AB].jpg	07[AB].jpg	sym_notredame13[AB].jpg
Dataset:	zubud	zubud	zubud	zubud	zubud	zubud	zubud	sym
Image A:								
Image dimensions:	640 × 480	640 × 480	640 × 480	640 × 480	640 × 480	640 × 480	640 × 480	460 × 600
Image B:								
Image dimensions:	640 × 480	640 × 480	640 × 480	640 × 480	640 × 480	640 × 480	640 × 480	460 × 600
Inlier error threshold:	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9
Inlier distance threshold:	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Standard TCs:	322	386	1326	562	1239	1325	740	185
m2m TCs:	2913	2845	3347	1502	3854	3381	4930	607
Filenames:	08[AB].jpg	09[AB].jpg	10[AB].jpg	11[AB].jpg	adam[AB].png	boat[AB].png	Boston[AB].jpg	
Dataset:	zubud	zubud	zubud	zubud	lebeda	lebeda	lebeda	
Image A:								
Image dimensions:	640 × 480	640 × 480	640 × 480	640 × 480	600 × 450	750 × 600	751 × 600	
Image B:								
Image dimensions:	640 × 480	640 × 480	640 × 480	640 × 480	600 × 450	750 × 600	751 × 600	
Inlier error threshold:	1.0	1.0	1.0	1.0	0.9	1.4	1.4	
Inlier distance threshold:	1.0	1.0	1.0	1.0	1.0	1.2	1.2	
Standard TCs:	1247	1110	748	733	257	381	321	
m2m TCs:	4122	3512	2067	3196	548	846	785	

Filenames:	BostonLib[AB].png	BruggeSquare[AB].jpg	BruggeTower[AB].png	Brussels[AB].jpg	CapitalRegion[AB].jpg	chal_cab[AB].jpg	chal_capitole[AB].jpg
Dataset:	lebeda	lebeda	lebeda	lebeda	lebeda	chal	chal
Image A:							
Image dimensions:	800 × 532	751 × 600	751 × 600	751 × 600	479 × 600	800 × 534	800 × 534
Image B:							
Image dimensions:	800 × 532	751 × 600	751 × 600	751 × 600	479 × 600	800 × 534	800 × 534
Inlier error threshold:	1.6	1.4	1.4	1.4	0.9	1.6	1.6
Inlier distance threshold:	1.3	1.2	1.2	1.2	1.0	1.3	1.3
Standard TCs:	334	335	362	516	31	207	134
m2m TCs:	1870	1202	789	1267	80	1772	833

Filenames:	chal_dagstuhl1[AB].jpg	chal_dagstuhl2[AB].jpg	chal_dagstuhl3[AB].jpg	chal_dagstuhl4[AB].jpg	chal_trondheim[AB].jpg	city[AB].png	dlazdice[AB].jpg
Dataset:	sattler	sattler	sattler	sattler	sattler	lebeda	own
Image A:							
Image dimensions:	450 × 600	800 × 600	800 × 600	800 × 600	800 × 600	329 × 278	800 × 600
Image B:							
Image dimensions:	800 × 600	800 × 600	800 × 600	800 × 600	800 × 600	329 × 278	800 × 600
Inlier error threshold:	0.9	1.6	1.6	1.6	1.6	0.3	1.6
Inlier distance threshold:	1.0	1.3	1.3	1.3	1.3	0.5	1.3
Standard TCs:	31	84	257	59	61	40	513
m2m TCs:	132	311	1374	203	253	85	2538

Filenames:	dlazky[AB].jpg	Eiffel[AB].png	ExtremeZoom[AB].png	graf[AB].png	LePoint1[AB].png	LePoint2[AB].png	LePoint3[AB].png
Dataset:	own	lebeda	lebeda	lebeda	lebeda	lebeda	lebeda
Image A:							
Image dimensions:	800 × 600	750 × 600	800 × 530	750 × 600	600 × 450	600 × 450	600 × 450
Image B:							
Image dimensions:	800 × 600	750 × 600	800 × 530	750 × 600	600 × 450	600 × 450	600 × 450
Inlier error threshold:	1.6	1.4	1.6	1.4	0.9	0.9	0.9
Inlier distance threshold:	1.3	1.2	1.3	1.2	1.0	1.0	1.0
Standard TCs:	74	188	69	934	114	33	28
m2m TCs:	1131	480	158	2277	262	83	61

Filenames:	Mona.Lisa[AB].jpg	Notre.Dame[AB].jpg	Starbucks[AB].jpg	sym_arch_easy[AB].jpg	sym_arch[AB].jpg	sym_bank[AB].jpg	sym_bdom[AB].jpg
Dataset:	jav	jav	jav	sym	sym	sym	sym
Image A:							
Image dimensions:	400 × 600	450 × 600	640 × 399	603 × 583	760 × 457	439 × 600	800 × 534
Image B:							
Image dimensions:	624 × 380	640 × 425	800 × 533	603 × 583	760 × 457	439 × 600	800 × 534
Inlier error threshold:	0.9	0.9	1.0	0.9	1.5	0.9	1.6
Inlier distance threshold:	1.0	1.0	1.0	1.0	1.2	1.0	1.3
Standard TCs:	17	52	56	171	46	81	64
m2m TCs:	50	123	171	388	131	292	210

Filenames:	sym_cars[AB].jpg	sym_chinesebuilding[AB].jpg	sym_eiffel[AB].jpg	sym_essighaus[AB].jpg	sym_londonbridge[AB].jpg	sym_notredame12[AB].jpg
Dataset:	sym	sym	sym	sym	sym	sym
Image A:						
Image dimensions:	800 × 533	406 × 600	638 × 600	363 × 550	800 × 592	460 × 600
Image B:						
Image dimensions:	800 × 533	406 × 600	638 × 600	363 × 550	800 × 592	460 × 600
Inlier error threshold:	1.6	0.9	1.0	0.8	1.6	0.9
Inlier distance threshold:	1.3	1.0	1.0	0.9	1.3	1.0
Standard TCs:	259	95	6	50	63	137
m2m TCs:	647	297	14	126	184	467

Filenames:	sym_notredame14[AB].jpg	sym_notredame15[AB].jpg	sym_notredame16[AB].jpg	sym_notredame[AB].jpg	ubrus[AB].jpg	WhiteBoard[AB].jpg
Dataset:	sym	sym	sym	sym	own	lebeda
Image A:						
Image dimensions:	460 × 600	460 × 600	460 × 600	481 × 600	800 × 600	800 × 532
Image B:						
Image dimensions:	460 × 600	460 × 600	460 × 600	481 × 600	800 × 600	399 × 600
Inlier error threshold:	0.9	0.9	0.9	0.9	1.6	1.6
Inlier distance threshold:	1.0	1.0	1.0	1.0	1.3	1.3
Standard TCs:	61	153	134	203	163	54
m2m TCs:	312	602	502	666	9140	124









Image	Qty↓	m2m full	m2m v2	m2m v1	m2m switching	LO-RANSAC
	I	273.9 ±2.5 (266-281)	272.3 ±3.1 (266-280)	269.2 ±26.3 (14-280)	77.8 ±0.9 (76-80)	77.8 ±0.9 (76-80)
	I samp	76.7 ±1.5 (73-80)	76.5 ±1.6 (72-80)	75.6 ±7.3 (6-80)	77.8 ±0.9 (76-80)	77.8 ±0.9 (76-80)
	I (%)	85.1 ±0.8 (83-87)	84.6 ±1.0 (83-87)	83.6 ±8.2 (4-87)	24.2 ±0.3 (24-25)	24.2 ±0.3 (24-25)
	Samp	500000.0 ±0.0 (500000-500000)	948.7 ±78.3 (791-1209)	1032.6 ±467.1 (791-4953)	888.7 ±46.6 (791-1036)	888.7 ±46.6 (791-1036)
	Time _(ms)	50271.8 ±0.06 (0.9-1.2)	424.6 ±0.07 (0.8-1.2)	262.0 ±0.08 (0.8-1.4)	60.5 ±0.10 (0.9-1.3)	58.0 ±0.10 (0.9-1.3)
	Error	0.96 ±3.0 (3-17)	3.7 ±1.5 (1-8)	3.6 ±1.4 (1-7)	1.04 ±1.4 (1-7)	1.04 ±1.4 (1-7)
	LO count	9.4 ±3.6 (123-141)	3.7 ±1.5 (1-8)	3.6 ±1.4 (1-7)	3.6 ±1.4 (1-7)	3.6 ±1.4 (1-7)
	LO count	132.3 ±3.6 (36-50)	132.0 ±3.8 (121-141)	132.1 ±3.8 (120-141)	47.8 ±1.6 (44-51)	47.8 ±1.6 (44-51)
	I samp	45.0 ±2.3 (32-37)	45.0 ±2.4 (34-50)	45.5 ±2.2 (30-54)	47.8 ±1.6 (44-51)	47.8 ±1.6 (44-51)
	I (%)	34.3 ±0.9 (32-37)	34.2 ±1.0 (31-37)	34.2 ±1.0 (31-37)	12.4 ±0.4 (11-13)	12.4 ±0.4 (11-13)
	Samp	500000.0 ±0.0 (500000-500000)	16527.4 ±3030.1 (11754-24484)	15942.6 ±2929.8 (10916-24373)	13419.2 ±1904.3 (10076-18309)	13419.2 ±1904.3 (10076-18309)
	Time _(ms)	20027.0 ±0.71 (3.0-5.2)	1365.5 ±0.67 (3.0-5.4)	585.5 ±0.61 (3.0-5.4)	244.6 ±0.34 (3.1-5.2)	237.3 ±0.34 (3.1-5.2)
	Error	3.71 ±2.8 (3-17)	6.3 ±2.2 (2-15)	5.8 ±2.1 (2-17)	3.59 ±2.0 (2-17)	3.59 ±2.0 (2-17)
	LO count	9.1 ±2.6 (1936-1949)	6.3 ±2.2 (2-15)	5.8 ±2.1 (2-17)	5.7 ±2.0 (2-17)	5.7 ±2.0 (2-17)
	I	1943.3 ±2.7 (774-787)	1941.3 ±2.9 (769-783)	1889.1 ±282.3 (14-1946)	922.1 ±1.1 (920-925)	922.1 ±1.1 (920-925)
	I samp	922.8 ±1.2 (919-926)	921.8 ±1.4 (918-925)	897.1 ±134.3 (7-925)	922.1 ±1.1 (920-925)	922.1 ±1.1 (920-925)
	I (%)	146.6 ±0.2 (146-147)	146.4 ±0.2 (146-147)	142.5 ±21.3 (1-147)	69.5 ±0.1 (69-70)	69.5 ±0.1 (69-70)
	Samp	517.4 ±2.6 (510-526)	26.1 ±11.5 (12-51)	26.1 ±11.5 (12-51)	26.1 ±11.5 (12-51)	26.1 ±11.5 (12-51)
	Time _(ms)	569.0 ±0.03 (0.4-0.5)	105.7 ±0.04 (0.4-0.6)	0.50 ±0.13 (0.4-1.2)	40.1 ±0.04 (0.4-0.6)	39.3 ±0.04 (0.4-0.6)
	Error	0.47 ±1.3 (1-6)	0.47 ±0.08 (0.4-0.9)	0.50 ±0.13 (0.4-1.2)	0.48 ±0.04 (0.4-0.6)	0.48 ±0.04 (0.4-0.6)
	LO count	3.0 ±1.3 (1-6)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	I	779.6 ±2.7 (774-787)	776.6 ±2.9 (769-783)	647.8 ±278.4 (10-782)	358.0 ±1.4 (355-362)	358.0 ±1.4 (355-362)
	I samp	359.4 ±1.3 (357-363)	357.9 ±1.3 (355-361)	298.5 ±128.1 (5-360)	358.0 ±1.4 (355-362)	358.0 ±1.4 (355-362)
	I (%)	138.7 ±0.5 (138-140)	138.2 ±0.5 (137-139)	115.3 ±49.5 (2-139)	63.7 ±0.3 (63-64)	63.7 ±0.3 (63-64)
	Samp	914.8 ±13.5 (878-939)	40.3 ±10.4 (20-53)	40.8 ±11.3 (20-80)	40.3 ±10.4 (20-53)	40.3 ±10.4 (20-53)
	Time _(ms)	342.9 ±0.06 (0.5-0.7)	51.1 ±0.08 (0.4-0.9)	43.6 ±0.31 (0.4-1.9)	20.7 ±0.07 (0.4-0.8)	20.6 ±0.07 (0.4-0.8)
	Error	0.56 ±1.6 (1-8)	0.55 ±0.08 (0.4-0.9)	0.69 ±0.31 (0.4-1.9)	0.55 ±0.07 (0.4-0.8)	0.55 ±0.07 (0.4-0.8)
	LO count	4.1 ±1.6 (1-8)	1.0 ±0.0 (1-1)	1.0 ±0.2 (1-2)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	I	1693.0 ±4.6 (1680-1703)	1687.0 ±5.9 (1671-1704)	1454.7 ±563.0 (11-1703)	753.8 ±2.7 (748-759)	753.8 ±2.7 (748-759)
	I samp	756.2 ±2.3 (750-762)	753.1 ±3.0 (746-762)	649.5 ±251.2 (5-761)	753.8 ±2.7 (748-759)	753.8 ±2.7 (748-759)
	I (%)	136.6 ±0.4 (136-137)	136.2 ±0.5 (135-138)	117.4 ±45.4 (1-137)	60.8 ±0.2 (60-61)	60.8 ±0.2 (60-61)
	Samp	2022.4 ±24.0 (1961-2090)	43.5 ±9.1 (22-54)	45.5 ±13.1 (22-108)	43.5 ±9.1 (22-54)	43.5 ±9.1 (22-54)
	Time _(ms)	1302.7 ±0.04 (0.6-0.8)	127.3 ±0.04 (0.6-0.8)	109.1 ±0.19 (0.6-2.3)	40.6 ±0.04 (0.6-0.8)	40.0 ±0.04 (0.6-0.8)
	Error	0.67 ±2.1 (1-10)	0.68 ±0.04 (0.6-0.8)	0.72 ±0.3 (1-2)	0.68 ±0.04 (0.6-0.8)	0.68 ±0.04 (0.6-0.8)
	LO count	4.6 ±2.1 (1-10)	1.0 ±0.0 (1-1)	1.1 ±0.3 (1-2)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	I	1858.0 ±9.2 (1841-1881)	1847.2 ±9.4 (1826-1878)	1588.7 ±615.0 (9-1879)	874.2 ±4.5 (860-888)	874.2 ±4.5 (860-888)
	I samp	878.4 ±4.6 (869-890)	873.2 ±4.7 (864-889)	750.8 ±290.9 (5-889)	874.2 ±4.5 (860-888)	874.2 ±4.5 (860-888)
	I (%)	140.2 ±0.7 (139-142)	139.4 ±0.7 (138-142)	119.9 ±46.4 (1-142)	66.0 ±0.3 (65-67)	66.0 ±0.3 (65-67)
	Samp	657.1 ±13.4 (623-686)	39.9 ±9.9 (18-53)	40.4 ±10.7 (18-67)	39.9 ±9.9 (18-53)	39.9 ±9.9 (18-53)
	Time _(ms)	633.7 ±0.59 ±0.02 (0.5-0.6)	115.5 ±0.03 (0.5-0.7)	97.9 ±0.16 (0.5-1.3)	43.6 ±0.03 (0.5-0.7)	42.6 ±0.03 (0.5-0.7)
	Error	0.59 ±1.6 (1-8)	0.60 ±0.04 (0.5-0.7)	0.67 ±0.2 (1-2)	0.59 ±0.03 (0.5-0.7)	0.59 ±0.03 (0.5-0.7)
	LO count	3.3 ±1.6 (1-8)	1.0 ±0.0 (1-1)	1.0 ±0.2 (1-2)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	I	1047.1 ±2.1 (1042-1051)	1044.2 ±3.0 (1035-1052)	912.4 ±335.3 (10-1051)	387.9 ±0.9 (386-390)	387.9 ±0.9 (386-390)
	I samp	388.2 ±1.0 (386-391)	387.1 ±1.5 (383-391)	338.9 ±124.0 (5-391)	387.9 ±0.9 (386-390)	387.9 ±0.9 (386-390)
	I (%)	141.5 ±0.3 (141-142)	141.1 ±0.4 (140-142)	123.3 ±45.3 (1-142)	52.4 ±0.1 (52-53)	52.4 ±0.1 (52-53)
	Samp	78061.9 ±775.4 (76011-80032)	50.8 ±2.2 (43-56)	62.3 ±24.7 (43-237)	50.8 ±2.2 (43-56)	50.8 ±2.2 (43-56)
	Time _(ms)	18833.7 ±0.05 (0.6-0.8)	145.7 ±0.05 (0.6-0.8)	0.78 ±0.30 (0.6-2.5)	26.2 ±0.07 (0.6-0.9)	25.1 ±0.07 (0.6-0.9)
	Error	0.70 ±2.4 (4-15)	0.67 ±0.08 (0.6-0.8)	0.78 ±0.30 (0.6-2.5)	0.68 ±0.07 (0.6-0.9)	0.68 ±0.07 (0.6-0.9)
	LO count	8.0 ±2.4 (4-15)	1.0 ±0.0 (1-1)	1.3 ±0.5 (1-2)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	LO count	8.0 ±2.4 (4-15)	1.0 ±0.0 (1-1)	1.3 ±0.5 (1-2)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)















Image	Qty↓	m2m full	m2m v2	m2m v1	m2m switching	LO-RANSAC
08		2145.2 ±2.5 (2139-2152)	2142.4 ±3.3 (2136-2153)	1950.1 ±612.3 (9-2149)	963.5 ±1.4 (961-968)	963.5 ±1.4 (961-968)
		964.5 ±1.3 (962-968) 172.0 ±0.2 (172-173) 998.6 ±5.0 (985-1009) 896.4 ±0.4 (0.4-0.6) 0.48 ±1.9 (1-10) 4.2 ±1.9 (1-10)	963.0 ±1.6 (960-968) 171.8 ±0.3 (171-173) 15.8 ±7.0 (7-38) 118.1 ±0.04 (0.4-0.6) 0.47 ±0.04 (1-1) 1.0 ±0.0 (1-1)	876.7 ±275.2 (4-967) 156.4 ±49.1 (1-172) 15.8 ±6.9 (7-38) 106.9 ±0.12 (0.4-1.1) 0.50 ±0.0 (1-1) 1.0 ±0.0 (1-1)	963.5 ±1.4 (961-968) 963.5 ±0.1 (77-78) 77.3 ±6.9 (7-38) 15.8 ±6.9 (7-38) 34.8 ±0.04 (0.4-0.6) 0.49 ±0.0 (1-1) 1.0 ±0.0 (1-1)	963.5 ±1.4 (961-968) 963.5 ±0.1 (77-78) 77.3 ±6.9 (7-38) 15.8 ±6.9 (7-38) 34.1 ±0.04 (0.4-0.6) 0.49 ±0.0 (1-1) 1.0 ±0.0 (1-1)
09		1853.1 ±2.2 (1847-1857)	1851.1 ±2.7 (1844-1857)	1614.2 ±604.1 (10-1857)	829.2 ±1.3 (827-832)	829.2 ±1.3 (827-832)
		830.0 ±1.2 (827-832) 166.9 ±0.2 (166-167) 960.0 ±5.6 (951-974) 639.7 ±0.03 (0.5-0.6) 0.51 ±1.7 (1-9) 3.7 ±1.7 (1-9)	828.9 ±1.3 (826-832) 166.8 ±0.2 (166-167) 15.7 ±7.2 (8-43) 97.0 ±0.02 (0.5-0.6) 0.50 ±0.0 (1-1) 1.0 ±0.0 (1-1)	722.6 ±270.5 (5-832) 145.4 ±54.4 (1-167) 15.7 ±7.2 (8-43) 87.1 ±0.20 (0.5-1.9) 0.56 ±0.0 (1-1) 1.0 ±0.0 (1-1)	829.2 ±1.3 (827-832) 829.2 ±0.1 (75-75) 74.7 ±7.2 (8-43) 15.7 ±7.2 (8-43) 30.1 ±0.02 (0.5-0.6) 0.50 ±0.0 (1-1) 1.0 ±0.0 (1-1)	829.2 ±1.3 (827-832) 829.2 ±0.1 (75-75) 74.7 ±7.2 (8-43) 15.7 ±7.2 (8-43) 30.1 ±0.02 (0.5-0.6) 0.50 ±0.0 (1-1) 1.0 ±0.0 (1-1)
10		1136.5 ±2.7 (1131-1144)	1133.3 ±3.6 (1124-1143)	983.0 ±377.3 (6-1141)	513.4 ±1.3 (510-517)	513.4 ±1.3 (510-517)
		514.6 ±1.4 (512-518) 151.9 ±0.4 (151-153) 779.0 ±8.2 (759-796) 427.8 ±0.04 (0.6-0.8) 0.73 ±1.4 (1-7) 3.6 ±1.4 (1-7)	513.2 ±1.9 (508-518) 151.5 ±0.5 (150-153) 29.0 ±10.0 (14-53) 65.9 ±0.04 (0.6-0.8) 0.73 ±0.0 (1-1) 1.0 ±0.0 (1-1)	445.1 ±170.7 (4-516) 131.4 ±50.4 (1-153) 29.0 ±10.0 (14-53) 57.0 ±0.18 (0.6-2.0) 0.78 ±0.0 (1-1) 1.0 ±0.0 (1-1)	513.4 ±1.3 (510-517) 513.4 ±0.2 (68-69) 68.6 ±10.0 (14-53) 29.0 ±10.0 (14-53) 24.9 ±0.04 (0.7-0.8) 0.72 ±0.0 (1-1) 1.0 ±0.0 (1-1)	513.4 ±1.3 (510-517) 513.4 ±0.2 (68-69) 68.6 ±10.0 (14-53) 29.0 ±10.0 (14-53) 24.9 ±0.04 (0.7-0.8) 0.72 ±0.0 (1-1) 1.0 ±0.0 (1-1)
11		578.6 ±3.7 (570-589)	575.1 ±5.3 (553-588)	568.1 ±54.7 (30-584)	228.6 ±2.0 (225-234)	228.6 ±2.0 (225-234)
		239.0 ±2.1 (224-235) 78.9 ±0.5 (78-80) 114270.4 ±4026.1 (103171-125032) 20252.4 ±0.10 (18.6-19.1) 18.80 ±2.3 (3-13) 8.4 ±2.3 (3-13)	227.6 ±2.3 (220-235) 78.5 ±0.7 (75-80) 321.9 ±13.2 (283-369) 327.9 ±0.10 (18.5-19.0) 18.80 ±1.3 (1-8) 2.9 ±1.3 (1-8)	224.6 ±21.7 (12-232) 77.5 ±7.5 (4-80) 354.8 ±115.3 (298-1089) 245.7 ±0.12 (18.1-19.1) 18.80 ±1.1 (1-6) 2.9 ±1.1 (1-6)	228.6 ±2.0 (225-234) 31.2 ±0.3 (31-32) 316.2 ±10.9 (288-337) 82.2 ±0.11 (18.6-19.1) 18.81 ±1.2 (1-7) 2.8 ±1.2 (1-7)	228.6 ±2.0 (225-234) 31.2 ±0.3 (31-32) 316.2 ±10.9 (288-337) 80.5 ±0.11 (18.6-19.1) 18.81 ±1.2 (1-7) 2.8 ±1.2 (1-7)
adam		390.1 ±3.9 (377-399)	387.5 ±3.9 (379-398)	381.6 ±38.2 (15-400)	189.7 ±2.0 (186-195)	189.7 ±2.0 (186-195)
		190.2 ±1.9 (184-194) 151.8 ±1.5 (147-155) 205.8 ±8.5 (190-235) 79.0 ±0.36 (1.1-2.7) 1.81 ±1.2 (1-8) 2.4 ±1.2 (1-8)	188.9 ±1.9 (184-194) 150.8 ±1.5 (147-155) 30.9 ±9.3 (13-50) 25.9 ±0.38 (1.1-2.6) 1.79 ±0.0 (1-1) 1.0 ±0.0 (1-1)	186.0 ±18.7 (7-195) 148.5 ±14.9 (6-156) 30.9 ±9.3 (13-50) 22.9 ±0.42 (0.7-3.5) 1.75 ±0.0 (1-1) 1.0 ±0.0 (1-1)	189.7 ±2.0 (186-195) 73.8 ±0.8 (72-76) 30.9 ±9.3 (13-50) 13.0 ±0.39 (0.9-2.6) 1.82 ±0.0 (1-1) 1.0 ±0.0 (1-1)	189.7 ±2.0 (186-195) 73.8 ±0.8 (72-76) 30.9 ±9.3 (13-50) 12.5 ±0.39 (0.9-2.6) 1.82 ±0.0 (1-1) 1.0 ±0.0 (1-1)
boat		561.6 ±3.8 (554-573)	559.4 ±4.2 (552-573)	540.5 ±93.7 (12-573)	278.2 ±2.1 (275-284)	278.2 ±2.1 (275-284)
		279.2 ±1.8 (275-284) 147.4 ±1.0 (145-150) 251.8 ±6.7 (235-267) 114.7 ±0.07 (1.1-1.5) 1.32 ±1.2 (1-6) 2.8 ±1.2 (1-6)	278.1 ±2.0 (275-285) 146.8 ±1.1 (145-150) 28.3 ±10.4 (11-53) 30.1 ±0.09 (1.1-1.5) 1.31 ±0.0 (1-1) 1.0 ±0.0 (1-1)	268.8 ±46.6 (6-285) 141.9 ±24.6 (3-150) 27.3 ±10.5 (11-54) 28.7 ±0.13 (1.1-2.1) 1.36 ±0.0 (1-1) 1.0 ±0.0 (1-1)	278.2 ±2.1 (275-284) 278.2 ±0.5 (72-75) 73.0 ±10.4 (11-53) 28.3 ±10.4 (11-53) 14.9 ±0.07 (1.2-1.5) 1.32 ±0.0 (1-1) 1.0 ±0.0 (1-1)	278.2 ±2.1 (275-284) 278.2 ±0.5 (72-75) 73.0 ±10.4 (11-53) 28.3 ±10.4 (11-53) 14.9 ±0.07 (1.2-1.5) 1.32 ±0.0 (1-1) 1.0 ±0.0 (1-1)
Boston		581.3 ±1.4 (579-584)	581.2 ±1.4 (578-584)	562.7 ±97.9 (8-584)	280.8 ±0.5 (280-282)	280.8 ±0.5 (280-282)
		280.9 ±0.3 (280-281) 181.1 ±0.4 (180-182) 181.2 ±0.7 (181-184) 79.5 ±0.04 (0.5-0.7) 0.54 ±1.1 (1-6) 2.4 ±1.1 (1-6)	280.7 ±0.5 (280-282) 175.3 ±0.4 (180-182) 9.8 ±5.7 (4-32) 26.9 ±0.05 (0.5-0.7) 0.55 ±0.0 (1-1) 1.0 ±0.0 (1-1)	271.9 ±30.5 (2-182) 175.3 ±30.5 (2-182) 9.8 ±5.7 (4-32) 25.6 ±0.14 (0.5-1.7) 0.57 ±0.0 (1-1) 1.0 ±0.0 (1-1)	280.8 ±0.5 (280-282) 280.8 ±0.1 (87-88) 9.8 ±5.7 (4-32) 12.5 ±0.05 (0.5-0.7) 0.54 ±0.0 (1-1) 1.0 ±0.0 (1-1)	280.8 ±0.5 (280-282) 280.8 ±0.1 (87-88) 9.8 ±5.7 (4-32) 12.5 ±0.05 (0.5-0.7) 0.54 ±0.0 (1-1) 1.0 ±0.0 (1-1)

Image	Qty↓	m2m full	m2m v2	m2m v1	m2m switching	LO-RANSAC
BostonLib	I	214.8 ±167.2 (198-1870)	198.1 ±0.2 (198-199)	211.1 ±169.6 (8-1870)	95.0 ±0.0 (95-95)	95.0 ±0.0 (95-95)
	I samp	97.4 ±23.9 (95-334)	95.0 ±0.0 (95-95)	95.6 ±27.1 (4-334)	95.0 ±0.0 (95-95)	95.0 ±0.0 (95-95)
BruggeSquare	I (%)	64.3 ±50.1 (59-560)	59.3 ±0.1 (59-60)	63.2 ±50.8 (2-560)	28.4 ±0.0 (28-28)	28.4 ±0.0 (28-28)
	Samp	453452.5±45505.2 (2951-458003)	461.2 ±28.0 (458-736)	465.9 ±67.9 (104-916)	461.2 ±28.0 (458-736)	461.2 ±28.0 (458-736)
	Time(ms)	11915.8	104.3	93.4	25.3	24.6
	Error	1.00 ±0.15 (0.0-1.4)	1.00 ±0.11 (0.8-1.4)	1.00 ±0.16 (0.0-1.5)	1.00 ±0.09 (0.8-1.3)	1.00 ±0.09 (0.8-1.3)
	LO count	8.5 ±2.4 (3-15)	2.3 ±1.0 (1-5)	2.3 ±0.9 (1-5)	2.2 ±0.9 (1-5)	2.2 ±0.9 (1-5)
BruggeTower	I	280.9 ±2.3 (262-284)	272.8 ±11.6 (241-284)	264.7 ±45.4 (12-284)	116.0 ±3.4 (102-120)	116.0 ±3.4 (102-120)
	I samp	117.9 ±1.1 (112-120)	115.3 ±3.9 (100-120)	112.0 ±18.8 (5-119)	116.0 ±3.4 (102-120)	116.0 ±3.4 (102-120)
	I (%)	83.8 ±0.7 (78-85)	81.4 ±3.5 (72-85)	79.0 ±13.6 (4-85)	34.6 ±1.0 (30-36)	34.6 ±1.0 (30-36)
	Samp	32698.8 ±1244.2 (30517-40265)	214.5 ±32.1 (181-377)	227.2 ±61.8 (187-650)	209.9 ±28.7 (181-348)	209.9 ±28.7 (181-348)
	Time(ms)	3232.2	111.2	91.4	38.4	37.3
Brussels	Error	2.82 ±0.13 (2.6-3.5)	3.23 ±1.18 (2.4-7.8)	3.30 ±1.46 (2.1-12.4)	3.64 ±1.52 (2.6-8.0)	3.64 ±1.52 (2.6-8.0)
	LO count	7.4 ±2.4 (3-15)	2.6 ±1.2 (1-6)	2.7 ±1.2 (1-6)	2.6 ±1.2 (1-6)	2.6 ±1.2 (1-6)
	I	441.4 ±4.5 (431-450)	436.3 ±6.9 (416-452)	432.2 ±14.0 (321-448)	214.7 ±3.7 (204-224)	214.7 ±3.7 (204-224)
	I samp	217.0 ±2.2 (212-221)	214.6 ±3.2 (205-222)	212.5 ±6.9 (158-220)	214.7 ±3.7 (204-224)	214.7 ±3.7 (204-224)
	I (%)	121.9 ±1.3 (119-124)	120.5 ±1.9 (115-125)	119.4 ±3.9 (89-124)	59.3 ±1.0 (56-62)	59.3 ±1.0 (56-62)
CapitalRegion	Samp	524.0 ±21.7 (486-575)	48.2 ±4.5 (33-54)	48.3 ±4.6 (33-60)	48.2 ±4.5 (33-54)	48.2 ±4.5 (33-54)
	Time(ms)	151.3	31.4	27.7	16.3	15.8
	Error	1.68 ±0.82 (1.2-3.6)	2.17 ±1.02 (1.2-4.8)	2.02 ±0.97 (1.2-5.9)	2.12 ±0.99 (1.2-3.9)	2.12 ±0.99 (1.2-3.9)
	LO count	3.4 ±1.4 (1-7)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	I	789.8 ±3.6 (782-798)	788.0 ±3.8 (776-797)	739.9 ±172.6 (7-801)	378.4 ±1.8 (374-383)	378.4 ±1.8 (374-383)
Chal.cab	I samp	379.2 ±1.8 (375-383)	378.3 ±1.9 (372-382)	355.2 ±82.9 (4-385)	378.4 ±1.8 (374-383)	378.4 ±1.8 (374-383)
	I (%)	153.1 ±0.7 (152-155)	152.7 ±0.7 (150-154)	143.4 ±33.4 (1-155)	73.3 ±0.4 (72-74)	73.3 ±0.4 (72-74)
	Samp	372.3 ±6.8 (358-390)	32.8 ±12.2 (12-52)	33.2 ±13.3 (12-89)	32.8 ±12.2 (12-52)	32.8 ±12.2 (12-52)
	Time(ms)	194.1	45.1	39.2	19.2	19.2
	Error	1.31 ±0.04 (1.2-1.4)	1.31 ±0.04 (1.2-1.4)	1.34 ±0.16 (0.9-2.6)	1.31 ±0.05 (1.2-1.4)	1.31 ±0.05 (1.2-1.4)
Chal.captole	LO count	3.0 ±1.4 (1-7)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	I	29.8 ±0.5 (28-31)	28.5 ±2.1 (21-30)	26.8 ±4.4 (9-30)	27.3 ±4.5 (13-31)	12.3 ±1.5 (8-14)
	I samp	12.4 ±0.7 (12-14)	12.3 ±1.1 (9-14)	11.7 ±2.0 (4-14)	12.5 ±1.4 (7-14)	12.3 ±1.5 (8-14)
	I (%)	96.3 ±1.7 (90-100)	91.8 ±6.7 (68-97)	86.6 ±14.3 (29-97)	88.1 ±14.6 (42-100)	39.8 ±4.7 (26-45)
	Samp	5565.5 ±1271.6 (3470-6626)	138.8 ±66.6 (68-448)	234.2 ±183.0 (68-747)	158.8 ±53.7 (68-362)	208.0 ±179.0 (68-747)
Chal.captole	Time(ms)	112.3	11.5	12.3	21.2	7.7
	Error	2.33 ±0.69 (1.8-4.8)	3.39 ±1.64 (1.9-8.0)	5.33 ±11.00 (1.9-106.1)	3.71 ±2.12 (1.9-14.2)	4.25 ±1.67 (1.7-12.1)
	LO count	4.7 ±1.6 (1-8)	1.6 ±0.7 (1-4)	1.6 ±0.7 (1-3)	4.1 ±1.3 (1-7)	1.7 ±0.7 (1-3)
	I	22.4 ±3.7 (17-43)	24.0 ±2.3 (20-31)	22.0 ±4.2 (11-33)	23.8 ±2.2 (20-30)	10.8 ±0.4 (10-11)
	I samp	6.3 ±1.5 (2-10)	9.1 ±1.3 (5-11)	9.4 ±1.4 (5-11)	9.1 ±1.2 (6-11)	10.8 ±0.4 (10-11)
Chal.captole	I (%)	10.8 ±1.8 (8-21)	11.6 ±1.1 (10-15)	10.6 ±2.0 (5-16)	11.5 ±1.0 (10-14)	5.2 ±0.2 (5-5)
	Samp	50000.0 ±0.0 (50000-50000)	487983.4 ±21499.3 (449684-500000)	486419.8 ±22333.9 (449684-500000)	494413.9 ±49820.4 (449684-612893)	462677.8 ±21877.5 (449684-500000)
	Time(ms)	9936.2	10124.3	4168.2	8246.5	3654.3
	Error	228.67 ±187.83 (59.5-1759.7)	198.15 ±99.40 (39.8-627.9)	178.70 ±65.88 (27.9-308.4)	215.12 ±90.68 (39.9-627.9)	158.90 ±57.94 (40.4-269.0)
	LO count	6.7 ±1.5 (3-10)	6.4 ±1.7 (3-10)	4.9 ±1.1 (2-7)	11.6 ±1.8 (8-16)	4.9 ±1.1 (2-7)
Chal.captole	I	47.5 ±1.2 (40-49)	45.4 ±3.4 (26-49)	45.0 ±4.4 (24-49)	18.6 ±8.9 (14-48)	15.3 ±1.4 (12-17)
	I samp	16.5 ±0.6 (14-17)	16.1 ±1.0 (12-17)	16.0 ±1.2 (11-17)	15.5 ±1.3 (11-17)	15.3 ±1.4 (12-17)
	I (%)	35.4 ±0.9 (30-37)	33.9 ±2.5 (19-37)	33.6 ±3.3 (18-37)	13.9 ±6.6 (10-36)	11.4 ±1.0 (9-13)
	Samp	50000.0 ±0.0 (50000-50000)	17642.3 ±6454.1 (12570-49937)	19792.1 ±11431.0 (12570-77711)	24106.7 ±11366.8 (12570-65350)	23384.3 ±9881.0 (12570-53800)
	Time(ms)	11604.6	438.5	270.2	267.9	212.3
Chal.captole	Error	530.37 ±9.69 (509.4-547.3)	523.36 ±16.04 (402.8-548.0)	516.96 ±49.60 (94.4-548.0)	490.52 ±24.14 (445.8-551.7)	490.52 ±24.14 (445.8-551.7)
	LO count	8.2 ±2.1 (4-14)	5.5 ±1.7 (2-10)	4.7 ±1.2 (2-8)	5.5 ±2.6 (2-16)	4.8 ±1.2 (2-8)



















Image	Qty↓	m2m full	m2m v2	m2m v1	m2m switching	LO-RANSAC							
 chal-dagstuh11	I	16.0	±1.4	15.2	±1.6	15.3	±1.4	15.3	±1.4	(12-16)	6.8	±0.4	(6-7)
	I samp	6.8	±0.4	6.8	±0.5	6.8	±0.4	6.8	±0.4	(5-7)	6.8	±0.4	(6-7)
 chal-dagstuh12	I (%)	51.6	±0.0	49.4	±4.6	49.0	±5.2	49.0	±5.2	(32-52)	21.8	±1.4	(19-23)
	Samp	500000.0	±0.0	1826.7	±621.5	1800.3	±586.1	1800.3	±586.1	(1346-4076)	1843.6	±568.4	(1346-2692)
 chal-dagstuh13	Time _(ms)	4131.9	±16.30	27.7	±32.72	27.50	±33.22	27.50	±33.22	(4.9-171.8)	14.0	±33.75	(4.9-171.8)
	Error	24.32	±1.5	2.8	±0.9	2.4	±0.7	2.4	±0.7	(1-5)	27.27	±0.7	(1-4)
 chal-dagstuh14	LO count	5.4	±1.1	15.2	±1.0	14.2	±1.3	14.2	±1.3	(14-18)	7.2	±0.4	(7-8)
	I	6.5	±0.8	6.9	±0.7	6.9	±0.6	6.9	±0.6	(6-9)	7.0	±0.4	(7-8)
 chal-trontheim	I (%)	18.2	±1.3	18.1	±1.2	16.9	±1.6	16.9	±1.6	(12-21)	8.6	±0.5	(8-10)
	Samp	500000.0	±0.0	83245.1	±25361.9	94597.3	±50076.0	94597.3	±50076.0	(27524-165150)	75088.9	±14240.1	(45874-82574)
 city	Time _(ms)	5835.6	±43.83	1051.4	±51.19	789.6	±57.88	789.6	±57.88	(69.5-302.0)	600.7	±47.01	(79.3-299.6)
	Error	135.05	±1.5	4.4	±1.2	2.9	±0.7	2.9	±0.7	(1-4)	166.58	±0.7	(1-4)
 dlazdice	LO count	5.7	±1.4	70.3	±1.6	68.9	±2.8	68.9	±2.8	(56-73)	25.1	±1.7	(21-28)
	I	22.4	±1.0	22.7	±1.5	23.1	±2.3	23.1	±2.3	(19-28)	25.1	±1.7	(21-28)
 city	I (%)	27.6	±0.6	27.4	±0.6	26.8	±1.1	26.8	±1.1	(22-28)	9.8	±0.7	(8-11)
	Samp	500000.0	±0.0	50110.2	±12419.5	49803.4	±17851.5	49803.4	±17851.5	(22394-109784)	37205.3	±10530.5	(22394-72714)
 city	Time _(ms)	11958.7	±1.24	1714.6	±4.0	734.7	±3.27	734.7	±3.27	(3.3-22.3)	435.8	±4.10	(3.1-15.0)
	Error	5.97	±2.6	7.1	±2.6	6.1	±1.2	6.1	±1.2	(1-12)	8.82	±2.1	(1-12)
 city	LO count	8.6	±1.6	4.0	±1.5	3.2	±1.0	3.2	±1.0	(1-5)	5.9	±0.9	(1-5)
	I	18.9	±0.7	17.7	±0.9	16.9	±2.8	16.9	±2.8	(10-20)	17.4	±1.9	(14-20)
 city	I (%)	9.0	±0.7	8.7	±0.9	8.4	±1.3	8.4	±1.3	(5-10)	8.6	±1.0	(7-10)
	Samp	31.9	±2.2	29.9	±2.9	28.7	±4.7	28.7	±4.7	(17-34)	29.5	±3.2	(24-34)
 city	Time _(ms)	500000.0	±0.0	9316.2	±4104.3	12034.3	±10902.9	12034.3	±10902.9	(4131-90895)	8696.1	±3804.6	(4131-19477)
	Error	4215.3	±26.43	110.9	±27.70	77.07	±42.81	77.07	±42.81	(10.0-311.8)	124.1	±28.04	(10.0-100.5)
 city	LO count	62.23	±1.6	4.0	±1.5	3.2	±1.0	3.2	±1.0	(1-4)	7.5	±1.6	(1-4)
	I	60.9	±1.1	60.1	±1.5	54.1	±15.8	54.1	±15.8	(6-62)	23.9	±0.4	(20-24)
 city	I (%)	24.0	±0.1	23.9	±0.3	21.7	±5.9	21.7	±5.9	(4-24)	23.9	±0.4	(20-24)
	Samp	99.9	±1.8	98.6	±2.4	88.7	±25.9	88.7	±25.9	(10-102)	39.2	±0.7	(33-39)
 city	Time _(ms)	39550.2	±752.0	125.7	±8.7	149.2	±66.1	149.2	±66.1	(123-636)	126.9	±18.0	(123-260)
	Error	2.83	±0.85	18.9	±0.75	17.2	±1.22	17.2	±1.22	(2.0-11.7)	10.1	±0.69	(2.0-5.8)
 city	LO count	6.8	±2.4	1.7	±0.8	1.7	±0.7	1.7	±0.7	(1-4)	2.48	±0.8	(1-4)
	I	51.4	±0.8	51.0	±0.9	40.5	±16.7	40.5	±16.7	(6-53)	24.1	±0.7	(20-25)
 city	I (%)	25.2	±0.4	25.0	±0.5	19.9	±8.1	19.9	±8.1	(4-26)	24.1	±0.7	(20-25)
	Samp	128.5	±1.9	127.6	±2.2	101.2	±41.8	101.2	±41.8	(15-132)	60.3	±1.7	(50-62)
 city	Time _(ms)	390.7	±25.6	47.0	±6.7	52.0	±14.8	52.0	±14.8	(21-113)	47.0	±6.7	(21-54)
	Error	0.50	±0.05	0.50	±0.09	0.76	±0.88	0.76	±0.88	(0.4-1.1)	6.9	±0.09	(0.3-1.2)
 city	LO count	3.0	±1.3	1.0	±0.0	1.1	±0.3	1.1	±0.3	(1-2)	1.0	±0.0	(1-1)
	I	173.1	±7.1	170.3	±11.8	170.8	±11.5	170.8	±11.5	(119-176)	48.2	±2.3	(41-49)
 city	I (%)	48.7	±1.6	48.1	±2.5	48.3	±2.3	48.3	±2.3	(41-49)	48.2	±2.3	(41-49)
	Samp	33.7	±1.4	33.2	±2.3	33.3	±2.3	33.3	±2.3	(23-34)	9.4	±0.4	(8-10)
 city	Time _(ms)	500000.0	±0.0	43532.2	±13423.7	42507.9	±12297.8	42507.9	±12297.8	(37099-83788)	42622.8	±12245.4	(37099-83788)
	Error	13020.5	±37.39	1239.7	±64.81	698.1	±64.93	698.1	±64.93	(202.6-530.3)	445.0	±87.06	(202.6-530.8)
 city	LO count	210.48	±2.2	225.07	±1.7	222.53	±1.4	222.53	±1.4	(2-9)	231.38	±1.4	(2-9)
	I	8.4	±1.7	5.9	±1.7	5.5	±1.4	5.5	±1.4	(2-9)	5.5	±1.4	(2-9)











Image	Qty↓	m2m full	m2m v2	m2m v1	m2m switching	LO-RANSAC
	I	20.4 ±7.4 (12-34)	31.6 ±0.7 (27-33)	31.2 ±2.4 (15-33)	31.5 ±1.9 (14-33)	9.8 ±0.7 (7-10)
	I samp	5.9 ±2.5 (2-10)	9.5 ±0.5 (8-10)	9.4 ±0.7 (8-10)	9.5 ±0.7 (8-10)	9.8 ±0.7 (7-10)
	I (%)	27.6 ±10.0 (16-46)	42.8 ±1.0 (36-45)	42.2 ±3.2 (20-45)	42.6 ±2.6 (19-45)	13.2 ±1.0 (9-14)
	Samp	50000.0 ±0.0 (50000-50000)	15789.2 ±9375.4 (10444-91911)	17713.6 ±11392.9 (10444-98484)	16262.8 ±5249.8 (10444-40952)	19391.2 ±12507.6 (10444-58880)
	Time _(ms)	10930.2 ±91.19 (2.4-364.4)	525.8 ±0.10 (2.4-3.2)	220.6 ±6.56 (2.4-68.1)	523.7 ±17.54 (2.4-178.0)	154.1 ±3.13 (2.6-32.9)
	Error	82.74 ±1.5 (3-10)	4.7 ±1.2 (1-8)	3.32 ±0.6 (2-5)	4.37 ±1.6 (4-13)	3.12 ±0.6 (1-4)
	LO count	6.2 ±1.0 (210-214)	211.1 ±1.2 (209-213)	165.6 ±81.8 (8-213)	97.7 ±0.4 (97-98)	97.7 ±0.4 (97-98)
	I samp	98.0 ±0.2 (97-99)	97.6 ±0.5 (97-98)	76.6 ±37.6 (4-98)	97.7 ±0.4 (97-98)	97.7 ±0.4 (97-98)
	I (%)	113.0 ±0.5 (112-114)	112.3 ±0.6 (111-113)	88.1 ±43.5 (4-113)	52.0 ±0.2 (52-52)	52.0 ±0.2 (52-52)
	Samp	1738.5 ±14.4 (1668-1810)	55.4 ±12.4 (41-106)	57.3 ±14.5 (41-108)	55.4 ±12.4 (41-106)	55.4 ±12.4 (41-106)
	Time _(ms)	92.7 ±0.16 (0.7-1.9)	17.3 ±0.44 (0.6-1.9)	17.0 ±0.60 (0.6-4.1)	9.6 ±0.37 (0.6-1.8)	9.8 ±0.37 (0.6-1.8)
	Error	1.64 ±1.5 (1-8)	1.29 ±0.0 (1-1)	1.41 ±0.2 (1-2)	1.34 ±0.0 (1-1)	1.34 ±0.0 (1-1)
	LO count	3.9 ±1.5 (1-9)	1.0 ±0.0 (1-1)	1.1 ±0.2 (1-4)	1.0 ±0.0 (1-4)	1.0 ±0.0 (1-4)
	I	53.0 ±0.0 (53-53)	53.0 ±0.3 (51-53)	46.1 ±15.8 (7-53)	26.3 ±0.4 (26-27)	26.3 ±0.4 (26-27)
	I samp	27.0 ±0.0 (27-27)	27.0 ±0.1 (26-27)	23.6 ±8.0 (4-27)	26.3 ±0.4 (26-27)	26.3 ±0.4 (26-27)
	I (%)	76.8 ±0.0 (77-77)	76.8 ±0.4 (74-77)	66.9 ±22.9 (10-77)	38.1 ±0.6 (38-39)	38.1 ±0.6 (38-39)
	Samp	3656.0 ±0.0 (3656-3656)	126.9 ±9.1 (125-186)	147.2 ±46.2 (125-392)	142.5 ±11.2 (125-186)	142.5 ±11.2 (125-186)
	Time _(ms)	68.6 ±0.18 (1.4-2.4)	12.9 ±0.20 (1.4-2.7)	12.4 ±0.49 (1.4-4.2)	9.5 ±0.43 (1.4-4.2)	9.5 ±0.43 (1.4-4.2)
	Error	1.75 ±1.5 (1-9)	1.77 ±0.7 (1-5)	1.88 ±0.7 (1-4)	1.81 ±0.7 (1-4)	1.81 ±0.7 (1-4)
	LO count	3.5 ±1.3 (1-7)	1.0 ±0.0 (1-1)	1.1 ±0.3 (1-2)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	I	1181.7 ±6.0 (1167-1198)	1175.3 ±9.2 (1152-1200)	1148.0 ±90.6 (493-1189)	561.6 ±4.3 (554-572)	561.6 ±4.3 (554-572)
	I samp	563.1 ±3.3 (556-572)	559.8 ±4.9 (547-574)	546.5 ±43.7 (229-568)	561.6 ±4.3 (554-572)	561.6 ±4.3 (554-572)
	I (%)	126.5 ±0.6 (125-128)	125.8 ±1.0 (123-128)	122.9 ±9.7 (53-127)	60.1 ±0.5 (59-61)	60.1 ±0.5 (59-61)
	Samp	800.8 ±18.1 (752-842)	49.8 ±2.7 (28-54)	51.6 ±8.7 (28-116)	49.7 ±2.7 (28-54)	49.7 ±2.7 (28-54)
	Time _(ms)	556.2 ±0.07 (1.7-2.1)	88.1 ±0.08 (1.7-2.1)	75.5 ±0.80 (1.7-9.7)	35.4 ±0.07 (1.8-2.1)	34.8 ±0.07 (1.8-2.1)
	Error	1.94 ±1.3 (1-7)	1.93 ±0.0 (1-1)	2.04 ±0.3 (1-2)	1.92 ±0.0 (1-1)	1.92 ±0.0 (1-1)
	LO count	3.5 ±1.3 (1-7)	1.0 ±0.0 (1-1)	1.1 ±0.3 (1-2)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	I	59.8 ±2.4 (54-64)	58.0 ±3.5 (49-64)	55.5 ±9.0 (6-64)	28.1 ±1.2 (25-30)	28.1 ±1.2 (25-30)
	I samp	28.8 ±1.0 (26-30)	28.1 ±1.4 (24-30)	27.0 ±4.3 (4-30)	28.1 ±1.2 (25-30)	28.1 ±1.2 (25-30)
	I (%)	52.5 ±2.1 (47-56)	50.9 ±3.0 (43-56)	48.6 ±7.9 (5-56)	24.7 ±1.0 (22-26)	24.7 ±1.0 (22-26)
	Samp	21726.4 ±2906.4 (18266-28072)	864.5 ±208.3 (634-1579)	971.6 ±353.6 (634-2257)	857.9 ±156.9 (634-1336)	857.9 ±156.9 (634-1336)
	Time _(ms)	764.3 ±0.50 (1.5-4.5)	65.4 ±0.97 (1.6-5.3)	52.8 ±1.65 (1.5-15.3)	37.6 ±0.97 (1.5-4.6)	36.8 ±0.97 (1.5-4.6)
	Error	2.08 ±2.0 (2-11)	2.46 ±1.4 (1-7)	2.68 ±1.3 (1-7)	2.52 ±1.4 (1-7)	2.52 ±1.4 (1-7)
	LO count	6.3 ±2.0 (12-14)	3.7 ±1.4 (11-14)	3.5 ±1.3 (6-14)	3.4 ±1.4 (11-14)	3.4 ±1.4 (11-14)
	I	13.9 ±0.4 (6-7)	13.1 ±1.0 (5-7)	10.5 ±1.7 (4-7)	13.2 ±1.0 (6-6)	6.0 ±0.0 (6-6)
	I samp	6.8 ±0.4 (36-42)	6.5 ±0.5 (33-42)	5.8 ±5.1 (18-42)	6.5 ±0.6 (33-42)	6.0 ±0.0 (18-18)
	I (%)	42.2 ±1.1 (78642-345528)	39.7 ±3.0 (1750-8171)	31.7 ±5.1 (1750-8171)	39.9 ±3.0 (1840-8171)	18.2 ±0.0 (3501-3501)
	Samp	96160.6 ±39183.8 (78642-345528)	2873.9 ±1199.2 (1.5-168.6)	3915.5 ±1741.0 (1.7-180.7)	2946.0 ±1174.5 (1.5-168.6)	3501.0 ±0.0 (1.7-180.7)
	Time _(ms)	824.2 ±13.33 (1.7-132.7)	37.6 ±60.62 (1.6-6)	64.55 ±52.49 (1.3-3)	41.7 ±55.35 (3-8)	26.9 ±52.75 (1.7-180.7)
	Error	5.12 ±1.1 (17-19)	3.0 ±1.0 (8-10)	1.9 ±0.4 (4-10)	5.6 ±1.0 (15-19)	72.15 ±0.4 (1-3)
	LO count	4.3 ±1.0 (9-10)	17.4 ±0.6 (61-68)	8.2 ±10.8 (54-68)	17.5 ±0.6 (54-68)	1.9 ±0.4 (1-3)
	I	18.1 ±0.5 (4736-7443)	293.5 ±84.4 (185-486)	293.7 ±60.3 (185-486)	9.2 ±0.4 (8-10)	9.2 ±0.4 (8-10)
	I samp	9.6 ±3.6 (61-68)	62.1 ±3.6 (61-68)	54.4 ±10.8 (54-68)	9.2 ±0.6 (8-10)	9.2 ±0.4 (8-10)
	I (%)	5954.1 ±1353.5 (4736-7443)	293.5 ±84.4 (185-486)	293.7 ±60.3 (185-486)	62.4 ±10.8 (54-68)	32.8 ±1.4 (29-36)
	Samp	70.76 ±10.16 (59.5-89.7)	14.1 ±9.39 (59.5-99.0)	72.32 ±11.66 (13.4-107.5)	314.5 ±72.0 (230-541)	280.0 ±47.6 (185-486)
	Time _(ms)	93.2 ±1.1 (1-5)	2.0 ±0.9 (1-4)	1.8 ±0.7 (1-3)	21.7 ±8.88 (59.5-99.0)	6.0 ±8.09 (59.5-87.8)
	Error	3.0 ±1.1 (1-5)	2.0 ±0.9 (1-4)	1.8 ±0.7 (1-3)	72.97 ±1.2 (1-7)	72.67 ±0.7 (1-3)
LO count	3.0 ±1.1 (1-5)	2.0 ±0.9 (1-4)	1.8 ±0.7 (1-3)	4.0 ±1.2 (1-7)	1.8 ±0.7 (1-3)	








Image	Qty↓	m2m full	m2m v2	m2m v1	m2m switching	LO-RANSAC
 Mona Lisa	I	12.0 ±0.0 (12-12)	11.6 ±3.4 (9-43)	10.2 ±2.1 (5-12)	11.7 ±2.9 (9-38)	5.5 ±0.5 (5-6)
	I samp	6.0 ±0.0 (6-6)	5.8 ±0.9 (5-13)	5.4 ±0.7 (4-6)	5.8 ±0.8 (4-6)	5.5 ±0.5 (5-6)
	I (%)	70.6 ±0.0 (71-71)	68.2 ±19.7 (53-253)	59.8 ±12.5 (29-71)	68.5 ±16.9 (53-224)	32.4 ±3.0 (29-35)
	Samp	19711.0 ±0.0 (19711-19711)	308.0 ±121.1 (174-474)	335.2 ±125.9 (203-474)	329.1 ±124.1 (57-652)	359.0 ±124.5 (203-474)
	Time _(ms) Error LO count	154.5 ±0.03 (2.6-2.7) 2.67 ±1.2 (1-6) 3.5 ±1.0 (21-26)	NaN 1.9 ±0.8 (1-4) 21.69 ±32.65 (2.6-103.6)	7.5 ±2.2 (1.7-3.4) 1.4 ±0.6 (1-3) 21.8 ±2.3 (17-26)	NaN 2.8 ±1.0 (1-5) 22.1 ±2.7 (15-26)	3.9 ±33.56 (2.6-103.6) 23.71 ±0.6 (1-3) 1.5 ±0.7 (9-13)
 Notre Dame	I	25.4 ±1.0 (21-26)	22.5 ±2.3 (18-26)	21.8 ±2.3 (17-26)	22.1 ±2.7 (15-26)	11.0 ±0.7 (9-13)
	I samp	11.6 ±0.7 (9-12)	11.3 ±0.9 (9-13)	11.1 ±0.7 (9-13)	10.9 ±1.0 (8-12)	11.0 ±0.7 (9-13)
	I (%)	48.8 ±1.9 (40-50)	43.3 ±4.5 (35-50)	42.0 ±4.4 (33-50)	42.6 ±5.2 (29-50)	21.1 ±1.4 (17-25)
	Samp	49397.0 ±14690.6 (38037-93074)	1706.5 ±641.3 (809-3861)	1746.2 ±554.1 (809-3861)	1954.7 ±632.0 (809-4487)	1786.2 ±587.7 (809-3861)
	Time _(ms) Error LO count	499.5 ±11.35 (4.3-74.8) 7.21 ±1.5 (2-10) 5.4 ±1.4 (49-54)	20.33 ±20.64 (4.3-75.0) 3.5 ±1.1 (1-6) 50.6 ±2.3 (36-54)	31.8 ±22.75 (4.3-123.5) 3.1 ±1.0 (1-5) 48.4 ±8.7 (8-54)	48.5 ±25.35 (3.7-110.9) 7.1 ±1.6 (4-11) 23.9 ±0.7 (18-25)	22.1 ±0.7 (9-13) 25.40 ±22.63 (4.0-123.5) 3.1 ±1.0 (1-5)
 Starbucks	I	52.1 ±1.4 (49-54)	50.6 ±2.3 (36-54)	48.4 ±8.7 (8-54)	23.9 ±0.7 (18-25)	23.9 ±0.7 (18-25)
	I samp	25.6 ±0.8 (24-27)	24.9 ±1.1 (24-27)	23.8 ±4.2 (5-27)	23.9 ±0.7 (18-25)	23.9 ±0.7 (18-25)
	I (%)	93.0 ±2.6 (88-96)	90.3 ±4.2 (84-96)	86.4 ±15.6 (14-96)	42.7 ±1.3 (32-45)	42.7 ±1.3 (32-45)
	Samp	6342.1 ±807.3 (5031-8143)	99.4 ±49.7 (62-368)	131.8 ±93.0 (62-593)	106.2 ±42.2 (73-283)	106.2 ±42.2 (73-283)
	Time _(ms) Error LO count	82.2 ±0.06 (1.8-2.2) 1.89 ±1.7 (1-10) 2.8 ±1.2 (25-27)	11.3 ±0.17 (1.7-3.4) 1.2 ±0.4 (1-3) 1.0 ±0.2 (1-1)	12.0 ±0.24 (1.7-3.4) 1.92 ±0.5 (1-3) 1.4 ±0.2 (1-2)	7.3 ±0.16 (1.7-3.2) 1.87 ±0.4 (1-3) 1.2 ±0.4 (1-3)	7.1 ±0.16 (1.7-3.2) 1.87 ±0.08 (1.0-1.4) 1.0 ±0.0 (1-1)
 sym-arch	I	261.5 ±1.2 (257-264)	260.5 ±1.4 (257-264)	234.7 ±72.8 (8-264)	123.8 ±0.8 (122-126)	123.8 ±0.8 (122-126)
	I samp	124.2 ±0.6 (122-125)	123.6 ±0.7 (122-125)	111.5 ±34.4 (4-125)	123.8 ±0.8 (122-126)	123.8 ±0.8 (122-126)
	I (%)	152.9 ±0.7 (150-154)	152.3 ±0.8 (150-154)	137.3 ±42.6 (5-154)	72.4 ±0.4 (71-74)	72.4 ±0.4 (71-74)
	Samp	284.9 ±6.0 (277-306)	37.7 ±11.7 (12-54)	38.0 ±12.2 (12-69)	37.7 ±11.7 (12-54)	37.7 ±11.7 (12-54)
	Time _(ms) Error LO count	62.3 ±0.07 (1.1-1.5) 1.24 ±1.2 (1-6) 2.8 ±1.2 (25-27)	17.4 ±0.08 (1.0-1.4) 1.0 ±0.0 (1-1) 1.0 ±0.0 (1-1)	16.0 ±0.30 (1.0-3.8) 1.28 ±0.2 (1-2) 1.0 ±0.2 (1-2)	10.5 ±0.08 (1.0-1.4) 1.23 ±0.0 (1-1) 1.0 ±0.0 (1-1)	10.2 ±0.08 (1.0-1.4) 1.23 ±0.0 (1-1) 1.0 ±0.0 (1-1)
 sym-arch-easy	I	26.2 ±0.6 (25-27)	25.2 ±1.3 (21-27)	24.1 ±3.9 (7-27)	25.3 ±1.3 (20-27)	11.7 ±0.8 (10-13)
	I samp	11.8 ±1.0 (11-13)	11.7 ±1.0 (10-13)	11.3 ±1.8 (4-13)	11.7 ±1.0 (9-13)	11.7 ±0.8 (10-13)
	I (%)	56.9 ±1.3 (54-59)	54.9 ±2.9 (46-59)	52.3 ±8.4 (15-59)	55.1 ±2.8 (43-59)	25.4 ±1.7 (22-28)
	Samp	54837.1 ±16755.6 (35064-70908)	857.9 ±316.1 (487-1966)	885.8 ±308.3 (487-1480)	920.2 ±242.5 (487-1480)	828.8 ±265.2 (487-1480)
	Time _(ms) Error LO count	726.1 ±9.13 (4.4-38.4) 16.03 ±2.0 (2-11) 5.8 ±2.0 (41-45)	32.3 ±11.80 (4.0-73.2) 3.1 ±1.6 (1-6) 41.4 ±2.8 (31-45)	46.4 ±19.00 (3.9-161.1) 2.7 ±0.9 (1-5) 39.6 ±6.4 (8-45)	20.16 ±12.47 (2.2-77.6) 6.9 ±1.5 (4-11) 20.4 ±9.4 (14-44)	15.9 ±15.35 (3.1-73.2) 2.7 ±0.9 (1-5) 15.9 ±1.6 (11-19)
 sym-bank	I	42.8 ±0.9 (41-45)	41.4 ±2.8 (31-45)	39.6 ±6.4 (8-45)	20.4 ±9.4 (14-44)	15.9 ±1.6 (11-19)
	I samp	16.0 ±1.1 (14-19)	15.9 ±1.5 (12-18)	15.6 ±2.4 (4-18)	16.1 ±1.3 (10-19)	15.9 ±1.6 (11-19)
	I (%)	52.9 ±1.1 (51-56)	51.1 ±3.4 (38-56)	48.9 ±7.8 (10-56)	25.1 ±11.6 (17-54)	19.7 ±2.0 (14-23)
	Samp	357533.8 ±93769.7 (183471-500000)	2476.7 ±982.0 (1285-6970)	3184.0 ±1885.1 (1285-10068)	2850.2 ±1563.2 (1285-7305)	2915.1 ±1840.6 (1285-10068)
	Time _(ms) Error LO count	5159.2 ±0.33 (3.8-5.1) 4.45 ±1.9 (4-13) 7.8 ±1.4 (71-77)	4.72 ±1.32 (3.9-11.2) 4.0 ±1.5 (1-7) 70.5 ±3.3 (62-76)	5.25 ±3.95 (3.6-40.4) 3.4 ±1.2 (1-6) 62.2 ±18.1 (8-76)	4.92 ±2.09 (3.6-12.2) 4.3 ±2.5 (1-12) 30.1 ±1.1 (28-31)	6.20 ±7.52 (3.6-53.9) 3.4 ±1.2 (1-6) 30.1 ±1.1 (28-31)
 sym-bldom	I	74.2 ±1.4 (71-77)	70.5 ±3.3 (62-76)	62.2 ±18.1 (8-76)	30.1 ±1.1 (28-31)	30.1 ±1.1 (28-31)
	I samp	30.0 ±0.3 (29-31)	29.6 ±0.9 (27-31)	26.4 ±7.2 (4-31)	30.1 ±1.1 (28-31)	30.1 ±1.1 (28-31)
	I (%)	115.9 ±2.2 (111-120)	110.2 ±5.1 (97-119)	97.2 ±28.2 (12-119)	47.0 ±1.7 (44-48)	47.0 ±1.7 (44-48)
	Samp	7415.8 ±258.2 (6559-7496)	64.4 ±10.1 (52-92)	80.1 ±24.0 (54-156)	61.0 ±9.7 (52-79)	61.0 ±9.7 (52-79)
	Time _(ms) Error LO count	174.1 ±0.15 (1.8-2.4) 2.17 ±2.0 (2-11) 5.3 ±2.0 (41-45)	12.9 ±2.46 (1.7-11.3) 1.1 ±0.4 (1-3) 1.1 ±0.4 (1-3)	13.3 ±5.99 (1.7-52.4) 4.59 ±0.5 (1-3) 1.3 ±0.5 (1-3)	7.8 ±3.74 (1.7-11.6) 3.73 ±0.3 (1-2) 1.1 ±0.3 (1-2)	7.6 ±3.74 (1.7-11.6) 3.73 ±0.3 (1-2) 1.1 ±0.3 (1-2)



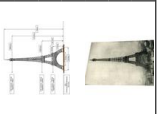



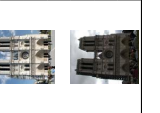












Image	Qty↓	m2m full	m2m v2	m2m v1	m2m switching	LO-RANSAC
 sym-cars	I	413.4 ±1.3	409-415	405.2 ±45.5	192.7 ±0.5	192.7 ±0.5
	I samp	192.8 ±0.5	(191-194)	189.2 ±21.2	192.7 ±0.5	192.7 ±0.5
	I (%)	159.6 ±0.5	(158-160)	156.5 ±17.6	74.4 ±0.2	74.4 ±0.2
	Samp	379.4 ±4.1	(370-394)	29.8 ±12.8	29.4 ±11.8	29.4 ±11.8
	Time _(ms)	105.9		23.1	12.2	12.4
	Error	2.66 ±0.03	(2.6-2.8)	2.66 ±0.15	±0.03	±0.03
	LO count	3.0 ±1.4	(1-7)	1.0 ±0.1	1.0 ±0.0	1.0 ±0.0
 sym-chinesebuilding	I	62.7 ±1.7	(58-66)	57.6 ±9.1	28.7 ±1.8	28.7 ±1.8
	I samp	29.4 ±1.0	(27-31)	27.3 ±4.1	28.7 ±1.8	28.7 ±1.8
	I (%)	66.0 ±1.8	(61-69)	60.6 ±9.6	30.2 ±1.9	30.2 ±1.9
	Samp	33011.7 ±4529.7	(26465-46481)	472.6 ±190.8	409.7 ±145.1	409.7 ±145.1
	Time _(ms)	793.1		35.1	21.9	21.9
	Error	6.10 ±1.36	(2.8-9.7)	6.78 ±4.96	±5.08	±5.08
	LO count	6.8 ±2.2	(2-13)	2.8 ±1.2	2.7 ±1.1	2.7 ±1.1
 sym-effeil	I	7.0 ±0.0	(7-7)	0.0 ±0.0	0.0 ±0.0	3.0 ±2.7
	I samp	2.0 ±0.0	(2-2)	0.0 ±0.0	0.0 ±0.0	3.0 ±2.7
	I (%)	116.7 ±0.0	(117-117)	0.0 ±0.0	0.0 ±0.0	50.5 ±45.4
	Samp	500000.0 ±0.0	(500000-500000)	500000.0 ±0.0	500000.0 ±0.0	500000.0 ±0.0
	Time _(ms)	3954.1		3723.9	3722.3	3661.7
	Error	94.01 ±0.00	(94.0-94.0)	NaN	NaN	NaN
	LO count	1.0 ±0.0	(1-1)	0.0 ±0.0	0.0 ±0.0	0.0 ±0.0
 sym-essighaus	I	17.1 ±0.5	(17-20)	15.3 ±2.7	16.7 ±1.2	8.2 ±0.5
	I samp	8.1 ±0.4	(8-10)	7.7 ±1.2	8.2 ±0.6	8.2 ±0.5
	I (%)	34.2 ±1.0	(34-40)	30.6 ±5.4	33.4 ±2.4	16.3 ±1.1
	Samp	231654.9 ±28258.4	(90862-237972)	5488.9 ±1526.7	5139.2 ±1138.0	5204.2 ±923.0
	Time _(ms)	2367.7		58.3	79.5	45.8
	Error	77.51 ±1.08	(77.3-87.2)	75.55 ±24.51	78.43 ±20.35	78.18 ±24.42
	LO count	4.2 ±1.2	(1-7)	2.1 ±0.8	6.6 ±1.4	2.1 ±0.8
 sym-londonbridge	I	44.2 ±0.4	(44-45)	43.0 ±4.9	19.8 ±0.8	19.8 ±0.8
	I samp	20.2 ±0.4	(20-21)	19.6 ±2.1	19.8 ±0.8	19.8 ±0.8
	I (%)	70.1 ±0.6	(70-71)	68.2 ±7.7	31.4 ±1.3	31.4 ±1.3
	Samp	22426.2 ±1571.6	(18927-23133)	392.1 ±211.9	343.7 ±125.5	343.7 ±125.5
	Time _(ms)	365.6		24.7	14.7	14.0
	Error	3.57 ±0.73	(2.7-5.5)	4.92 ±8.88	±1.22	±1.22
	LO count	6.3 ±2.0	(1-11)	2.8 ±1.3	3.68 ±1.1	3.68 ±1.1
 sym-notredame12	I	119.9 ±1.6	(117-126)	114.8 ±17.3	46.5 ±0.7	46.5 ±0.7
	I samp	45.9 ±1.0	(44-48)	44.4 ±6.4	46.5 ±0.7	46.5 ±0.7
	I (%)	87.5 ±1.2	(85-92)	83.8 ±12.6	33.9 ±0.5	33.9 ±0.5
	Samp	32526.9 ±2592.7	(27661-39335)	272.9 ±104.5	227.1 ±15.6	227.1 ±15.6
	Time _(ms)	1015.7		40.8	21.7	21.5
	Error	1.05 ±0.14	(0.8-1.5)	1.11 ±0.57	±0.14	±0.14
	LO count	6.5 ±2.3	(2-14)	2.6 ±1.1	2.5 ±1.2	2.5 ±1.2
 sym-notredame13	I	134.3 ±2.1	(121-137)	125.7 ±19.2	56.8 ±1.8	56.8 ±1.8
	I samp	57.8 ±0.9	(54-59)	54.2 ±8.2	56.8 ±1.8	56.8 ±1.8
	I (%)	72.6 ±1.1	(65-74)	68.0 ±10.4	30.7 ±1.0	30.7 ±1.0
	Samp	37303.6 ±2621.0	(34406-49194)	410.1 ±184.7	350.2 ±78.5	350.2 ±78.5
	Time _(ms)	1198.2		53.4	27.7	27.4
	Error	2.37 ±0.15	(1.8-2.8)	2.55 ±2.31	±0.35	±0.35
	LO count	6.9 ±2.4	(3-15)	2.7 ±1.2	2.7 ±1.1	2.7 ±1.1

Image	Qty↓	m2m full	m2m v2	m2m v1	m2m switching	LO-RANSAC
	I	48.7 ±0.5 (47-49)	47.5 ±1.6 (43-49)	46.0 ±6.3 (17-49)	47.6 ±1.5 (43-49)	11.2 ±1.2 (8-12)
	I samp	13.0 ±0.0 (13-13)	12.8 ±0.6 (11-13)	12.5 ±1.1 (8-13)	12.8 ±0.5 (11-13)	11.2 ±1.2 (8-12)
	I (%)	79.8 ±0.8 (77-80)	77.8 ±2.6 (70-80)	75.4 ±10.3 (28-80)	78.0 ±2.5 (70-80)	18.3 ±1.9 (13-20)
	Samp	500000.0 ±0.0 (500000-500000)	1846.0 ±570.5 (1561-3850)	3290.5 ±2419.7 (1561-12406)	2949.2 ±467.6 (2706-4504)	4348.4 ±2406.0 (2185-12406)
	Time _(ms)	5531.8 ±0.04 (0.9-1.1)	60.6 ±0.43 (0.9-2.7)	58.5 ±65.72 (0.9-600.2)	78.9 ±0.36 (0.9-2.7)	40.0 ±0.36 (0.9-2.7)
	Error	0.94 ±2.6 (3-16)	4.0 ±1.3 (1-7)	3.5 ±0.9 (1-5)	7.7 ±1.5 (5-11)	18.19 ±46.38 (1.0-164.2)
	LO count	8.2	4.0	3.5	7.7	3.7
	I	142.1 ±1.4 (138-147)	140.2 ±1.8 (135-144)	134.5 ±25.1 (11-144)	54.0 ±0.8 (52-56)	54.0 ±0.8 (52-56)
	I samp	54.3 ±0.7 (52-56)	53.5 ±0.9 (51-55)	51.4 ±9.3 (6-55)	54.0 ±0.8 (52-56)	54.0 ±0.8 (52-56)
	I (%)	92.9 ±0.9 (90-96)	91.7 ±1.1 (88-94)	87.9 ±16.4 (7-94)	35.3 ±0.5 (34-37)	35.3 ±0.5 (34-37)
	Samp	46547.2 ±2223.9 (41089-51322)	200.8 ±13.8 (178-248)	255.7 ±157.1 (178-1308)	192.1 ±11.6 (165-224)	192.1 ±11.6 (165-224)
	Time _(ms)	1472.5 ±0.05 (0.6-0.9)	47.4 ±0.71 (0.6-0.8)	42.5 ±0.27 (0.6-2.6)	18.6 ±0.06 (0.6-0.9)	18.6 ±0.06 (0.6-0.9)
	Error	0.70 ±2.5 (2-14)	2.2 ±1.0 (1-5)	0.76 ±0.9 (1-6)	2.1 ±1.0 (1-6)	2.1 ±1.0 (1-6)
	LO count	7.5	2.2	0.76	2.1	2.1
	I	146.1 ±1.0 (144-149)	144.7 ±1.1 (142-148)	135.9 ±31.3 (9-148)	53.6 ±0.6 (53-55)	53.6 ±0.6 (53-55)
	I samp	52.4 ±0.8 (51-54)	52.1 ±0.9 (50-54)	49.1 ±10.9 (4-54)	53.6 ±0.6 (53-55)	53.6 ±0.6 (53-55)
	I (%)	109.0 ±0.8 (107-111)	107.9 ±0.8 (106-110)	101.4 ±23.4 (7-110)	40.0 ±0.4 (40-41)	40.0 ±0.4 (40-41)
	Samp	25568.8 ±1525.3 (22965-28931)	128.8 ±8.9 (112-153)	160.3 ±64.1 (112-463)	116.4 ±5.8 (104-150)	116.4 ±5.8 (104-150)
	Time _(ms)	916.6 ±0.13 (0.8-1.4)	34.9 ±1.05 ±0.12 (0.8-1.3)	34.2 ±0.25 (0.8-2.6)	15.3 ±0.14 (0.8-1.5)	15.0 ±0.14 (0.8-1.5)
	Error	1.04 ±2.5 (2-14)	1.9 ±1.0 (1-5)	2.1 ±0.8 (1-5)	1.8 ±0.9 (1-5)	1.8 ±0.9 (1-5)
	LO count	7.3	1.9	2.1	1.8	1.8
	I	201.9 ±1.5 (199-207)	200.6 ±2.2 (192-206)	188.6 ±41.8 (9-206)	82.6 ±1.1 (78-85)	82.6 ±1.1 (78-85)
	I samp	83.5 ±0.8 (82-86)	82.9 ±1.2 (80-86)	78.0 ±17.0 (5-86)	82.6 ±1.1 (78-85)	82.6 ±1.1 (78-85)
	I (%)	99.5 ±0.8 (98-102)	98.8 ±1.1 (95-101)	92.9 ±20.6 (4-101)	40.7 ±0.5 (38-42)	40.7 ±0.5 (38-42)
	Samp	12304.8 ±476.4 (10933-13243)	106.5 ±5.8 (91-123)	153.9 ±92.4 (91-533)	108.6 ±6.0 (101-136)	108.6 ±6.0 (101-136)
	Time _(ms)	749.5 ±0.06 (2.1-2.4)	41.3 ±0.07 (2.1-2.4)	40.8 ±0.13 (2.0-3.2)	17.3 ±0.06 (2.1-2.5)	17.2 ±0.06 (2.1-2.5)
	Error	2.28 ±2.5 (2-15)	1.7 ±0.8 (1-4)	1.9 ±0.7 (1-4)	1.6 ±0.8 (1-4)	1.6 ±0.8 (1-4)
	LO count	6.8	1.7	1.9	1.6	1.6
	I	108.8 ±912.3 (11-9140)	32.0 ±17.4 (15-59)	32.2 ±18.5 (10-59)	29.0 ±16.7 (15-58)	11.7 ±3.1 (7-16)
	I samp	4.4 ±16.2 (0-163)	9.1 ±4.2 (4-16)	9.7 ±3.9 (4-16)	9.1 ±4.2 (4-16)	11.7 ±3.1 (7-16)
	I (%)	66.8 ±559.7 (7-5607)	19.6 ±10.7 (9-36)	19.7 ±11.3 (6-36)	17.8 ±10.2 (9-36)	7.2 ±1.9 (4-10)
	Samp	500000.0 ±0.0 (500000-500000)	372976.7 ±171811.9 (35674-500000)	375673.6 ±161351.7 (35674-500000)	363247.7 ±171859.3 (35674-500000)	294355.3 ±161910.4 (35674-500000)
	Time _(ms)	50695.5 ±74.70 (0.0-709.7)	35979.4 ±32.61 (66.1-243.4)	3888.8 ±54.17 (69.2-462.1)	29841.9 ±44.38 (66.1-365.2)	2346.9 ±35.17 (107.0-231.6)
	Error	164.09 ±1.4 (3-8)	5.9 ±1.6 (3-9)	4.1 ±0.9 (2-7)	9.3 ±2.4 (2-14)	166.76 ±0.8 (2-6)
	LO count	5.6	5.9	4.1	9.3	4.0
	I	75.0 ±0.0 (75-75)	75.0 ±0.0 (75-75)	60.5 ±26.7 (8-75)	36.0 ±0.0 (36-36)	36.0 ±0.0 (36-36)
	I samp	36.0 ±0.0 (36-36)	36.0 ±0.0 (36-36)	29.1 ±12.8 (4-36)	36.0 ±0.0 (36-36)	36.0 ±0.0 (36-36)
	I (%)	138.9 ±0.0 (139-139)	138.9 ±0.0 (139-139)	112.0 ±49.5 (15-139)	66.7 ±0.0 (67-67)	66.7 ±0.0 (67-67)
	Samp	425.0 ±0.0 (425-425)	21.6 ±9.8 (13-53)	7.8 ±11.0 (13-74)	21.6 ±9.8 (13-53)	21.6 ±9.8 (13-53)
	Time _(ms)	19.4 ±0.09 (0.7-1.3)	8.3 ±0.09 (0.7-1.1)	0.92 ±0.37 (0.7-3.2)	0.85 ±0.09 (0.7-1.1)	0.85 ±0.09 (0.7-1.1)
	Error	0.86 ±1.1 (1-6)	1.0 ±0.0 (1-1)	1.0 ±0.1 (1-2)	1.0 ±0.0 (1-1)	1.0 ±0.0 (1-1)
	LO count	1.7	1.0	1.0	1.0	1.0

Bibliography

- [1] David G. Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”. In: *Int. J. Comput. Vision* 60.2 (Nov. 2004), pp. 91–110. ISSN: 0920-5691. DOI: 10.1023/B:VISI.0000029664.99615.94. URL: <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>.
- [2] Jiri Matas et al. “Robust Wide Baseline Stereo from Maximally Stable Extremal Regions”. In: *Proceedings of the British Machine Vision Conference 2002, BMVC 2002, Cardiff, UK, 2-5 September 2002*. 2002, pp. 1–10. DOI: 10.5244/C.16.36. URL: <http://dx.doi.org/10.5244/C.16.36>.
- [3] P. Pritchett and A. Zisserman. “Wide baseline stereo matching”. In: *Computer Vision, 1998. Sixth International Conference on*. Jan. 1998, pp. 754–760. DOI: 10.1109/ICCV.1998.710802.
- [4] D. Nister. “Preemptive RANSAC for live structure and motion estimation”. In: *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. Oct. 2003, 199–206 vol.1. DOI: 10.1109/ICCV.2003.1238341.
- [5] J. Philbin et al. “Object Retrieval with Large Vocabularies and Fast Spatial Matching”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2007.
- [6] Martin A. Fischler and Robert C. Bolles. “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”. In: *Commun. ACM* 24.6 (June 1981), pp. 381–395. ISSN: 0001-0782. DOI: 10.1145/358669.358692. URL: <http://doi.acm.org/10.1145/358669.358692>.
- [7] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [8] Krystian Mikolajczyk and Cordelia Schmid. “Scale & affine invariant interest point detectors”. In: *International journal of computer vision* 60.1 (2004), pp. 63–86.
- [9] O. Chum, T. Werner, and J. Matas. “Two-view geometry estimation unaffected by a dominant plane”. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. June 2005, 772–779 vol. 1. DOI: 10.1109/CVPR.2005.354.
- [10] O. Chum and J. Matas. “Matching with PROSAC - progressive sample consensus”. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. June 2005, 220–226 vol. 1. DOI: 10.1109/CVPR.2005.221.
- [11] J. Matas and O. Chum. “Randomized RANSAC with sequential probability ratio test”. In: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*. Vol. 2. Oct. 2005, 1727–1732 Vol. 2. DOI: 10.1109/ICCV.2005.198.
- [12] Ondřej Chum, Jiří Matas, and Josef Kittler. “Locally optimized RANSAC”. In: *Pattern Recognition*. Springer, 2003, pp. 236–243.

- [13] K. Lebeda, J. Matas, and O. Chum. “Fixing the Locally Optimized RANSAC.” In: *BMVC*. 2012, pp. 1–11.
- [14] Wei Zhang and J. Kosecka. “Generalized RANSAC Framework for Relaxed Correspondence Problems”. In: *3D Data Processing, Visualization, and Transmission, Third International Symposium on*. June 2006, pp. 854–860. DOI: 10.1109/3DPVT.2006.67.
- [15] Andrej Mikul’k et al. “Learning a fine vocabulary”. In: *Computer Vision–ECCV 2010*. Springer, 2010, pp. 1–14.
- [16] *LAPACK Linear Algebra Package*. <http://www.netlib.org/lapack/>.
- [17] Daniel Cabrini Hauage and Noah Snavely. “Image matching using local symmetry features”. In: *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE. 2012, pp. 206–213.
- [18] L. Van Gool H. Shao T. Svoboda. *ZuBuD - Zurich Buildings Database for Image based Recognition*. Tech. rep. 2003. URL: <http://www.vision.ee.ethz.ch/showroom/zubud>.
- [19] Andrea Cohen et al. “Discovering and exploiting 3d symmetries in structure from motion”. In: *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE. 2012, pp. 1514–1521.