The Imaging Science Journal
Volume 65, 2017, Issue 7, pp. 428-437
DOI: 10.1080/13682199.2017.1361665

Evaluating 3D Local Descriptors for Future LIDAR Missiles with Automatic Target Recognition Capabilities

Future Light Detection and Ranging seeker missiles incorporating 3D Automatic Target Recognition (ATR) capabilities can improve the missile's effectiveness in complex battlefield environments. Considering the progress of local 3D descriptors in the computer vision domain, this paper evaluates a number of these on highly credible simulated air-to-ground missile engagement scenarios. The latter take into account numerous parameters that have not been investigated yet by the literature including variable missile – target range, 6-Degrees-of-Freedom missile motion and atmospheric disturbances. Additionally, the evaluation process utilizes our suggested 3D ATR architecture that compared to current pipelines involves more post-processing layers aiming at further enhancing 3D ATR performance. Our trials reveal that computer vision algorithms are appealing for missile oriented 3D ATR.

Keywords: 3D Automatic Target Recognition; Light Detection and Ranging; Missile engagement scenarios

Introduction

Extending ATR from the 2D to the 3D domain offers numerous advantages. Indicatively, features (data) extracted from the 3D domain are less affected by illumination variation and target pose changes [1] and describe the underlying structure of an object. With respect to future LIDAR based missiles, 3D ATR can improve weapon effectiveness against camouflage, concealment and deception techniques. In addition, the LIDAR's short operating wavelength provides high-resolution data and the capability to acquire details of the target reinforcing recognition applications.

Current computer vision algorithms suggest a number of local feature description techniques appropriate for 3D ATR. These techniques describe patches around a point of interest and thus afford object recognition for partially visible targets. Current local based descriptors include SHOT [2], Spin Images [3], RoPS [4], Tri-Spin

Images [5], FPFH [6], 3DSC [7], USC [8] and HoD [1]. A common principle of all these descriptors is encoding a spherical volume $\boldsymbol{V}$ of radius r that is centred on a keypoint p. The 3D descriptors evaluated in this paper are:

- Histogram of Distances. HoD [1] calculates the probability mass density of the normalized point-pair L2-norm distance distributions within $\boldsymbol{V}$. L2-distances are encoded in a coarse and a fine manner by using different quantizing bin sizes, which are concatenated in a single descriptor. This strategy enhances ATR performance for noise and subsampling perturbations. In contrast to current 3D descriptors, HoD does not require a LRF/A. This is important because the repeatability of a LRF/A is prone to noise and point cloud subsampling [2]. Additionally, HoD has an adaptive description radius that depends on average point cloud resolution of each scene rather than the template, which is the norm for a 3D descriptor.
- Histogram of Distances – Short. HoD-S is a compact version of HoD that relies only on the coarse part of HoD. Except that, HoD-S and HoD share the same principles.
- Signatures of Histograms of Orientations. SHOT [2] divides $\boldsymbol{V}$ into a number of sub-volumes along the azimuth, the elevation and the radius. For every sub-volume, a 1-dimensional local histogram is computed considering the normal variation produced by the keypoint p (including its surrounding vertices) and the vertices within each sub-volume.
- Rotational Projection Statistics. RoPS [4] establishes on $\boldsymbol{V}$ a LRF. Then $\boldsymbol{V}$ is rotated around each axis of the LRF's coordinate frame and is then projected on each of the coordinate planes. Finally, each projection undergoes a statistical

analysis based on low order moments and entropy, which are converted into a histogram named the RoPS descriptor.

- 3D Shape Context. 3DSC [7] establishes on $\boldsymbol{V}$ a LRA that is aligned with the normal of the vertices enclosed in $\boldsymbol{V}$. Then $\boldsymbol{V}$ is divided into several bins along the azimuth, elevation and radial dimension, and the 3DSC descriptor is created by accumulating a weighted sum of the points within each bin. Weights are proportional to the bin to centre-of-$\boldsymbol{V}$ distance. 3DSC is LRA based and compensates for 360° azimuthal rotation by describing $\boldsymbol{V}$ in multiple azimuthal orientations.
- Unique Shape Context. USC [8] is an extension of the 3DSC where the LRA is substituted by a LRF. Except that, USC and 3DSC identical.
- Fast Point Feature Histograms. FPFH [6] establishes on $\boldsymbol{V}$ a *Darboux* LRF. Then for each point belonging to $\boldsymbol{V}$, FPFH encodes the angular relationship between the keypoint p and its neighbours as provided by the LRF. Finally, that angular relationship is transformed into a histogram.

Driven by the appealing advantages of missile 3D ATR we extend the typical 3D object recognition architecture [9,10] and use it to evaluate the performance of the aforementioned 3D descriptors on complex simulated air-to-ground missile engagement scenarios. We extend the object recognition architecture by involving a smooth surface filtering procedure during the pre-processing phase and a quadruple post-processing scheme aiming at further enhancing the architecture's 3D ATR performance. In contrast to current pipelines that use up to three post-processing layers [4,5,11,12], we use four. That is, a dual-layered hypothesis generation scheme based on Geometric consistency checks and RANSAC, and a dual-layered hypothesis verification scheme that relies on ICP and on a geometrical cue based acceptance criterion.

Compared to current literature that evaluates the recognition performance of 3D descriptors, ours differ as it exploits military rather than computer vision scenarios [13,14] and it simulates scale changes and atmospheric disturbances affecting LIDAR. It is worth noting that literature has not investigated yet how atmospheric disturbances affect ATR tasks. Additionally, opposing to current 3D descriptor evaluations [13,14] or to purely military oriented 3D ATR manuscripts [15–18] that consider a constant target scale, we investigate the descriptor's performance in various missile-target ranges. It should be noted that, in contrast to current literature that correlates *scale* with the description radius [14,19] we refer to its distance related meaning that affects the spot size of the laser beam on the scene and forces the scene to have simultaneously a different size and resolution. An example of how the missile – target range affects the scene's size and resolution is presented in Figure 1.

3D ATR Missile Seeker Architecture

We extend the typical 3D object recognition architecture [9,10] by introducing additional pre and post-processing operations aiming to enhance the architecture's recognition performance. The suggested pipeline comprises of an offline and an online phase. During the former, the ideal 3D point cloud of a template MBT is manipulated such as to emulate a realistic representation for missile based applications. The online pre-processing phase aims at filtering the scene's smooth surfaces and then implements keypoint description and matching depending on the 3D descriptor evaluated. During the post-processing phase, the pipeline exploits a dual-layered hypothesis generation scheme that is based on Geometric consistency checks and RANSAC, and a dual-layered hypothesis verification scheme that relies on ICP and on an acceptance criterion that relies on geometrical cues. The 3D ATR architecture is presented in Figure 2 (a). It should be noted that the online adaptive description radius of HoD and HoD-S imply adding to their ATR architecture and *online-template* module as presented in Figure 2 (b). The differences between the two architectures are highlighted in Figure 3.

Offline phase

During the offline stage, the input is a 3D point cloud $\boldsymbol{P_m}$ of a Leopard C2 MBT to be recognized. Typical anti-tank missiles attack the MBT in its top and side view pose to defeat the target where armour is thinnest. Thus, we reject the lower part of $\boldsymbol{P_m}$ by applying the Hidden Point Removal algorithm [20] in order to emulate a synthesized view of various MBT poses as observed by the LIDAR missile seeker during the final engagement. The remaining part $\boldsymbol{P_{HPR}}$ is shown in Figure 4. $\boldsymbol{P_{HPR}}$ is then uniformly subsampled at 0.3-meter resolution to enhance descriptiveness and computational efficiency. For simplicity, instead of using a keypoint detection strategy, we describe all vertices of the subsampled $\boldsymbol{P_{HPR}}$ with the 3D descriptors of Table 1. The description radius used is the standard $15\,\mathrm{Res}$ where Res is the average point cloud resolution of the template [2,4,21] (for HoD and HoD-S it is the scene resolution [1]). Due to the large number of points to be described, all descriptors are assembled into a FLANN structure [22] that will be used during the matching stage. Finally, $\boldsymbol{P_m}$ is also uniformly subsampled at 0.3-meter resolution and stored as $\boldsymbol{P_{mt}}$.

Online phase

The input to the online phase is a scene point cloud $\boldsymbol{P}$ that is uniformly subsampled at 0.3-meter resolution that comprises of the vertices $P_c, c \in \aleph$.

Smooth surface filtering

We associate to each P_c a normal by estimating the best fitting local plane to its six closest neighbours. Then, for each P_c acting as a centroid, we extract a spherical volume $\boldsymbol{P_d}$ with 10-meter radius which is equal to the MBT length used during simulations. Then, the following cost function defines whether each P_c is rejected as

being part of a greater smooth area or is preserved:

$$P_c = \begin{cases} accepted & if\ \sigma\left(\left\langle P_{dc} \right\rangle\right) > 10^o \\ rejected & elsewhere \end{cases} \quad (1)$$

where $\sigma\left(\left\langle P_{dc} \right\rangle\right)$ is the standard deviation of the normals enclosed in $\boldsymbol{P_d}$ which is centred at P_c. The accepted P_c comprise the filtered point cloud scene $\boldsymbol{P_f}$:

It is worth noting that simply filtering the scene by comparing $\left\langle P_{dc} \right\rangle$ with the average normal of its surrounding vertices has questionable performance because it is not robust even to minor perturbations. An example highlighting the effectiveness of the suggested standard deviation based smooth surface filtering compared to the simplistic average based technique is shown in Figure 5. Despite the effectiveness of a smooth surface filtering scheme, surprisingly, current military literature either does not exploit such a procedure at all [15] or discards only planar ground surfaces [18,23].

Keypoint description, matching and consistency checks

The scene vertices $\boldsymbol{P_f}$ are then described by the descriptors of Table 1. For the feature matching, we extend [24] and de-correlate kNNDR with k=10 by shifting the matching burden to a number of Geometric consistency checks [25]. These checks repeat until all correspondences from the kNNDR stage are grouped into one of the clusters $H_a, a \in \aleph$ with $\boldsymbol{H} = \{H_1, H_2, ..., H_a\}$. Clusters that have a cardinality greater than 66% of the largest cluster in $\boldsymbol{H}$ are maintained and comprise $H_b, b \in \aleph, b < a$ with $H_b \in \boldsymbol{H'} \subset \boldsymbol{H}$, while the rest are discarded as too small.

Hypothesis generation and verification

Each cluster of $\boldsymbol{H'}$ defines a transformation hypothesis T between the model and the target. Inconsistent correspondences within the same transformation hypothesis $\boldsymbol{H'}$ are

discarded based on the RANSAC algorithm using 1000 iterations. Finally, a geometrical cue verification task rejects false transformation hypotheses by applying each T on the point cloud model $\boldsymbol{P_{mt}}$ creating $\boldsymbol{P_{mf}}$. The latter is fine aligned with the scene point cloud $\boldsymbol{P_f}$ via the ICP technique using 1000 iterations. Finally, a hypothesis H_b is accepted if the following constraint is fulfilled:

$$\frac{\sum \left\| P_{mf} - P_f \right\|}{\sum P_{mf}} > 1\% \;\; s.t. \;\; \left\| P_{mf} - P_f \right\| < 2 \cdot \mathrm{Re}\, s \qquad (2)$$

The acceptance criteria of 1% is experimentally chosen such as to afford recognizing the target even if it is heavily occluded.

Experimental Setup

Synthetic engagement scenarios

Real military laser scans are restricted and thus we use OpenFlight [26] to simulate highly credible air-to-ground missile engagement scenarios. All scenarios consider the missile flying at several altitudes and headings and under various pitch, roll and yaw angles. In addition, the missile-target range is variable while occlusion, clutter, Gaussian noise and atmospheric disturbances interfering the laser propagation are considered. It is worth noting that compared to [15,17,18,23] our scenarios are more realistic and challenging since they are affected by a greater number of parameters. The parameters per scenario are shown in Table 2.

Evaluation criteria

Currently, 3D descriptors are evaluated based on a precision-recall curve [2,4,8,14,21] that relies on the TP, FP, TN, FN matches. This curve is created by varying the kNNDR threshold value in the region of [0,1] during the feature matching process [4], which for

this paper this is not possible as we do not exploit a variable kNNDR threshold but rather reject matches based on Geometric consistency checks.

Similarly to [4] we do not exploit the simplistic case of just comparing the ATR prediction state with the actual state e.g. claiming a TP match if both states coincide, but consider as well the Euclidean distance based translational error T_{error} between the ground truth and the estimated target position in the scene. Estimating accurately the position of the target within the scene is mandatory for missile applications so that the missile can navigate itself towards the target. Therefore, we define as a TP match the case where the algorithm provides a transformation hypothesis for a scene in which a target does exist and T_{error} is less than 2-meters. The 2-meter T_{error} threshold is linked to the effective miss-distance range of a missile. Accordingly, for a FP match the algorithm provides a hypothesis for a scene in which a target does not exist or the scene does include a target but T_{error} is more than 2-meters. Finally, for a FN match the algorithm does not provide a hypothesis for a scene that contains a target and for a TN rejection, the algorithm does not provide a hypothesis for a scene that does not contain a target.

It is worth noting that for our scenarios the TN rejection case is not always applicable because in several runs per scenario the target is always present. Therefore, we avoid biasing the ATR performance of each descriptor by exploiting the F1-score that still encapsulates precision and recall information in a single value:

$$F1-score=\frac{2\#TP}{2\#TP+\#FP+\#FN} \tag{3}$$

where # denotes the number of the metric that follows i.e. *TP*, *FP* and *FN*.

Experiments

The 3D ATR pipeline is implemented in MATLAB while descriptors are implemented either in MATLAB or in C++ using a MEX wrapper. The parameters of each descriptor are fixed either to the ones originally proposed by their authors or to their point cloud library implementation [1,13,14]. For better readability, besides the scale invariance trials, noise, subsampling and the detailed performance metrics are investigated only for the benchmark Scenario 3.

Scenarios 1-3

Scenario 1 considers the 50-meter missile-target range. All descriptors evaluated excel at 15°-75° obliquity because the target pose provides distinctive details for description. In contrast, in the 0° obliquity case (side view), the MBT's features lack of distinctiveness and clutter objects interfere the missile-target Line-Of-Sight. Despite that, HoD-S, FPFH, 3DSC and RoPS still manage close to 0.9 F1-score, while the rest competitors up to 0.72 (Figure 6 (a)).

Scenario 2 doubles the missile-target range to 100-meters. This scenario is challenging, as the size of all objects in the scene and the resolution are half the one of scenario 1. For the 15°-75° obliquity, HoD-S, HoD and 3DSC excel with the rest of the descriptors closely following. In contrast, for the 0° case, the scale and resolution combined with occlusion and the non-distinctive MBT's features at that view impose a vast ATR performance reduction for all descriptors (Figure 6 (b)).

Scenario 3, considers the 200-meter missile-target range at 30° obliquity. This scenario is even more challenging as the missile-target range has quadrupled, further affecting the scene's size and resolution. Despite that, HoD-S and HoD are quite stable managing a 0.95 F1-score while 3DSC achieves 0.9 (Figure 6 (c)). This confirms [2] stating that it is very challenging to establish a robust LRF and thus neglecting it can be

beneficial. For example, the 3DSC is LRA based and is not robust to azimuthal rotation, but if 3DSC describes a keypoint *p* multiple times to compensate all possible 360° azimuthal rotations, then 3DSC can afford a high ATR performance. The number that 3DSC has to describe the same local patch depends on the size of the azimuthal bin size.

Processing efficiency

Despite being MATLAB implemented, the fastest 3D descriptor generated is the HoD-S due to overriding a LRF/A and having a small descriptor size. Next are HoD and FPFH as the former neglects a LRF/A while the latter has the smallest descriptor size among all competitors. Least efficient is RoPS due to its complex LRF algorithm and MATLAB implementation. The processing time per descriptor is shown in Figure 7 (a) and per scene in Figure 7 (b). The former plot is important as it highlights the processing burden that is purely implied by each descriptor. The latter plot considers the processing time of the entire ATR pipeline including the common modules, which require on average 1268ms. It should be noted that the relative ratios between Figure 7 (a) and (b) differ as each descriptor produces a different number of kNNDR matches and thus a number of Hypothesis to be tested. The computational breakdown of the common procedures is presented in Figure 7 (c) showing that Hypothesis Verification imposes the vast computational burden because of the RANSAC and ICP iterative processes.

Qualitative Matching performance

Although the 2-meter distance threshold that defines a *TP* match is sufficient, descriptors achieving a translational error T_{error} less than that can offer to the missile pinpoint targeting. Our trials show that RoPS affords the smallest average T_{error} closely followed by FPFH and USC. This is due to their LRF, provided that they have positively recognized the target. Since USC is LRF based while 3DSC is LRA, the

higher accuracy we demonstrate for USC is reasonable. Largest error, but still less than 0.5-meters is given by HoD-S due to neglecting a LRF (Figure 7 (d)).

Compactness

This metric reveals the descriptive power per element of the descriptor vector. Computer vision literature calculates compactness as the fraction of the area under the precision - recall curve divided by the number of elements that each descriptor has [10]. For the suggested 3D ATR architecture calculating the precision - recall curve is not realistic as we reject matches based on Geometric consistency checks. Therefore, we define compactness as:

$$compactness = \frac{F1 - score}{\#\, descriptor\ elements} \tag{4}$$

Figure 7 (e) clearly shows that HoD-S achieves highest compactness followed by FPFH. This happens as both descriptors combine an appealing F1-score with a small descriptor length.

Robustness to Sensor noise

The noise levels of real military scenarios are classified and therefore we investigate the robustness of each descriptor to the Gaussian noise levels suggested by the computer vision community. Compared to [2,4,14] that apply Gaussian noise with zero mean and σ up to $0.5\mathrm{Res}$, we consider up to $1\mathrm{Res}$ of the point cloud resolution of the benchmark Scenario 3, which is $\sigma = 30\,\mathrm{cm}$. Hence, we evaluate the ATR performance of the descriptor presented in Table 1 on Gaussian noise with zero mean and $\sigma = \{10,20,30\}$ cm.

The first experiment concerns $\sigma = 10\,\mathrm{cm}$ where HoD, HoD-S, 3DSC, USC, RoPS and SHOT are almost unaffected. Interestingly, the latter three have a stable poor

performance throughout scenario 3 as they cannot manage the combined scale and resolution changes. Next trial doubles noise, where both HoD variants, USC and 3DSC are still minor affected. In contrast, SHOT and RoPS exceed their noise invariance limits and therefore perform quite poor. Finally, noise triples to $\sigma = 30$ where all descriptors exceed their robustness capabilities. Despite that, HoD still gains a 0.5 F1-score. An example of a scene affected by Gaussian noise is presented in Figure 8 (b) while the noise affected recognition performance is shown in Figure 9 (a).

Laser propagation related effects

In this trial we simulate the laser atmospheric interferences of beam wandering and scintillation [27]. Both these effects reduce irregularly and randomly the number of reflected laser beams and thus affect the spatial location of the scene vertices. Therefore, trials involve non-uniform and random subsampling of the scene point clouds of Scenario 3 to $\{1/2, 1/8, 1/16\}$ the original scene point cloud ***P***. This is novel as current literature considers only uniform subsampling where the template and the target are at the same scale [2,14,21].

For the $1/2$ case, most robust are HoD and HoD-S followed by 3DSC due to neglecting the prone to perturbations LRF [2]. For the $1/8$ case all competitors except HoD and HoD-S fail to gain an appealing performance because the simultaneous combination of scale, resolution and subsampling change exceeds the limits of a repeatable LRF/A. For the $1/16$ case all competitors fail as the combined subsampling and scale change are quite excessive. Despite that, it is worth noting that both variants of HoD are the only ones with F1-score close to 0.5. An example of a scene affected by atmospheric perturbations is presented in Figure 8 (c) while detailed recognition performance is shown in Figure 9 (b).

Overall assessment

Figure 10 presents the overall performance per descriptor based on all trials along with the corresponding computational requirement of each descriptor. Highest performance is gained by HoD closely followed by HoD-S because they neglect a LRF/A and thus offer a more stable local encoding in such an extended combination of disturbances. This conclusion is enforced by the fact that the second best preforming descriptor is 3DSC that relies on a LRA. From the computational aspect, both HoD variants along with FPFH are the most efficient. The performance of each descriptor is explained as:

- FPFH: Its LRF relies on the immediate neighbours of each keypoint and therefore is prone to disturbances confirming [14].
- RoPS: Current trials combine disturbances that exceeded the invariance limits of its LRF, confirming that it is very challenging to establish a robust LRF.
- USC, 3DSC: These are robust because their description process relies on a normalized weighted point-count per description bin.
- SHOT: It involves normal estimation based on small groups of points within the description radius and therefore SHOT is prone to the large disturbances investigated this paper.
- HoD, HoD-S: The normalized histogram of the point-pair distances encoded can withstand large and combined nuisances because it is matched to the requirements of each scene.

Conclusion

In this paper, we evaluate a number of current computer vision oriented 3D local feature based descriptors in the context of future LIDAR based missiles with ATR capabilities. Experiments are conducted on a computer vision 3D ATR architecture that is properly

extended to facilitate the requirements of missile engagement scenarios.

Trials are performed on a number of synthetic but highly realistic and credible air-to-ground scenarios and consider the case of a single template scheme while the missile has 6-Degrees-of-Freedom motion capability and is flying at various altitudes, obliquities and missile – target distances. The realistic nature of the trials is enhanced by simulating atmospheric perturbations that affect the LIDAR's laser beam. Under these conditions, all 3D descriptors are evaluated for their ATR performance processing efficiency, scale change, qualitative matching accuracy, compactness and their robustness to atmospheric effects.

Overall performance in terms of ATR capability and of processing efficiency clearly shows that HoD-S can be a promising option towards future LIDAR based missiles with ATR capabilities.

Acknowledgements: The authors would like to thank MBDA UK for providing the evaluation scenarios.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Kechagias-Stamatis O, Aouf N. Histogram of distances for local surface description. 2016 IEEE Int. Conf. Robot. Autom. IEEE; 2016. p. 2487–2493.

[2] Salti S, Tombari F, Di Stefano L. SHOT: Unique signatures of histograms for surface and texture description. Comput. Vis. Image Underst. 2014;125:251–264.

[3] Johnson AE, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Trans. Pattern Anal. Mach. Intell. 1999;21:433–449.

[4] Guo Y, Sohel F, Bennamoun M, et al. Rotational Projection Statistics for 3D Local Surface Description and Object Recognition. Int. J. Comput. Vis. 2013;105:63–86.

[5] Guo Y, Sohel F, Bennamoun M, et al. A novel local surface feature for 3D object recognition under clutter and occlusion. Inf. Sci. (Ny). 2015;293:196–213.

[6] Rusu RB, Blodow N, Beetz M. Fast Point Feature Histograms (FPFH) for 3D registration. 2009 IEEE Int. Conf. Robot. Autom. IEEE; 2009. p. 3212–3217.

[7] Frome A, Huber D, Kolluri R, et al. Recognizing Objects in Range Data Using Regional Point Descriptors. ECCV. 2004. p. 224–237.

[8] Tombari F, Salti S, Di Stefano L. Unique shape context for 3d data description. Proc. ACM Work. 3D object Retr. - 3DOR '10. New York, New York, USA: ACM Press; 2010. p. 57.

[9] Aldoma A, Marton Z-C, Tombari F, et al. Tutorial: Point Cloud Library: Three-Dimensional Object Recognition and 6 DOF Pose Estimation. IEEE Robot. Autom. Mag. 2012;19:80–91.

[10] Guo Y, Bennamoun M, Sohel F, et al. 3D Object Recognition in Cluttered Scenes with Local Surface Features: A Survey. IEEE Trans. Pattern Anal. Mach. Intell. 2014;36:2270–2287.

[11] Bariya P, Novatnack J, Schwartz G, et al. 3D geometric scale variability in range images: Features and descriptors. Int. J. Comput. Vis. 2012;99:232–255.

[12] Taati B, Greenspan M. Local shape descriptor selection for object recognition in range data. Comput. Vis. Image Underst. 2011;115:681–694.

[13] Alexandre L. 3D descriptors for object and category recognition: a comparative evaluation. Work. Color. Camera Fusion Robot. IEEE/RSJ Int. Conf. Intell. Robot. Syst. 2012;1.

[14] Guo Y, Bennamoun M, Sohel F, et al. A Comprehensive Performance Evaluation of 3D Local Feature Descriptors. Int. J. Comput. Vis. 2016;116:66–89.

[15] Vasile A, Marino R. Pose-independent automatic target detection and recognition using 3D laser radar imagery. Lincoln Lab. J. 2005;15:61–78.

[16] Grönwall C. Ground object recognition using laser radar data: geometric fitting, performance analysis, and applications. Linköping, Sweden; 2006.

[17] Xiaofeng Li XL, Jun Xu JX, Jijun Luo JL, et al. Ground target recognition based on imaging LADAR point cloud data. Chinese Opt. Lett. 2012;10:S11002-

311005.

[18] Roy S, Maheux J. Baseline processing pipelilne for fast automatic target detection and recognition in airborne 3D ladar imagery. Sadjadi FA, Mahalanobis A, editors. Autom. target recogntion XXI. 2011;8049.

[19] Tombari F, Salti S, Di Stefano L. Performance evaluation of 3D keypoint detectors. Int. J. Comput. Vis. 2013;102:198–220.

[20] Katz S, Tal A, Basri R. Direct visibility of point sets. ACM Trans. Graph. 2007;26:24.

[21] Guo Y, Sohel F, Bennamoun M, et al. TriSI : A Distinctive Local Surface Descriptor for 3D Modeling and Object Recognition. 8th Int. Conf. Comput. Graph. Theory Appl. Barcelona, Spain; 2013.

[22] Muja M, Lowe DG. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. Int. Conf. Comput. Vis. Theory Appl. (VISAPP '09). 2009;1–10.

[23] Gronwall C. Ground target recognition using rectangle estimation. IEEE Trans. Image Process. 2006;15:3400–3408.

[24] Aldoma A, Tombari F, Di Stefano L, et al. A Global Hypotheses Verification Method for 3D Object Recognition. Comput. Vision--ECCV 2012. Springer; 2012. p. 511–524.

[25] Chen H, Bhanu B. 3D free-form object recognition in range images using local surface patches. Pattern Recognit. Lett. 2007;28:1252–1262.

[26] Presagis. OpenFlight visual simulation [Internet]. [cited 2016 Aug 1]. Available from: http://www.presagis.com/products_services/standards/openflight/.

[27] Dios F, Rodrigues A, Comeron A. Scintiallation and beam-wander analysis in an optical ground station-satellite uplink. Appl. Opt. 2004;43:3866–3873.

Nomenclature

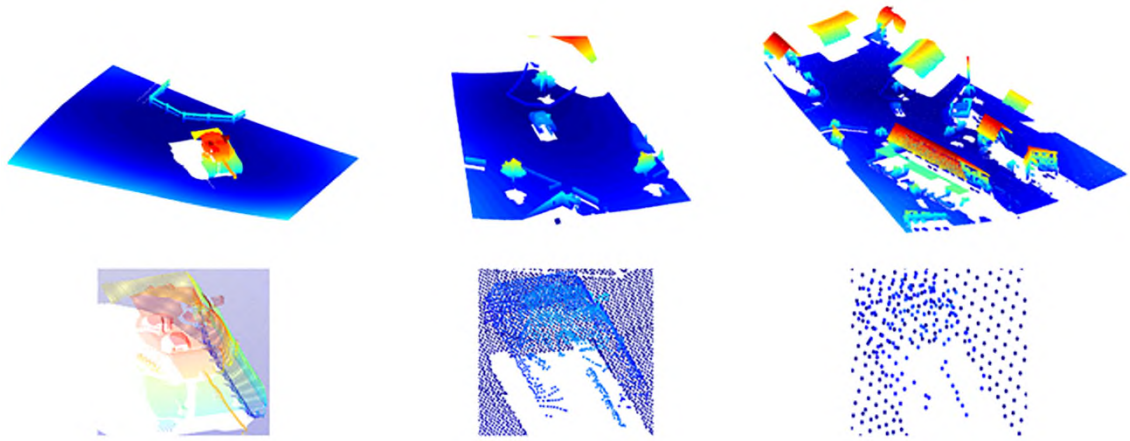
3DSC	3D Shape Context	LIDAR	Light Detection and Ranging
ATR	Automatic Target Recognition	LRF/A	Local Reference Frame / Axis
FLANN	Fast Library for Approximate Nearest Neighbours	MBT	Main Battle Tank
FN	False Negative match	RANSAC	Random Sample Consensus
FP	False Positive match	RoPS	Rotational Projection Statistics
FPFH	Fast Point Feature Histograms	SHOT	Signatures of Histograms of Orientations
HoD	Histogram of Distances	TN	True Negative match
HoD-S	Histogram of Distances - Short	TP	True Positive match
ICP	Iterative Closest Point	TriSI	Tri-Spin Images
kNNDR	k-Nearest Neighbour Distance Ratio	USC	Unique Shape Context

Table 1. 3D descriptors evaluated

Descriptor	**Descriptor Length**	**Basic concept**	**LRF/A**	**Implementation platform**
SHOT	352	Angular variations	LRF	C++
USC	1980	Accumulating points	LRA	C++
HoD-S	40	L2-norm distances	---	MATLAB
HoD	240	L2-norm distances	---	MATLAB
FPFH	33	Angular variations	LRF	C++
3DSC	1980	Accumulating points	LRA	C++
RoPS	135	Low order statistics	LRF	MATLAB

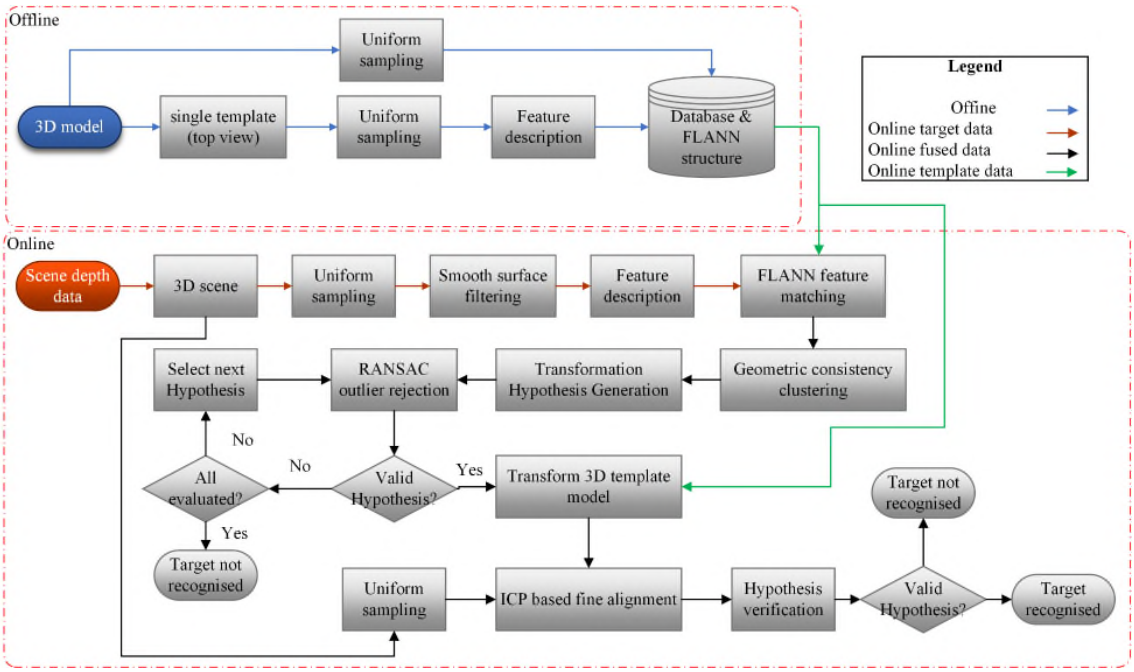
Table 2. Parameters per Scenario

Scenario	1	2	3
Runs	6	6	1
Obliquity (°)	0°–75° per 15°	0°–75° per 15°	30°
Range (m)	50	100	200
Resolution (cm)	11	18	30
Scenes with target/out of total	334/345	327/364	78/78

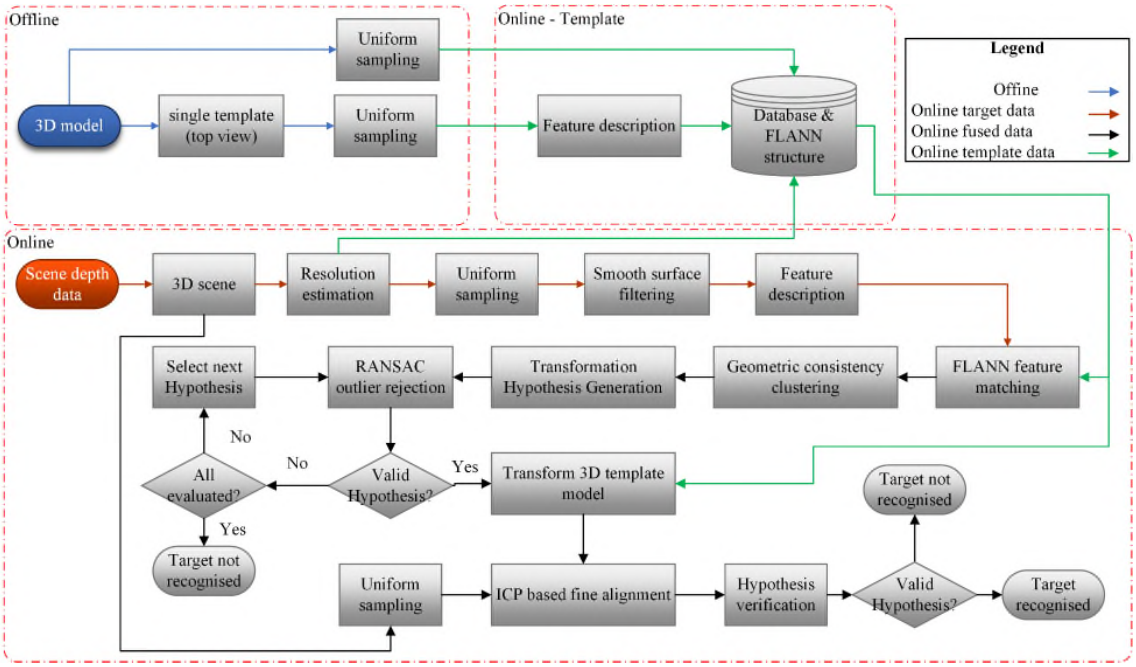


(a) (b) (c)

Figure 1. Example showing the distance related scene resolution, top row shows the scene enclosed in the missile's seeker Field-of-View and bottom row the corresponding MBT point cloud patch at a missile-target range of (a) 50-meters (b) 100-meters (c) 200-meters



(a)



(b)

Figure 2. 3D ATR pipeline for the descriptors evaluated (a) SHOT, USC, FPFH, 3DSC and RoPS (b) HoD and HoD-S

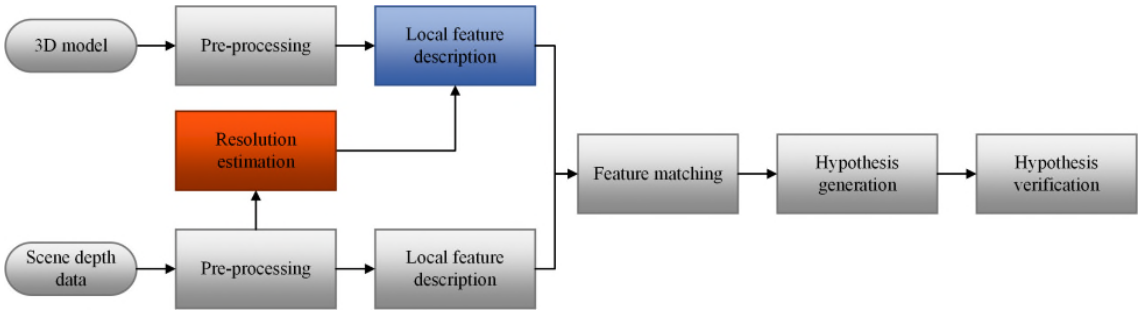
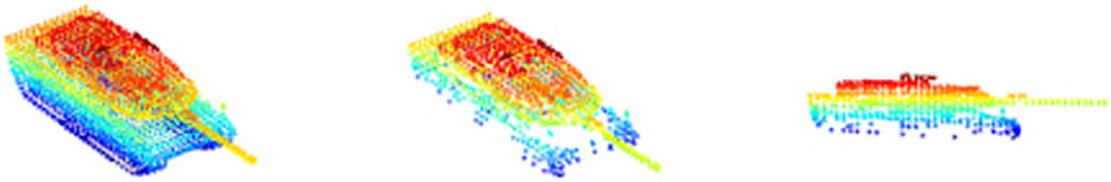
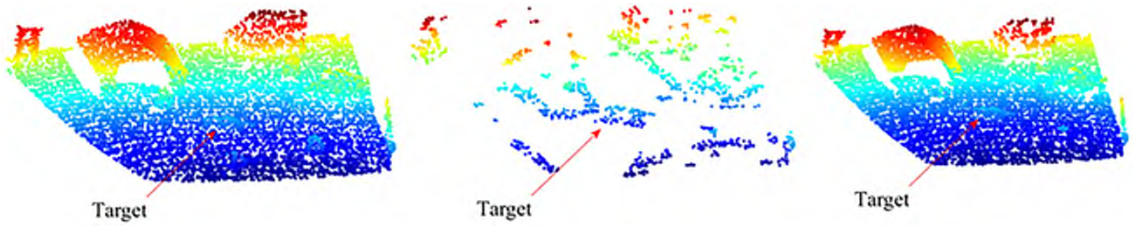


Figure 3. Diagram showing the differences highlighted in colour between the ATR algorithm used for SHOT, USC, FPFH, 3DSC, RoPS and both HoD variants. HoD and HoD-S include an online template description scheme rather than an offline, which is the norm (in blue) and an online scene resolution calculation (in red).



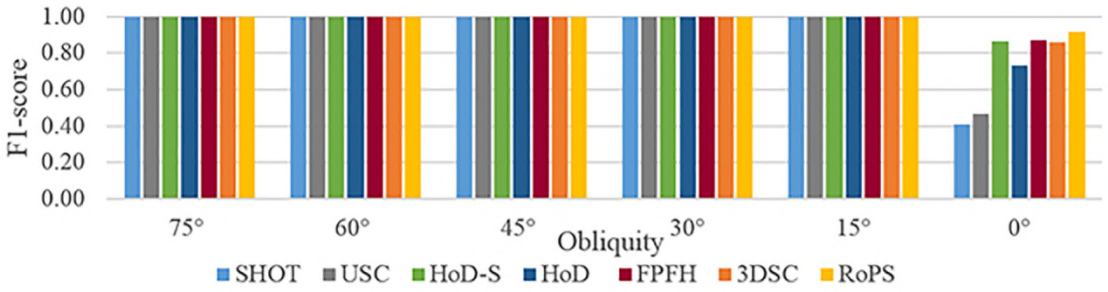
(a) (b)

Figure 4. Colour coded MBT (a) ideal point cloud (b) HPR processed

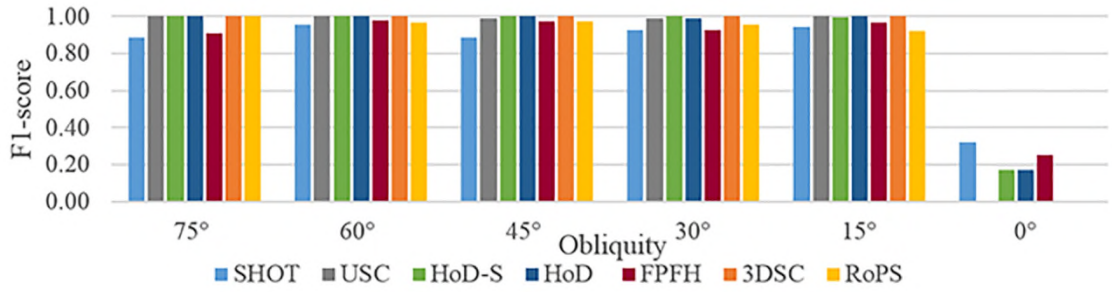


(a) (b) (c)

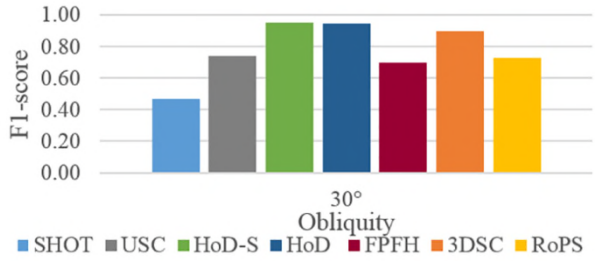
Figure 5. (a) LIDAR point cloud with σ=10cm Gaussian noise (b) proposed standard deviation filtering (c) average smooth surface filtering (height-related colour coding for better visualization)



(a)



(b)



(c)

Figure 6. Recognition performance under various obliquities for (a) scenario 1 (b) scenario 2 and (c) scenario 3

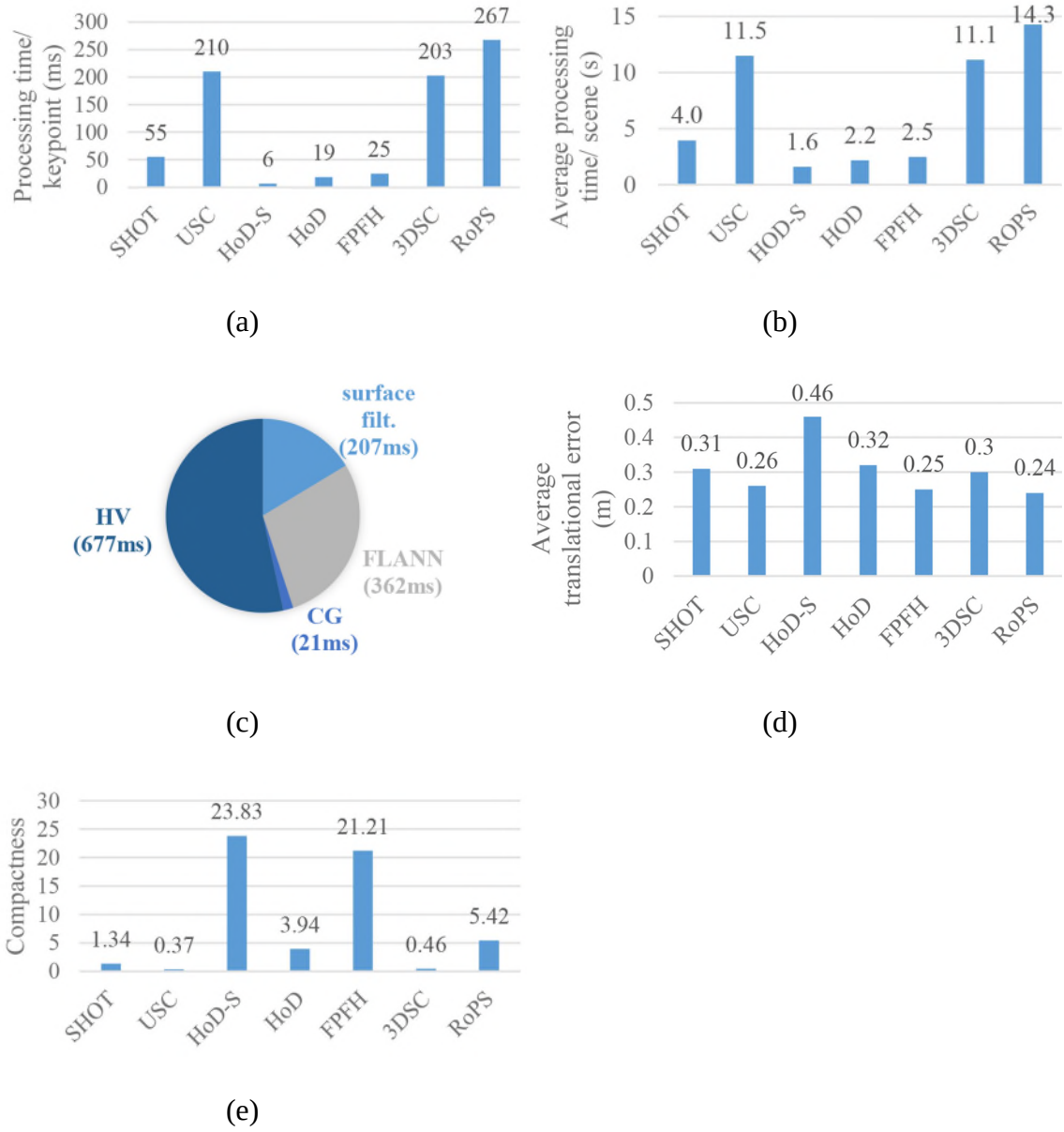
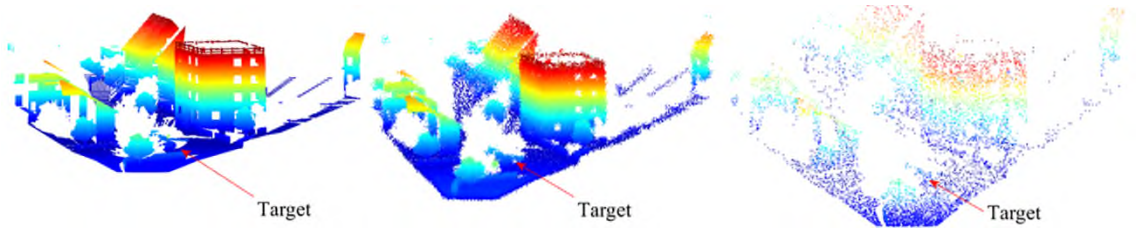
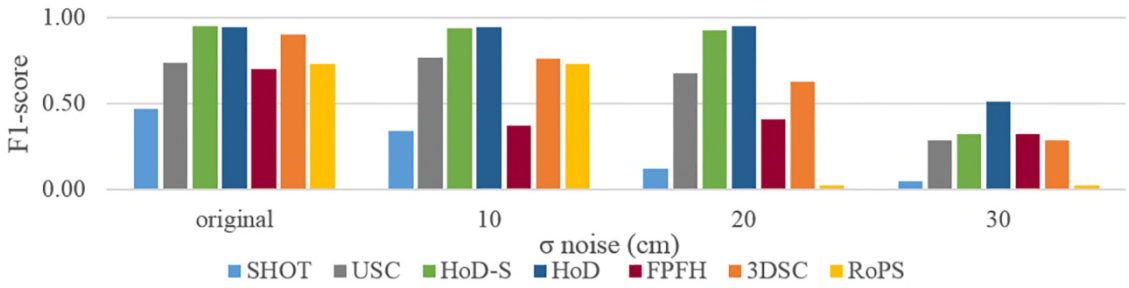


Figure 7. Performance metrics (a) processing efficiency (b) average processing time per descriptor per scene (c) computational breakdown in milliseconds excluding description time (d) average translational error (e) compactness

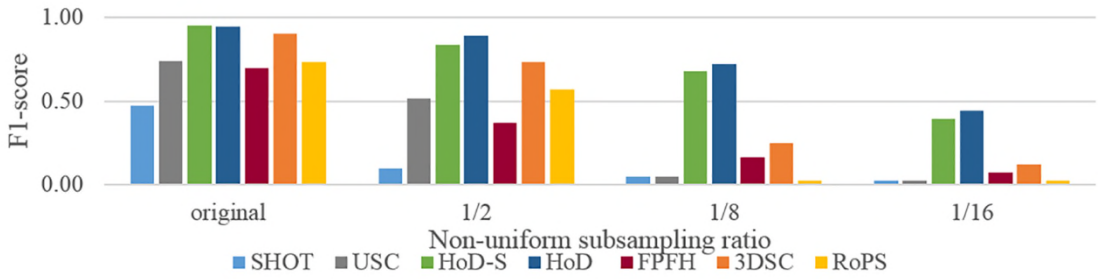


(a) (b) (c)

Figure 8. Example of a point cloud scene at 200-meters missile – target range (a) ideal (b) with σ=20cm Gaussian noise (c) with 1/16 non-uniform subsampling

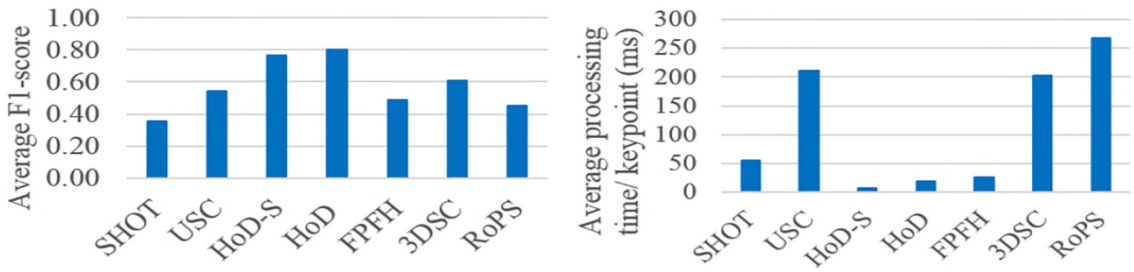


(a)



(b)

Figure 9. Robustness to (a) noise (b) non-uniform subsampling



(a) (b)

Figure 10. Overall performance of the 3D descriptors evaluated (a) recognition performance (b) processing time