



NAPSAC: HIGH NOISE, HIGH DIMENSIONAL ROBUST ESTIMATION - IT'S IN THE BAG

D.R. Myatt P.H.S. Torr S.J. Nasuto J.M. Bishop R. Craddock

Abstract

A number of the most powerful robust estimation algorithms, such as RANSAC, MINPRAN and LMS, have their basis in selecting random minimal sets of data to instantiate hypotheses. However, their performance degrades in higher dimensional spaces due to the exponentially decreasing probability of sampling a set that is composed entirely of inliers. In order to overcome this, rather than picking sets at random, a new strategy is proposed that alters the way samples are taken, under the assumption that inliers will tend to be closer to one another than outliers. Based on this premise, the NAPSAC (N Adjacent Points Sample Consensus) algorithm is derived and its performance is shown to be superior to RANSAC in both high noise and high dimensional spaces.

1 Introduction

Robust estimation has found many uses in computer vision. Indeed, one of the most powerful robust estimation algorithms, RANSAC [1] was developed for registering 3D to 2D point sets. RANSAC has also been used extensively for such tasks as estimation of epipolar geometry[2] and motion model selection [3], and has spawned a variety of robust algorithms all with their basis in random sampling being used to minimize different criteria e.g. LMS [4], MINPRAN [5], MLESAC [2].

Whilst the robust estimation problem has been largely solved in low dimensional cases by random sampling methods, there is no known generic robust estimator that will work for large numbers (>10) of parameters. This paper explores one way of extending the random sampling formalism to higher dimensions.

Random sampling methods are explained in Section 2. These algorithms sample minimal sets¹ of the data, such that each point has an equal likelihood of being selected. These sets are then used to form hypotheses as to the parameters of the model, θ . A cost function, $C(\hat{\theta})$, is used to determine the fit of the estimated hypersurface to the data. The idea is to draw m minimal sets, assuming at least one set will be found with all inliers, and thus yield optimal $C(\hat{\theta})$. As the proportion of outliers and the minimal set size increase, m must be increased exponentially to retain a significant chance of selecting good $\hat{\theta}$.

¹A minimal set is the minimum number of data, n , required to estimate the parameters θ of the manifold.



The key concept introduced in Section 3 is that, given the outliers possess a diffuse distribution, the selection of minimal sets based on proximity can significantly improve the probability of selecting a set of inlying data, and thus drastically reduce the samples, m , required.

This selection method is investigated more formally in Section 4, leading to a proof showing how it is beneficial in 2 dimensions. In Section 5 it is then shown experimentally that this approach is valid in higher dimensions. An enhanced point selection algorithm is described and integrated into the RANSAC algorithm - the result is named N Adjacent Points SAmple Consensus (NAPSAC). This algorithm is compared with standard random sampling algorithms, showing the superior behaviour of NAPSAC in high noise and higher dimensions. Section 6 presents conclusions and avenues of future work discussed.

Conventions: presented here is a clarification of the conventions used throughout this paper.

\mathbf{x} - bold type indicates a vector.

\underline{x} - an underline represents noise free data.

$\hat{\theta}$ - a hat indicates an estimate.

2 Robust Estimation

Fundamentally, the problem is that of trying to estimate the parameters, θ , of some manifold, \mathcal{M} , defined by $\mathbf{g}(\underline{\mathbf{x}}_i, \theta) = 0, i = 1 \dots n$, where $(\underline{\mathbf{x}}_i)$ are the noise free points. Outliers are points of gross error generated independently of θ with distribution $\psi_o(x)$, which in absence of any information are assumed to be uniform. The overall data set is therefore a mixture containing a proportion, μ , that arise from the manifold, but are perturbed by noise $\mathbf{x}_i = \underline{\mathbf{x}}_i + \psi$ (here assumed Gaussian), the inlier distribution is denoted $\psi_i(x, \theta)$. In order to estimate θ effectively it is necessary to data generated by $\psi_i(x, \theta)$. Let γ_i be an indicator variable, such that

$$\gamma_i = \begin{cases} 1 & \text{if the } i\text{th datum is generated by } \psi_i(x, \theta), \\ 0 & \text{if the } i\text{th datum is generated by } \psi_o(x) \end{cases} \quad (2.1)$$

and therefore the actual probability of selecting an inlier from a data set with size S is $\underline{\gamma} = 1/S \sum \gamma_i$. However, if we assume S is large then this probability will tend towards μ .

RANdom SAmple Consensus (RANSAC)[1], MINimise Probability of RANdomness (MINPRAN)[5] and Least Median Squares (LMS) are among the most effective algorithms in this field. They all use the same general paradigm which can be summarised as follows:

for $i = 1$ to m

 Sample a minimal set of data to create an estimate of θ , $\hat{\theta}_i$.

 Use a cost function, $C(\hat{\theta}_i)$, to determine the robust error of $\hat{\theta}_i$.

next

Select the model, $\hat{\theta}'$, for which $C(\hat{\theta}_i)$ was minimal, $i = 1 \dots m$

Use classified inliers to make a final estimate of $\hat{\theta}$.



3 Minimal Set Sampling

It is apparent that, in order to estimate θ effectively, the set of inliers must be obtained. This is achieved through the repeated sampling of minimal sets of n points, where n is the minimum number of data points required to instantiate the model's free parameters. The probability of selecting n inlying points at random is monotonically decreasing with respect to increasing n , thus the minimal set paradigm *maximises* the probability of selecting a set of inliers from which a hypothesis may be derive. Currently, these minimal sets are selected uniformly over all data, giving each datum an equal probability of selection - *uniform point sampling*. So, given that there is a probability, μ , of selecting an inlier generated by $\psi_i(x, \theta)$, then the probability of selecting all inliers, from n point selections, is clearly μ^n , and consequently the efficacy of uniform point sampling degrades exponentially as dimensionality increases. Table 1 shows the theoretical number of samples required for an algorithm using uniform point sampling to have a 95% chance of selecting one of more sets of inliers. This can be calculated using simple binomial distribution and is a reformulation in higher dimensionality of a table found in [6]. It is evident that the required number of samples increases exponentially with dimensionality, and therefore this technique is impractical in higher dimensions.

Dimensionality	Percentage of outliers		
	30%	40%	50%
2	5	7	11
3	8	13	23
4	11	22	47
5	17	38	95
6	24	63	191
7	35	106	382
8	51	177	766
9	73	296	1533
10	105	494	3067
20	3753	81936	3.1×10^6
30	132910	1.4×10^7	3.2×10^9
40	4.7×10^6	2.2×10^9	3.3×10^{12}

Table 1: *The number of samples, using uniform point sampling, required to achieve a probability 0.95 of selecting one or more sets of inliers, for three levels of outlying noise*

The failure, in higher dimensions, of uniform point sampling may be related to its neglect of the spatial relationship between the inlying data points. Using the distribution of the inlying data within the multi-dimensional space to modify the point sampling may improve hypothesis selection. Such models are already used to determine the quality of an estimation. The method proposed here is to use a similar technique to *select* hypotheses through improved point selection.



4 Mathematical derivation

Within this section we analyse the probability of selecting inliers based on proximity in 2D. It is assumed that a point, \mathbf{x}_0 , has already been selected that lies on the manifold, so therefore the marginal density of inliers at a distance r from \mathbf{x}_0 can be calculated, and similarly for the outlier distribution. Then, by comparing these marginal densities, it can be determined whether selecting by proximity increases the probability of sampling inliers over uniform random sampling. Without loss of generality the manifold is taken to be the line $y = 0$, such that the noise free points $\underline{\mathbf{x}}_i = (\underline{x}_i, \underline{y}_i)$ satisfy $\underline{y}_i = 0$. We shall examine the worst case scenario, where the points are uniformly distributed on the manifold between limits $-v < x < v$, such that the probability density function of \underline{x} is $p_i(\underline{x}) = \frac{1}{2v}$. The inliers are then independently perturbed orthogonally to the manifold by Gaussian noise, so $x = \underline{x}$, $y = \underline{y} + \delta y$, where δy is generated by $N(0, \sigma)$. This distribution is truncated $-v < \underline{y} < v$, but because it is assumed that $v \gg \sigma$, the truncation will have negligible effect on subsequent calculations. Thus, the joint probability density of the inliers is

$$p_i(x, y) = \begin{cases} \frac{1}{2v\sigma\sqrt{2\pi}} \exp\left(-\frac{y^2}{2\sigma^2}\right) & \text{if } -v < x < v, -v < y < v, \\ 0 & \text{otherwise.} \end{cases} \quad (4.1)$$

Outliers are distributed uniformly in a hypersphere centred on the origin with radius v . Their density is given by

$$p_o(x, y) = \begin{cases} \frac{1}{\pi v^2} & \text{if } \sqrt{x^2 + y^2} < v, \\ 0 & \text{otherwise.} \end{cases} \quad (4.2)$$

The overall distribution of data is therefore a mixture of $p_i(x, y)$ and $p_o(x, y)$ with inlier component, μ :

$$p(x, y) = \mu p_i(x, y) + (1 - \mu) p_o(x, y) \quad (4.3)$$

In order to find the conditional probability density of selecting an inlying point from a given point on the manifold as a function of their mutual distance, a co-ordinate transform from Cartesian to Polar is required:

$$p_i(r, \phi) = p_i(r \cos \phi, r \sin \phi) \left| \begin{bmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{bmatrix} \right| \quad (4.4)$$

$$p_i(r, \phi) = Kr \exp\left(-\frac{r^2 \sin^2 \phi}{2\sigma^2}\right), \text{ where } K = \frac{1}{2v\sigma\sqrt{2\pi}} \quad (4.5)$$

The marginal density, $m_i(r)$, is the integral wrt ϕ .

$$m_i(r) = \int_0^{2\pi} Kr \exp\left(-\frac{r^2 \sin^2 \phi}{2\sigma^2}\right) \delta\phi \quad (4.6)$$

However, using the trigonometric identities $\sin^2 \phi = \sin^2(\phi + \pi)$ and $\sin^2 \phi = \frac{1}{2}(1 - \cos 2\phi)$ this can be rearranged to



$$m_i(r) = 2Kr \exp\left(-\frac{r^2}{4\sigma^2}\right) \int_0^\pi \exp\left(\frac{r^2}{4\sigma^2} \cos 2\phi\right) \delta\phi \quad (4.7)$$

Integrating (4.7) gives

$$m_i(r) = \mu \frac{1}{v} \sqrt{\frac{\pi}{2}} \left[\frac{r}{\sigma} \exp\left(-\frac{r^2}{4\sigma^2}\right) I_0\left(-\frac{r^2}{4\sigma^2}\right) \right] \quad (4.8)$$

where I_o is a Modified Bessel Function of the First Kind. Similarly, transforming the outlier joint density (4.2) from Cartesian to Polar co-ordinates yields

$$p_o(r, \phi) = \frac{1}{\pi v^2} r \quad (4.9)$$

Thus, integrating (4.9) wrt ϕ results in the outlier marginal density:

$$m_o(r) = \int_0^{2\pi} \frac{1}{\pi v^2} r \delta\phi = \frac{2}{v^2} r \quad (4.10)$$

The left panel of figure 1 shows a plot of the marginal inlier and outlier densities. It can be seen that the inlier density is indeed higher closer to the manifold. Also, $m_i(r)$ tends to a constant value that is independent of σ , as can be seen from the right panel.

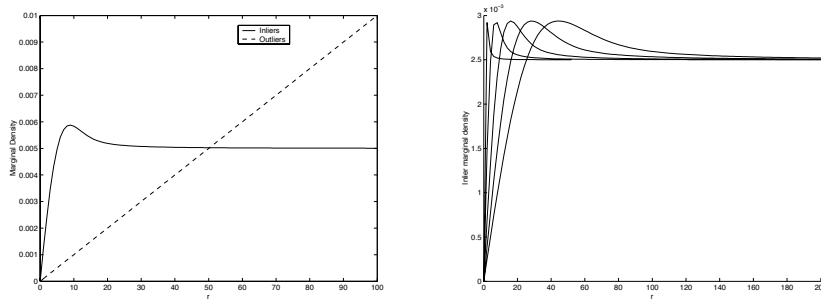


Figure 1: *Left panel - Plot of marginal densities, given $v = 100$, $\sigma = 3$, $\mu = 0.5$. Right panel - Plot of $m_i(r)$ for $\sigma = i^2$, $i = 1 \dots 5$, when $\mu = 0.5$, $v = 200$*

It can be shown geometrically that, as r increases, $m_i(r) \rightarrow \frac{1}{v}$. Consequently, if r_c denotes a distance at which both marginal densities are equal, then

$$r_c \approx \frac{v}{2} \frac{\mu}{(1 - \mu)} \quad (4.11)$$

Figure 2 shows that the estimation of r_c given by (4.11) is very accurate, given that $\sigma \ll v$. However, for large σ or low μ (< 0.1) this estimate will degrade significantly. For uniform sampling, the probability of selecting an inlier is constant at μ , the inlying component of the mixture. Therefore, by selecting a single point A initially and then selecting the rest of the points based on their relative proximity to A , the probability of achieving a set of all inliers will be greatly increased over unguided random selection. For tutorial purposes this section has concentrated on a 2D analysis. However, the result can be generalised to higher dimensions.

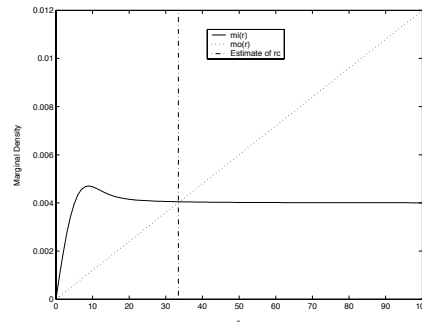


Figure 2: Accuracy of estimation of r_c , for $v = 100$, $\sigma = 5$, $\mu = 0.4$

5 NAPSAC

Non-uniform sampling has been shown to provide a theoretical advantage over uniform sampling, but will now be shown experimentally to be just as effective in high noise and higher dimensions. To demonstrate this, a simple enhanced sampling algorithm was created. The following algorithm can be used in place of the uniform point sampling process in any of the robust estimation algorithms discussed.

1. Select an initial point \mathbf{x}_0 randomly from all points
2. Find the set of points, $\mathcal{S}_{\mathbf{x}_0}$, lying within a hypersphere of radius r centred on \mathbf{x}_0 .
3. If the number of points in $\mathcal{S}_{\mathbf{x}_0}$ is less than the minimal set size then fail.
4. Select points from $\mathcal{S}_{\mathbf{x}_0}$ uniformly until the minimal set has been selected, inclusive of \mathbf{x}_0 .

This results in a cluster of points being selected from a ball. If the initial point, \mathbf{x}_0 , lies on the manifold, then the rest of the points sampled adjacently will theoretically have a significantly higher probability of being inliers. If there are not enough points within the hypersphere to estimate the manifold, then that sample is considered a failure.

The enhanced sampling algorithm was integrated with the RANSAC consensus set cost function [1] to facilitate experimentation. This amalgamation was named N Adjacent Points SAmples Consensus (NAPSAC).

5.1 Experimental results

In order to estimate the epipolar geometry of a scene at least two images are required from different positions. In general, most vision applications use images from a video sequence or similar, such that the disparity is small and a given feature moves only a few pixels. This allows feature matching methods to find correspondences effectively. In wide baseline stereo matching applications [7], or in the case of object recognition [8], however, the motion between images is large

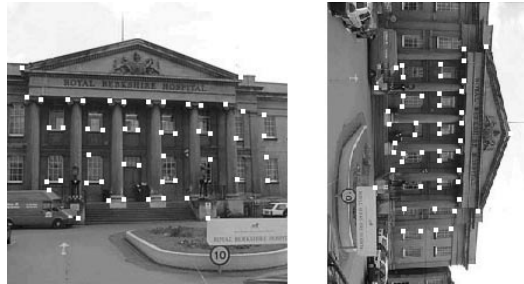


Figure 3: *The images used for testing, with the manually defined correspondences shown*

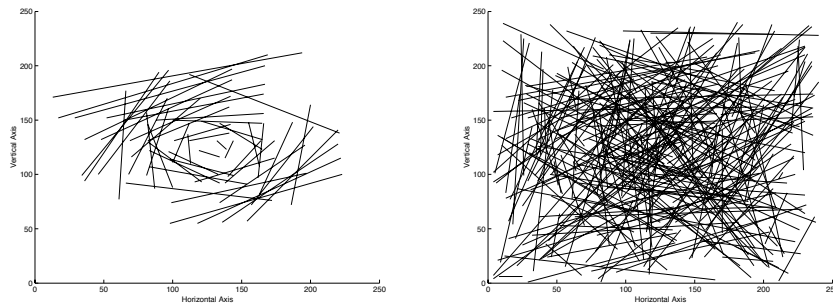


Figure 4: *Left Panel - the 50 manually defined inlying correspondences. Right Panel - a test data set with 200 outlying correspondences*

e.g. a 90 degree rotation and scaling, and this results in a very large proportion of incorrect correspondences, as only a weak metric can be used for matching. Also, because matches can be found between features detected anywhere on the image, the outlying data will tend towards uniformity [8], which has the advantage of making the problem simpler when using NAPSAC. This is because the inliers and outliers will be more separated in the 4D joint image space created by concatenating the 2D co-ordinate vectors from both images [9]. On small baseline matching problems, conversely, the outliers distribution will contain more structure and consequently NAPSAC will not be as effective.

Standard RANSAC was compared against NAPSAC on a general wide baseline stereo matching problem: figure 3 shows two 240×240 images, where the second is a 90 degree rotation and 75% scaling of the first. Fifty correct correspondences between the images were added manually, thus allowing varying amounts of outlying data to be added. Incorrect correspondences were generated uniformly within the entire span of the image. A number of outlying matches ranging from 0 to 450 were added to the inlying correspondences that resulted in outlier proportions from 0 to 0.9, in increments of 0.05. The left panel of figure 4 shows the inlying correspondences, and the right panel a data set with 0.8 outlying noise. Although the transformation between the images can be described by a

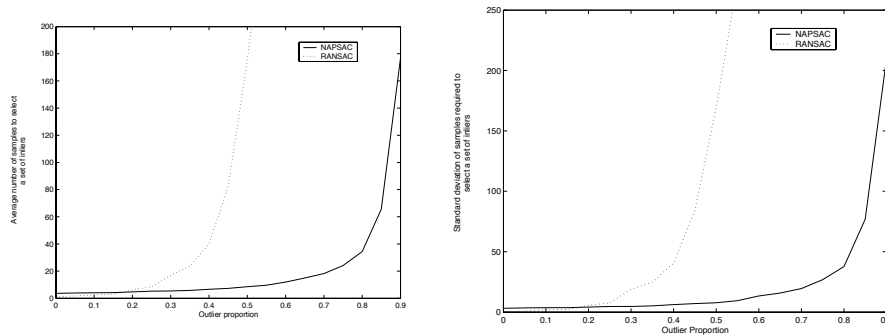


Figure 5: *Left Panel - mean number of samples required to select a set of inliers for NAPSAC and RANSAC. Right Panel - standard deviation of number of samples required to select a set of inliers.*

homography, which can be defined by 4 correspondences, 7 data were selected to demonstrate the possibility of estimating epipolar geometry in appropriate cases. The experiment was designed to show that the enhanced sampling of NAPSAC has a much greater probability of selecting an inlying set than RANSAC. Further, it will be shown that the inlying sets selected are not degenerate and can effectively recover the inlying data.

The number of samples required for RANSAC and NAPSAC (using a hypersphere of radius 50) to select a fully inlying set was recorded over 1000 trials for each level of noise. Figure 5 shows the mean and standard deviation of the number of samples required for a successful estimation with respect to the proportion of outliers.

As the proportion of noise increases the number of samples required by RANSAC follows an exponential increase as expected, and reaches 6000 for 70% noise. NAPSAC, by comparison, shows much more graceful degradation and requires less than 200 samples on average even in 90% noise. The standard deviation for both algorithms follows a similar pattern. So, this example shows that NAPSAC requires many fewer samples than RANSAC to select an inlying set. For simplicity, the breaking point criteria here will be assumed to be the proportion of noise for which the mean required samples exceed 200. These values are approximately 50% and 90% for RANSAC and NAPSAC respectively. However, the optimal radius of the NAPSAC hypersphere varies slightly for different data sets. Further work is in progress to address this issue. Also, since NAPSAC relies on hyper-spherical clusters there needs to be regions on the image that have several good correspondences, rather than a uniform scattering. This assumption is already made by several algorithms, including MINPRAN [5]. NAPSAC was also used to estimate a homography for the hospital data set to demonstrate that the minimal sets were not degenerate despite their selection by proximity (estimating the fundamental matrix would be inappropriate as all \mathbf{F} would be degenerate in this case). Figure 6 shows the best minimal sets selected by NAPSAC in three consecutive trials and the resulting inlying set. Out of 100 trials, 94% yielded a homography that recovered 20 or more inlying data with at maximum one Type II error (an

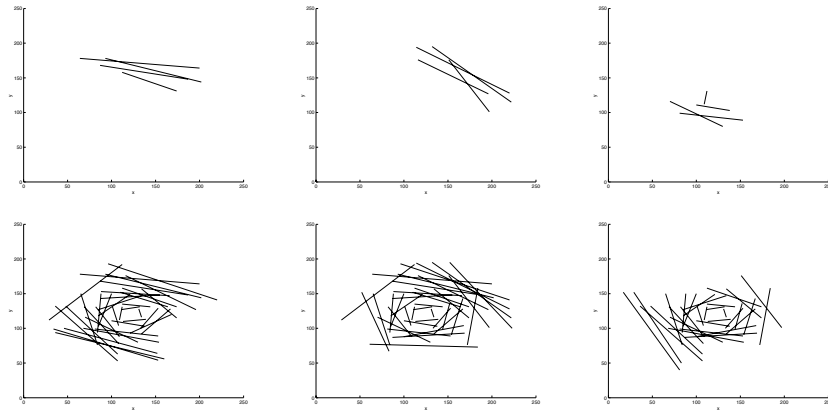


Figure 6: *Three examples of minimal sets selected by NAPSAC to form a homography in 80% outlying noise and the recovered inlying set*

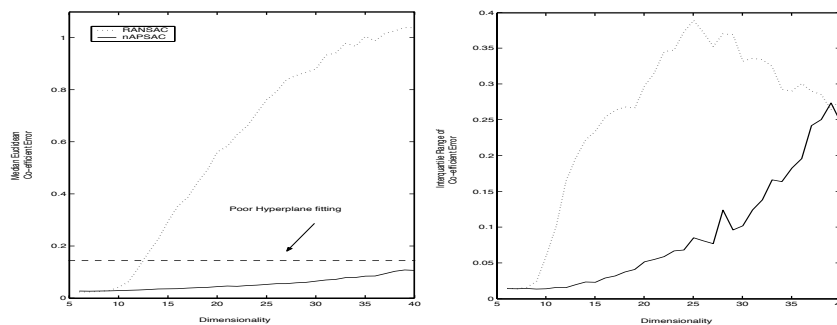


Figure 7: *A comparison between the RANSAC and NAPSAC algorithms in higher dimensions. Upper Panel - Mean error of NAPSAC and RANSAC estimated hyperplanes. Lower Panel - Standard deviation of errors.*

outlier classified as an inlier). This is sufficient inlying data to accurately recover the homography through post-processing. It could mistakenly be concluded that NAPSAC performs effectively on this example because of a tendency to bias in 2D - this is because the 2D transform exactly preserves the adjacency of the features between the images. Figure 6 shows that, on the contrary, the minimal set correspondences can cover a significant region of the image densely populated by outliers, and consequently many 3D transforms could be tolerated with an appropriately sized hypersphere.

NAPSAC is also more effective in high dimensionality than RANSAC. Figure 7 shows a comparison of the two approaches on a simulated hyperplane estimation problem in 50% outlying noise over 1000 trials - it can be seen that NAPSAC is more robust. Selecting points based on proximity increases the probability of sampling inliers both when a large sample of inliers is required and when a smaller sample is required in high noise.



6 Conclusions

This paper has introduced a new method for sampling points for minimal set algorithms such as RANSAC and LMS. Currently such algorithms rely on uniform sampling from the data, Using NAPSAC, a modification of RANSAC, which samples sets of adjacent points in a hypersphere, it has been shown experimentally that the probability of selecting an inlying set is significantly increased in both high outlying noise and higher dimensions, as demonstrated using a wide baseline stereo matching application and a high-dimensional hyperplane estimation problem, respectively. This suggests a shift from research into effective robust error functions to an exploration of how minimal set selection, and therefore hypothesis selection, may be most effectively enhanced using the distribution of the data in higher dimensional space, and an alternative approach to NAPSAC has already been proposed by Tordoff, 2002 [10]. Possible further work includes calculation of the optimal hypersphere radius for a given data set. An exploration of different methods of selecting points to enhance inlier selection, not just hyperspherical, is also suggested.

References

- [1] M.A. Fischler and R.C. Bolles. ‘Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography’, *Comm. Assoc. Comp. Mach.*, 24(6): 381–395, 1981
- [2] P.H.S. Torr and A. Zisserman, ‘MLESCAC: A new robust estimator with application to estimating image geometry’, *Comp.Vision and Imag.Underst.*, vol. 78, pp. 138–156, 2000.
- [3] P.H.S. Torr and D.W. Murray. ‘Outlier detection and motion segmentation’. In *SPIE 93*, 1993.
- [4] P. J. Rousseeuw. *Robust Regression and Outlier Detection*. Wiley, New York, 1987.
- [5] C.V. Stewart. ‘MINPRAN: A new robust operator for computer vision’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(10): 925–938, 1995.
- [6] P.H.S. Torr, A. Zisserman, and S.J. Maybank. ‘Robust detection of degenerate configurations while estimating the fundamental matrix’, *Computer Vision and Image Understanding*, 71(3): 312–333, 1998.
- [7] P.C. Pritchett and A.P. Zisserman. ‘Wide baseline stereo matching’. In *Proc. 6th Int’l Conf. on Computer Vision*, Bombay, January 1998.
- [8] D.G. Lowe. ‘Local Feature View Clustering for 3D Object Recognition’. In *Proc. Int’l Conf. on Computer Vision*, 2001.
- [9] W. Triggs. ‘The geometry of projective reconstruction i: Matching constraints and the joint image’. In *Proc. 5th Int’l Conf. on Computer Vision*, Boston, pages 338–343, 1995.
- [10] B. Tordoff and D. Murray. ‘Guided sampling and consensus for motion estimation’. *ECCV2002*.