

A hamming distance and spearman-correlation based star identification algorithm

Mehta Deval Samirbhai*, Shoushun Chen*, *Senior Member, IEEE* and Kay Soon Low[#], *Senior Member, IEEE*

*Nanyang Technological University, [#]National University of Singapore

Abstract – This paper presents a novel star identification algorithm for a “Lost-In-Space” mode star tracker. The spearman correlation approach provides reliable recognition even when the captured images are swayed and biased from the onboard star pattern database. The hamming distance approach provides a shortlisted list of star IDs. Thus, the proposed combination of hamming distance and spearman correlation provides a reliable and fast recognition. The achievable performance is evaluated by testing on simulated and real images.

Index Terms— star identification; lost-in-space mode; hamming distance; spearman correlation; star tracker

I. INTRODUCTION

Satellite missions require a high accuracy for determining their orientation in space. Thus, developing an accurate and reliable attitude determination system is one of the most important requirements for satellite missions. Sun sensor, magnetometer, and earth sensor are some of the attitude sensors that have been developed in the past decades for determining the orientation of the satellite [1]. Star trackers, also called as star sensors have become very popular for attitude determination in the last two decades because they provide accuracy of orientation down to arc seconds [2], [3].

A star tracker consists of an image sensor for capturing the image of the stars within a defined field of view (FOV) and an on-board stored star pattern database (SPD), which is prepared offline. It operates in two modes: Lost-In-Space (LIS) mode when no initial attitude information is available and later it shifts to the tracking mode. In the LIS mode, after acquiring the image of the stars, firstly, star positions are determined in the image. Later, stars are identified in the image and finally, by comparing the star vectors obtained from the image with the reference star catalog, attitude is calculated. This sequence of events is called centroiding process, star identification, and

attitude determination respectively. Of the above events, star identification is the most important part of a star tracker. In the past few decades, achieving a reliable and fast star identification has evolved into a challenging problem because of the magnitude uncertainty, false stars, and distortion present in the image captured. Hence many star pattern recognition techniques have been proposed to solve the above problems and identify the stars correctly in the image over the past two decades [4]-[21].

Basically, star pattern recognition techniques adopt two different approaches: geometric approach and pattern based approach. The former approach visualizes small clusters of stars in the image and uses features such as inter-star angle, distances between two stars for identification. One of the earliest techniques [4] adopting this approach utilized the features formed by a three-star cluster for identification. Later, Mortari *et al.* [5] extended this idea to a four-star cluster. An approach was also suggested in [6] to utilize the area formed by the stars as a feature for identification. In recent years, geometric voting technique [7] suggested considering the distances of all the stars that lie within the FOV of the reference star for identification. A probabilistic and iterative approach utilizing these geometric distances was also recently suggested in [8].

The pattern based approach was first proposed in the grid algorithm [9]. In this type of approach, a well-defined pattern code is formed by utilizing positions of all the neighboring stars which lie in the FOV of the reference star. The same operation is applied to the image and the pattern code generated from the image is compared with the patterns stored in the SPD. This approach was widely adopted by the research community and many star identification algorithms have been developed based on this approach [10]-[14]. The technique in [10] proposed to construct the pattern code from translation and rotation invariance of the nearest k -neighboring stars. A modified grid algorithm was also developed in [11] which made the existing grid algorithm more robust by using a polar grid and multiple reference stars.

In addition to providing a reliable and robust recognition, star pattern recognition techniques are also expected to provide a high speed of recognition. Some star recognition techniques [15]-[18] were developed specifically for achieving the purpose of fast recognition. Mortari *et al.* [15] proposed

This paper titled “A hamming distance and spearman-correlation based star identification algorithm” was accepted for publication on April 23, 2018. This work was supported in part by the S13-1109-NRFOSTIn-SRP.

Author’s address: School of Electrical & Electronic Engineering, Nanyang Technological University, Block S2, 50 Nanyang Avenue, Singapore 639798, Singapore. E-mail: (meht0004@e.ntu.edu.sg)

the idea of k -vector which did not require to search for the pattern or features for comparison. An optimized search tree database construction was also proposed in [16], [17] for fast recognition.

In this paper, firstly, we present a study performed on *real images* to identify and analyze the problems associated with attaining a high recognition reliability for star pattern recognition techniques. To solve the identified problems, we then propose a novel combination of hamming distance and spearman-correlation to achieve a reliable as well as a fast star identification. The proposed idea is to construct a discrete signal from the features extracted from the star image. We present a generalized framework for the process of feature extraction and construction of Look-Up-Table (LUT) and SPD for the proposed technique. Later, with the help of an illustrative example, we describe the process of using hamming distance for shortlisting the star IDs and spearman correlation for identifying the correct star ID from the shortlisted list. The hamming distance approach leads to fast recognition and the spearman correlation helps in achieving a high recognition reliability. The subsequent sections describe the details of the research problem, proposed approach, simulation and testing, and benchmarking with state-of-the-art star pattern recognition techniques. The paper is concluded by analyzing the performance of the proposed technique.

II. RESEARCH PROBLEM

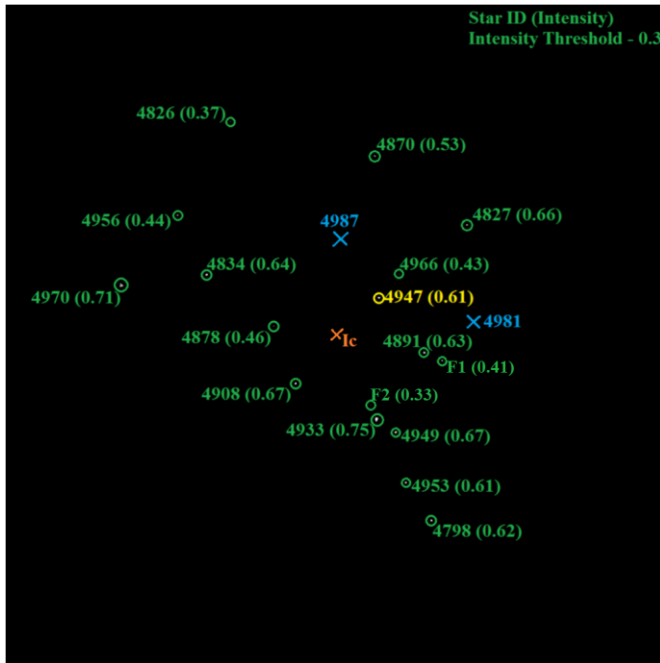


Fig. 1. A real image from the dataset of SST-20S star tracker showing star ID and their corresponding intensity.

To identify the problems associated with attaining a high recognition reliability for star pattern recognition techniques, we perform a study on 1019 identified real images captured by a star tracker SST-20S currently mounted on a low earth orbit satellite, VELOX-CI [22]. The specifications of the SST-20S star tracker and the real image are listed in TABLE I. For explaining the process of identifying and analyzing the

problems, we will be using an identified real image as shown in Fig. 1. In Fig. 1, the center of the image is marked with orange and labeled as I_c . The star nearest to the center of the image is chosen as a reference star (star ID: 4947, marked with yellow). The other stars present in the image are marked in green and labeled by their star ID (intensity) in the image. The intensity in the image varies from 0 (black) to 1.0 (255) being the white intensity. The intensity of each star in Fig. 1 is the mean of the intensity of the pixels occupied by that star in the image. As can be seen from Fig. 1, the lowest intensity is that of star F2 (0.33) and the highest intensity is that of star ID 4933 (0.75). The standard deviation (σ) of all the star intensities is 0.122. The maximum, minimum, and standard deviation of the star intensities for all the 1019 images are similar to those for the example in Fig. 1. Thus, the intensity threshold was chosen to vary from 0.3 to 0.6 in the steps of 0.1 for the three different analysis performed on an image, namely – the number of nearest neighboring stars to the reference star, false stars in the image, and the amount of patch mismatch in the image.

TABLE I

Specifications of the SST-20S star tracker and the real image

SST-20S star tracker	
Parameter	Value
Field of View (FOV)	$15^\circ \times 15^\circ$
Sensitivity (M_v)	6.3
Star catalog	SAO J2000 [23]
Resolution ($w \times h$)	1024 x 1024 pixels
Pixel size (ρ)	13 μm
Focal length (f)	50 mm
Bits per pixel	8
Position accuracy	40 arcsec

We will now explain the calculation to check how many nearest stars to the reference star are present with different intensity thresholds applied to the image using Fig. 1. For an intensity threshold of 0.3 in Fig. 1, the reference star ID 4947 has only two nearest stars (out of four) present i.e. the star ID 4966 and star ID 4891. Star IDs 4987 and 4981 (both marked by a blue cross) are present in the SPD of star ID 4947, however, they are missed in the image captured. If we increase the intensity threshold to 0.5 in Fig. 1, then star ID 4966 will also cease to be present in the image, which will imply only one nearest star (out of four) is present for star ID 4947 in the image. A similar analysis was done for all the 1019 images and results are shown in Fig. 2. As can be seen from the results in Fig. 2, even for a low-intensity threshold of 0.3, the cases that all the *four* nearest stars are present are still less (only 344 out of 1019 images). There were only half of the total images (615 out of 1019) which had three nearest neighboring stars in the image for intensity threshold of 0.4. The number of cases (22 out of 1019) when none of the nearest four neighboring stars are present in the image is low for the intensity threshold of 0.3 compared to that of higher intensity thresholds of 0.4, 0.5 and 0.6. If the intensity threshold is increased, the cases of having all the *four* nearest stars to the reference star in the image decreases. Star pattern

recognition techniques such as Liebe, Pyramid, and Oriented triangles [4]-[6] use the features constructed from the *nearest* stars to the reference star for identification. Clearly, these techniques will fail in the cases even when one of the *nearest* stars is missing from the image. As can be seen from the study in Fig. 2, this happens quite often in the real images. This problem occurs because of the *magnitude uncertainty* of the star tracker which results in the low magnitude nearest stars to be missed in the image.

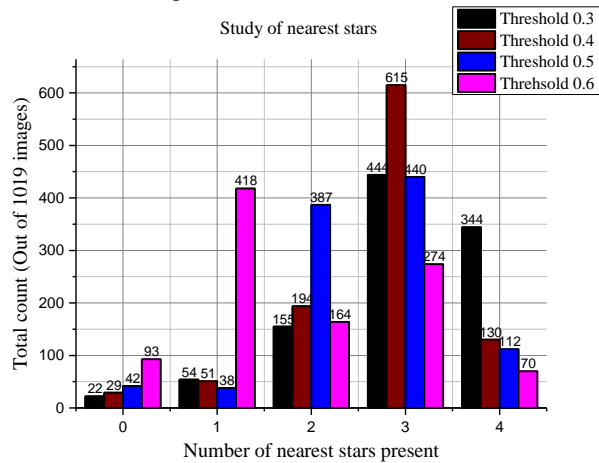


Fig. 2. Study on real images for nearest stars.

The second study performed is to check the number of false stars present in the image with varying intensity threshold. Planets, meteors, nearby satellites, space debris can be wrongly identified as stars while capturing the image in space. For the case in Fig. 1, at an intensity threshold of 0.3, there are two false stars present in the image labeled as F1 and F2. If the intensity threshold is increased to 0.4, only one false star F1 will remain in the image; and if the intensity threshold is further increased to 0.5, both the false stars F1 and F2 will cease to exist in the image. A similar analysis was performed for all the images and the results are shown in Fig. 3.

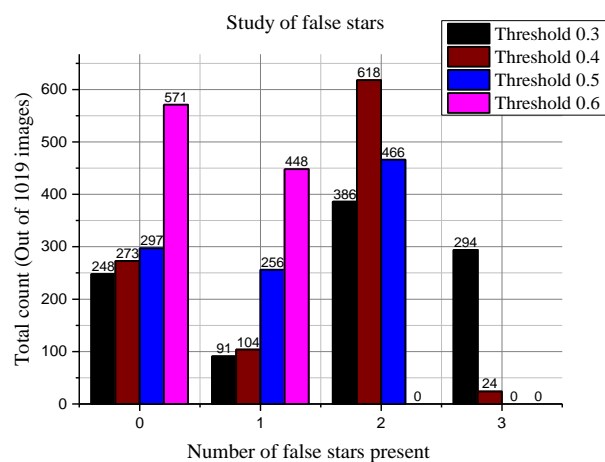


Fig. 3. Study on real images for false stars.

As can be seen from Fig. 3, irrespective of a lower or higher intensity threshold, there is quite often a scenario when at least one false star is present in the image. If the intensity threshold is high (0.6), there are less number of cases of having false

stars in the image. In fact, when the intensity threshold is 0.6, there is not a single case when there is more than one false star in the image. If the intensity threshold is low (say 0.3), there are as many as 294 cases when *three* false stars are present in the image. This problem of false stars in the image can add to the earlier mentioned problem of missing of nearest stars, thus elevating the difficulty of correct star identification.

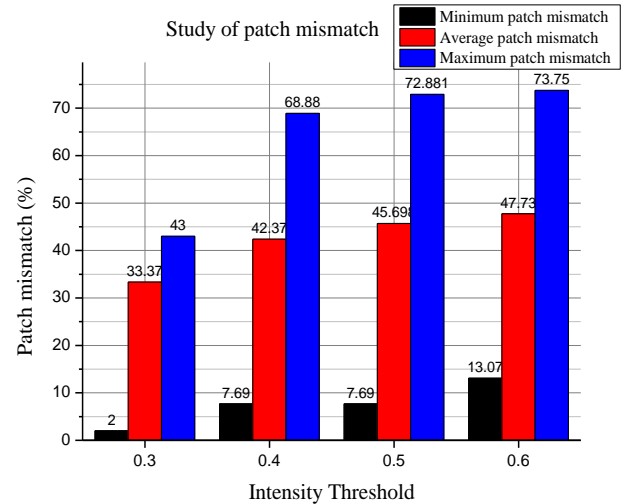


Fig. 4. Study on real images for patch mismatch.

The final study aims to analyze the amount of patch mismatch present in the image. The image captured from the star tracker is translated and rotated as compared to the onboard SPD. The above situation leads to a certain number of stars been missed in the FOV of the reference star (in the image) when compared to the number of stars in the FOV of the reference star (in the SPD). The number of stars missed is termed as the amount of patch mismatch and is calculated in patch mismatch (%). For the example in Fig. 1, the reference star ID 4947 has 24 neighboring stars in its SPD. However, because the reference star does not lie exactly at the center of the image only 14 (excluding the two false stars) of those 24 stars are present in the FOV of the star 4947. Thus, the patch mismatch (%) for star 4947 is around 41% for intensity threshold of 0.3. If the intensity threshold is increased to 0.4 (and later to 0.5 and 0.6) the number of stars in the FOV of the reference star will decrease which results in a further increase of the patch mismatch. This problem of patch mismatch was also reported in [10]-[13] and its visualization is shown later in this paper in Fig. 9. In Fig. 4, we show the results of the analysis of the patch mismatch problem with varying intensity threshold on the real images. It can be seen from Fig. 4 that even with a low-intensity threshold, the average amount of patch mismatch (or) missing stars is 33.37%. If the intensity threshold is increased to 0.6, the number of stars missed will be more and thus the average patch mismatch rises to 47.73% in this case. One of the images out of the 1019 real images, registered a patch mismatch as high as 73.75% with intensity threshold of 0.6. Star pattern recognition techniques such as geometric voting [7], grid algorithm [9], [11], search tree with optimized database [16], which use the number of neighboring stars in the image as a criterion for comparison will fail drastically because of the problem of patch mismatch.

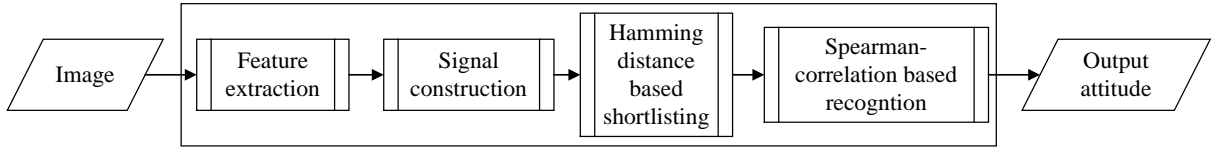


Fig. 5. Overall flow of proposed combination of hamming distance based shortlisting and spearman correlation based star recognition.

The brightness of the stars can also be utilized as one of the features and a technique proposed in [19] in fact selects its reference star based upon brightness. Some techniques [20], [21] also proposed to utilize the magnitude of the stars for identification. However, we can conclude from the study performed that the patch mismatch and number of false stars varies with intensity threshold. Decreasing the intensity threshold might decrease the patch mismatch gently, but at the same time, the number of false stars in the image inflate and vice-versa. Thus, utilizing the intensity of the stars is not a reliable characteristic for comparison. It was also studied in [24], [25] that the techniques which utilize the brightness of stars are more prone to failure because the order of brightness of stars in the image and the SPD differs. Thus, we analyzed from the above study performed that the patch mismatch, magnitude uncertainty, and false stars are the major problems faced in achieving a high recognition reliable performance for star pattern recognition techniques. Moreover, the intensity of the stars is not a reliable parameter for comparison purposes.

In addition to providing a high recognition reliability, the star pattern recognition techniques are also required to be fast. The star pattern recognition techniques which consider every star in the FOV will be slower as compared to the ones which consider only some selected stars (say *nearest* three or four stars) for identification. However, the latter case star pattern recognition techniques may not be as recognition reliable as the former case star pattern recognition techniques.

Having identified the problems faced by the star pattern recognition techniques and analyzed the existing approaches, we propose a star pattern recognition technique which does not rely on the brightness of the stars. It selects and identifies the star nearest to the center of the image. A discrete signal is constructed from the features extracted from the neighboring stars of this center star in the image. Before applying the robust approach of spearman correlation for star identification, hamming distance is used for shortlisting the star IDs. The complete algorithm flow is shown in Fig. 5.

The rest of the paper is structured as follows. In Section III, we describe the feature extraction and the proposed discrete signal construction utilizing the features extracted. In this section, we also explain the generalized framework for the construction of the look-up-table (LUT) and star pattern database (SPD) for the proposed technique. In Section IV, we describe the process of hamming distance used for shortlisting the star IDs and spearman correlation for identifying the stars correctly in the image. Section V deals with the simulation results of the proposed technique and comparison with state-of-the-art star pattern recognition techniques. The performance of the proposed technique on real images is also shown in Section V. Finally, we conclude the paper in Section VI.

III. FEATURE EXTRACTION, AND SIGNAL, LOOK-UP-TABLE AND STAR PATTERN DATABASE CONSTRUCTION

A. Feature extraction

The proposed technique uses two rotation and translation invariant features – euclidean distances from the center star and the relative angle that the consecutive neighboring stars make with the center star, as shown in Fig. 6. We use the star catalog SAO J2000 [23] and every star ID which has a relative magnitude M_v less than 6.0 is considered for forming the LUT and the SPD for the proposed star recognition technique. Each star ID is considered at the center of the FOV and all the neighboring stars which lie within the pattern radius (the equivalent of the FOV in terms of the number of pixels) from the center are considered for extracting the above-mentioned two features. The relation between the pattern radius (PR) and the FOV is given by (1). Both the extracted features are calculated by (2) and (3) respectively.

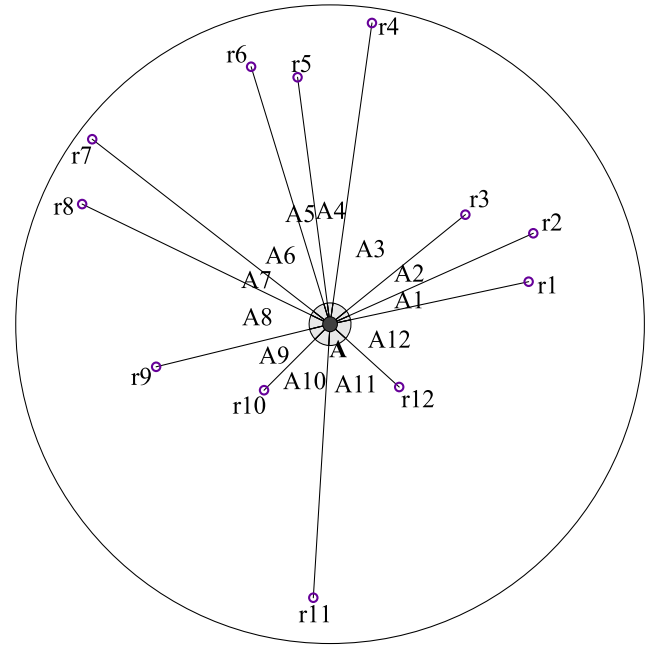


Fig. 6. Feature extraction for the proposed technique.

$$FOV = 2 \tan^{-1} \left(\frac{PR}{2(f/\rho)} \right) \quad (1)$$

where FOV is the field of the view of the star tracker, f is the focal length, ρ is the pixel dimension, and PR is the maximum distance in terms of pixels from the center.

$$r_n = \sqrt{(x_n - w/2)^2 + (y_n - h/2)^2} \quad (2)$$

$$A_n = \tan^{-1} \frac{(y_n - h/2)}{(x_n - w/2)} \quad (3)$$

where x_n and y_n are the pixel co-ordinates of the star in the image, and w and h are the dimensions of the image sensor.

B. Signal, LUT and SPD construction

1) Signal construction

As explained in Fig. 6, there are two features extracted – relative angles and euclidean distances. We propose that the relative angles extracted can be visualized as time domain and the corresponding radial distances can be visualized as the amplitude of the signal. For example, in Fig. 6, there are twelve stars surrounding the reference star A. Thus, there are twelve radial distances and relative angles, namely – {r1, A1, r2, A2...r12, A12}. Every radial distance rn corresponds to its relative angle An i.e. r1 corresponds to A1, r2 corresponds to A2 and similarly r12 corresponds to A12. Thus, we construct a signal containing data points as shown in Fig. 7, which depicts the star pattern of the reference star A. As shown in Fig. 7, the signal has relative angles (An) in an increasing order of their value on the x-axis (visualized as time domain) and their corresponding radial distance value (rn , visualized as amplitude). The point Sn in Fig. 7 corresponds to the star n, with its coordinates being (An , rn). For example, A5, the smallest angle in Fig. 6 and its corresponding amplitude r5 is represented as the point S5 in Fig. 7. The maximum relative angle between two stars in the image can be 180° or π radians. The maximum radial distance between any of the neighboring star from the center star can be PR pixels. This constructed signal is now utilized for LUT and SPD preparation as explained later in this section.

2) Look-Up-Table (LUT) construction

The purpose of the LUT construction is for shortlisting the star IDs which seem to be present in the image. Thus, LUT will contain a *coded signal* for every star ID. The LUT can be constructed as an N_{LUT} -sample signal for every star ID, where N_{LUT} is given by (4). The entire signal is then broken into N_{LUT} samples (bins). At the same time, the amplitude will also fall in either of slabs from 0 to S (total number of slabs are $S+1$). The division of the amplitude into slabs corresponds to describing the radial distance of the data point when compared to the maximum distance (pattern radius). Thus, S is calculated by (5). Hence we divide the signal into N_{LUT} bins and each bin can take a value from 0 to S .

$$N_{LUT} = 2^k ; k = 1, 2, 3, \dots \quad (4)$$

$$S = \text{round}(\log_2 PR) \quad (5)$$

Depending upon An , every data point of the signal will fall into a certain bin number given by (6) and the amplitude of that data (rn) will fall into one of the $S+1$ slabs given by (7).

$$\text{bin}_{LUT} = \text{round}\left(\frac{\text{Angle}^0}{180^\circ/N_{LUT}}\right) \quad (6)$$

$$\text{slab}_{LUT} = \text{round}\left(\frac{\text{Radial distance}}{(PR/(S+1))}\right) \quad (7)$$

For visualization and better explanation, we will describe the construction of the LUT for the parameters listed in TABLE I. For the parameters listed in TABLE I, PR is calculated to have a value of 512 pixels and the corresponding S calculated is 9. Thus, the total number of slabs are 10 i.e. 0 to 9. We select the value of k as 5 in (4), which gives us the value of N_{LUT} as 32. The reason for this selection is explained later in part A of Section V in detail. For now, we will explain the construction of LUT for the reference star A (in Fig. 6) with N_{LUT} as 32.

If there is no data point or a star in a bin, it is allotted a slab_{LUT} value of 0, as shown in Fig. 7. If there is a single data point in a bin, it is assigned the slab_{LUT} calculated by (7). For example, in Fig. 7, only one point, S5 lies in a certain bin. Thus, the slab_{LUT} value of point S5 calculated by (7) is equal to 9. If there is more than one point in a bin, the mean of their amplitude values (rn) is considered for calculation. For example, in Fig. 7, the slab_{LUT} value calculated for two points S1 and S7 falling in the same bin is 6. Hence using (6) and (7), a 32-sample signal – {9, 6, 8, 7 ... 0} as shown in Fig. 7 is prepared for the pattern of reference star A. Similarly, the 32-sample signal for each star ID is constructed for preparing the LUT.

3) Star Pattern Database (SPD) construction

The SPD is also constructed as an N_{SPD} -sample signal for every star ID, where N_{SPD} is given by (8). The sample signal constructed for the SPD is expected to have a higher number of samples because this signal will be used for identifying the correct star from the shortlisted entries. Thus, the value of N_{SPD} must be high compared to that of N_{LUT} . For the parameters listed in TABLE I, we select the value of k as 9 in (8), and thus, the N_{SPD} calculated is 512. The reason behind this selection is explained later in part A of Section V. For now, we will explain the construction of the SPD for the reference star A (in Fig. 6) with N_{SPD} as 512.

$$N_{SPD} = 2^k ; k = 1, 2, 3, \dots \quad (8)$$

For the SPD, the signal is broken into 512 samples (bins) and the amplitude remains the same as the radial distance of the data point. For a bin, if there is no data point, it is assigned a value of zero; if there is a single data point, it is assigned the value of the radial distance of that data point; if there is more than one data point, it is assigned a value equal to the average of the data points, all the three cases been shown in Fig. 7. The bin number for a data point for the SPD is given by (9).

$$\text{bin}_{SPD} = \text{round}\left(\frac{\text{Angle}^0}{(180^\circ/N_{SPD})}\right) \quad (9)$$

As there are 512 bins, it is obvious that there will be many bins which will have zero value in between two data points. So, if there are more than three bins with zero value between two data points, they are represented by $z_number_of_zeroes$. Hence the 512-sample signal for the pattern of reference star A (in Fig. 6) is prepared as – {z26, r5, 0, 0, mean(r1,r7) ... z358} as shown in Fig. 7. Similarly, the 512-sample signal is constructed for every star ID and the SPD is prepared.

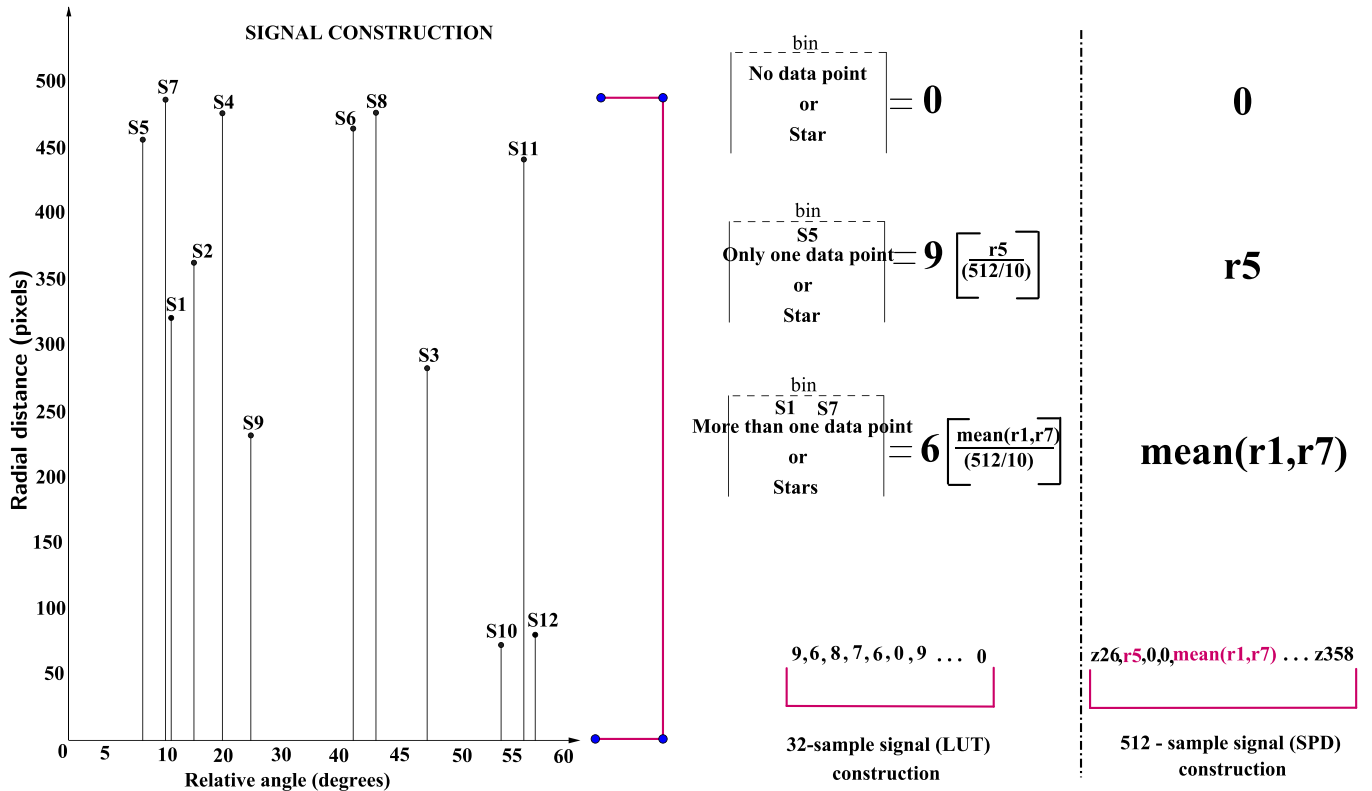


Fig. 7. Signal, LUT and SPD construction from the extracted features.

As explained earlier, the 32-sample signal and the 512-sample signal is constructed for every star ID. The LUT and SPD prepared for the proposed technique are shown in Fig. 8. The reason for selecting 32-sample for LUT and 512-sample for SPD is explained later in part A of Section V of this paper.

Look-Up-Table								Star Pattern Database							
Index	32 sample signal							512 sample signal							
0	6	6	4	9	...	0	0	0	333	z4	452	z8	...	z291	
1	7	8	2	5	...	0	0	z5	414	z7	348	z10	...	z301	
2	6	6	2	5	...	0	0	z5	108	338	0	403	...	z265	
.	
.	
.	
4955	4	6	9	7	...	0	0	0	138	282	0	0	...	z313	

Fig. 8. Look-Up-Table and Star Pattern Database for the proposed technique.

IV. SHORTLISTING AND STAR IDENTIFICATION

For explaining the process of shortlisting using hamming distance and star identification using spearman correlation, we will be using the example shown in Fig. 9. In Fig. 9, the image captured is represented by the dotted circle, with I_c being the center of the image. Let us say that star A is the closest to I_c and thus, it will be selected as a reference star. The SPD of star A is represented by the solid circle in Fig. 9, which is the same as was shown in Fig. 6.

It should also be noted from Fig. 9 that the captured image does not contain the stars 4, 7, 8, 10 and 12 which are a part of the SPD of star A (in Fig. 6). Also, the image contains a false star F which is not a part of the SPD of star A. Stars 4, 7 and 8 are missed due to patch mismatch and stars 10 and 12, which

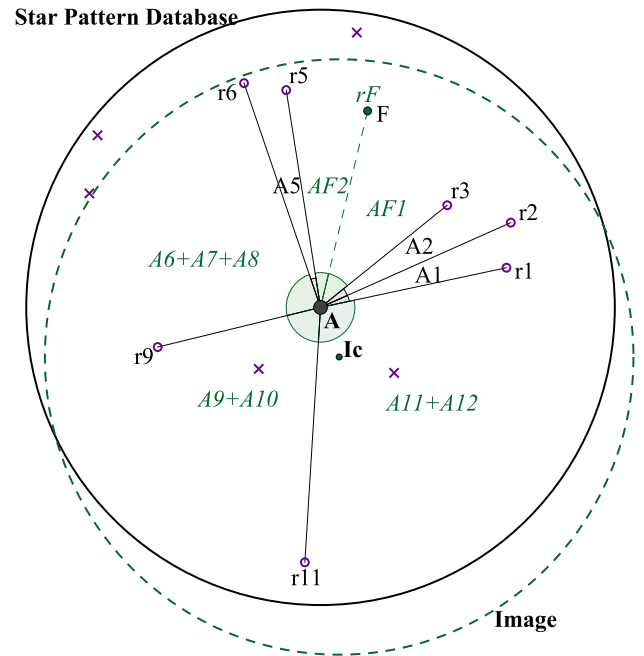


Fig. 9. An example of the image captured for explaining the proposed idea.

are also nearest to the reference star A, are missed due to magnitude uncertainty. Thus, the example shown in Fig. 9 has all the problems which a star pattern recognition technique will face (also identified in Section II of this paper). Hence the aim of the proposed technique is to identify the stars correctly in the image along with a low time complexity. We will explain the process of shortlisting using a 32-sample signal and star identification using a 512-sample signal, as both LUT and SPD are prepared using the same parameters.

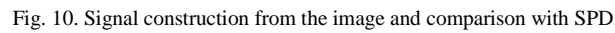


Fig. 11. Calculation of hamming distance between I_{32} and LUT_{A32}

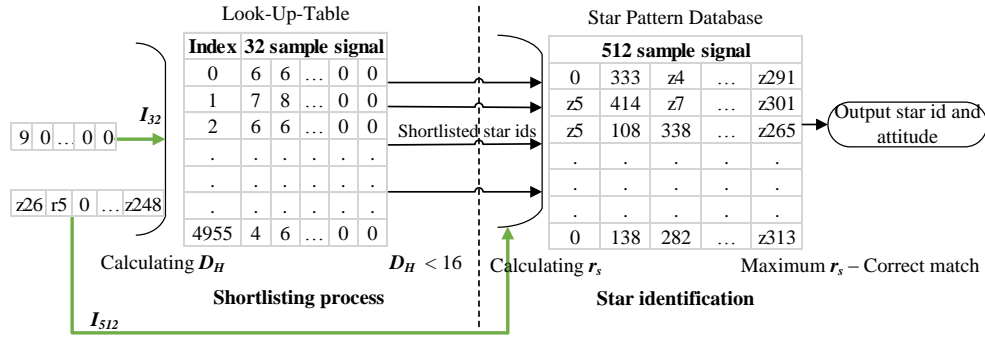


Fig. 12. Shortlisting and star identification process.

If there is a very high amount of patch mismatch and magnitude uncertainty, the number of data points can also be very less in the signal constructed from the image acquired. Owing to the above-mentioned scenario, the best way to measure the match of the image with the SPD is by calculating the *spearman correlation co-efficient* between the signal constructed from the image acquired and the signals of the shortlisted star IDs in the SPD. Normally, correlation refers to calculating the *pearson co-efficient*, however, the data variables must have constant variance and exhibit a linear relationship for the pearson correlation co-efficient. Moreover, the number of samples must be significant for calculating the pearson co-efficient [28]. As the signal constructed from the image acquired does not satisfy the above-mentioned conditions, we do not prefer pearson correlation for star identification. On the other hand, spearman correlation is a non-parametric measure of correlation and is appropriate for ordinal scale signals [29]. It does not assume a normal distribution of data or any definite relationship between the data points. Thus, in the above-mentioned conditions, spearman correlation co-efficient will be an appropriate measure for calculating the match between the image and the SPD. The spearman correlation coefficient (r_s) is calculated by (12) and (13).

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (12)$$

$$d_i = rg(X_i) - rg(Y_i) \quad (13)$$

where $rg(X_i)$ and $rg(Y_i)$ are the ranks of each observation in the samples X and Y , and n is the number of samples.

The maximum spearman correlation co-efficient calculated between the 512-sample signals of the shortlisted star IDs in the SPD and the I_{512} (signal constructed from the image acquired) is regarded as the correct match for the reference star in the image. The entire process of shortlisting and star identification is shown in Fig. 12. Once the star ID which gives the maximum spearman correlation co-efficient is identified as the reference star, the nearest one or two stars to the reference star in the image can be identified from the sequence of the SPD of the reference star ID. The existing QUEST or TRIAD method [30] requires two to three correctly identified stars in the image for determining the attitude.

V. RESULTS – PERFORMANCE AND BENCHMARKING

The performance of the star pattern recognition techniques was tested on 16200 simulated images generated by the MATLAB software. We have tried to make the simulated images like real images and thus, we select the image parameters listed in TABLE I (except for M_v being 6.0) for the generation of simulated images. The images were generated by varying the ascension angle (α) from 0° to 360° and the declination angle (δ) from -90° to $+90^\circ$ at an interval of every two degrees increment. This simulation covers the full sky scan of all the stars having a magnitude threshold M_v less than 6.0 from the star catalog SAO J2000 [23]. The platform PC specifications were - Intel core i7, 3.4 GHz processor, and 4GB RAM. In this section, firstly we provide the rationale for selecting 32 samples and 512 samples for LUT and SPD construction respectively for the image parameters listed in TABLE I. Later, benchmarking¹ of the proposed technique with state-of-the-art star pattern recognition techniques is shown with details of simulation image setup followed by discussion and analysis of the performance of the compared techniques and our proposed technique. Finally, the performance of the proposed idea on real images is also shown.

A. Selecting 32 samples for shortlisting and 512 samples for star identification

Firstly, we will show the reason behind the selection of 32 samples (value of k as 5 in (4) for shortlisting) and 512 samples (value of k as 9 in (8) for star identification) based on the test results for different sample sizes. Concerning with the shortlisting process, hamming distance performance for different sample sizes (different of values of k in (4)) is shown in Fig. 13. We vary the value of k from 3 to 7, as the number of samples of the discrete signal required for shortlisting is not that high. The performance is measured in terms of two parameters – the number of shortlisted IDs (N_{hd}) (out of total 4956 star IDs) and the probability that the shortlisted entries contain the correct star ID (p_{hd}). Lower the N_{hd} and higher the p_{hd} , better the performance. As can be seen in Fig. 13, the N_{hd} is least (1071) for 8-sample signal ($k = 3$), however, the p_{hd} is also least (0.537) for this case. On the contrary, the p_{hd} is maximum (1.0) for 128-sample signal ($k = 7$), however, the N_{hd} is also very high (3565).

¹ Codes for simulation and testing of the star pattern recognition techniques are available at https://github.com/DevD1092/spr_ham_spear.

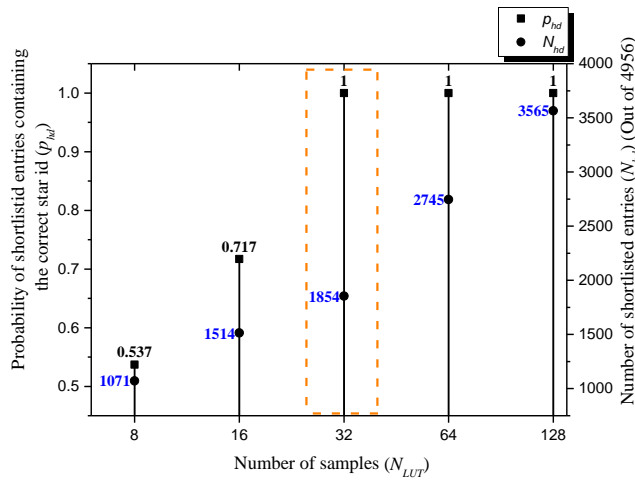


Fig. 13. Comparison of performance of the proposed shortlisting technique for a different number of samples (N_{LUT}).

When the number of samples is greater than or equal 32 ($k \geq 5$), the p_{hd} is 1. However, if the sample size still goes on increasing the N_{hd} will also go on increasing. Thus, an optimum sample size of 32 gives a perfect p_{hd} and a low N_{hd} (almost one-third of the total star IDs). Thus, it is evident from Fig. 13 that the 32-sample signal ($k = 5$ in (4)) is the perfect choice for the shortlisting technique to attain the best performance.

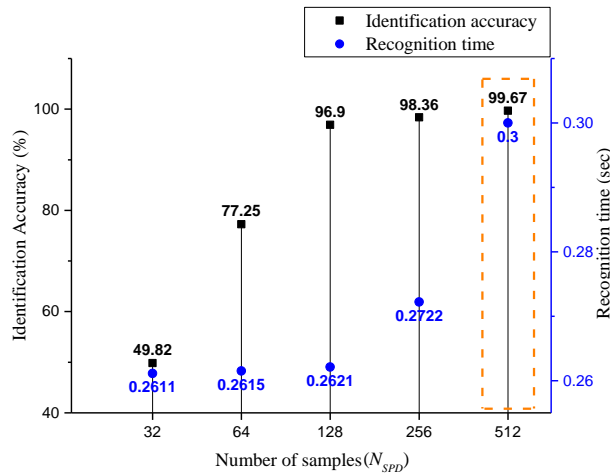


Fig. 14. Comparison of performance of the proposed star identification technique for a different number of samples (N_{SPD}).

Concerning with the star identification process, spearman correlation performance for different sample sizes (different values of k in (8)) is shown in Fig. 14. We vary the value of k from 5 to 9 as the star identification requires a high number of samples to achieve a reliable identification. The performance is measured in terms of identification accuracy and recognition time required for different sample sizes. As can be seen from Fig. 14, 32-sample signal ($k = 5$) gives a very low identification accuracy and the identification accuracy increases with the number of samples. The recognition time also increases with the number of samples, however, the increase in the recognition time is not drastic. We choose 512-sample size ($k = 9$ in (8)) because the identification accuracy is very high, and the recognition time is reasonable in this case. Because of the shortlisting due to hamming distance and the

performance achieved by it (as explained earlier), this recognition time for the 512-sample size will be reduced. The overall performance of the shortlisting and the star identification technique is depicted and benchmarked in TABLE II.

B. Benchmarking with state-of-the-art techniques

State-of-the-art star pattern recognition techniques which claim to provide high recognition reliability and high speed of recognition were implemented. We compare a high identification accuracy algorithm polestar [13], a high speed of recognition technique based on search tree with optimized database (STOD) [16], and a high recognition reliable technique geometric voting algorithm (GMV) [7] with our proposed technique. All the above three methods select the star nearest to the center of the image as the reference star to be identified and none of them use the ambiguous brightness of the star as a feature for comparison which is in compliance with our proposed technique. Hence the above-mentioned three methods are used for a justified benchmarking. To begin with the testing, an ideal image was considered. Later deviation in star positions, magnitude uncertainty, and false stars were added to the generated images to check the recognition reliability of the star pattern recognition techniques. The details of all the above stages of benchmarking are described in their corresponding part. The results and analysis of the performance of the star pattern recognition techniques are shown below.

1) Ideal case

In an ideal case scenario, there is no noise, positional deviation or magnitude uncertainty in the image simulated. The simulated image parameters are the same as in TABLE I. However, the problem of patch mismatch is by default present in this case. TABLE II summarizes the performance of the star pattern recognition techniques in an ideal case scenario. The identification accuracy corresponds to the number of correctly identified star images (out of the total of 16200). The run-time is measured after the process of star centroiding until the IDs of the stars in the image are returned. This time measurement method is in compliance with the one described in STOD [16].

TABLE II

Benchmarking of star pattern recognition methods – Ideal case

Technique	Identification accuracy (%)	Run-time (sec)	SPD size (MB)
GMV [7]	94.3	0.567	1.31
STOD [16]	90.3	0.21	1.4
Polestar [13]	97.86	0.386	0.9
Proposed technique	99.67	0.112	1.21

The results in TABLE II depict that the proposed technique outperforms every other star pattern recognition technique in the aspect of identification accuracy and run-time. The identification accuracy achieved by the proposed technique shows an increase of almost 2% when compared to the Polestar technique which claims to have a very high identification accuracy. Comparing the speed of recognition, our technique outperforms the fastest STOD technique by almost 0.1 seconds (almost twice as fast as STOD). STOD has

a very fast recognition because of the parallel search and optimized tree database adopted. Our technique outperforms STOD because of the fast shortlisting achieved by hamming distance comparison. GMV and the Polestar algorithm have a comparatively high run-time because of the voting process adopted by both the techniques. SPD size (including LUT) of our technique is 1.21 MB (only 0.3 MB more than that of the Polestar). However, current on-board computers can easily store up to 5MB in their on-chip memory. The primary focus is on attaining a high identification accuracy and less run-time. The test for robustness of the star pattern recognition techniques in the scenario of noise and magnitude uncertainty in the simulated images is shown later in this section.

2) Positional deviation of stars

The spacecraft usually undergoes unexpected velocity changes and acceleration. It also experiences vibrations and disturbances throughout the period in its orbit. Due to this, the image captured by the star tracker is distorted. This leads to deviation in position of stars in the image, which further leads to an error in features extracted from the image. To simulate this scenario, we introduce positional deviation (also known as positional noise) of stars in the simulated images. In our simulations, the positional deviation is introduced in the images from 0.1 pixels to 1.0 pixels in the steps of 0.1. This introduction of the pixel positional deviation is equivalent to the addition of random Gaussian noise to the projected locations of the stars in the image (as described in Polestar [13]). In [13], the authors added the positional deviation with a nominal standard deviation (σ) of 0.062 pixels and a maximum of 0.31 pixels. In our simulations, we assign a limit of 1.0 pixel as the maximum positional deviation for a star is usually around 1.0 pixel in the space environment according to a recent finding [31].

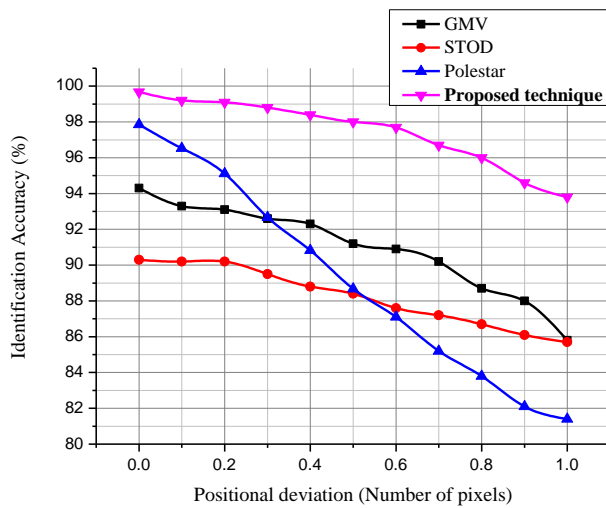


Fig. 15. Recognition reliability of star pattern recognition techniques with positional deviation.

Geometric distance between the stars is the most affected feature due to the positional deviation of stars. The recognition reliability of the star recognition techniques in the scenario of positional deviation is shown in Fig. 15. It can be seen from Fig. 15 that the proposed technique robustly maintains the identification accuracy to 94% whereas the identification

accuracy of the other techniques falls below 85% when the positional deviation tends to 1.0 pixel. Polestar technique fails drastically to recognize stars in this scenario because it utilizes a coded binary string constructed from the geometric distance feature. GMV utilizes voting of the distances, and STOD utilizes the distance as the only feature for comparison. Thus, both GMV and STOD also fail in this scenario. Our technique utilizes both the distance and relative angle between the stars as the features extracted. Data points in the signal constructed do not undergo a high shift in position because the relative angles are used as the time domain in our proposed technique. Thus, spearman correlation co-efficient calculated does not change considerably. Hence the proposed technique maintains the recognition reliability to 94%.

3) Magnitude uncertainty of stars

As was described in Section II of this paper, magnitude uncertainty is one of the major problems faced by star pattern recognition techniques. The magnitude uncertainty adds missing stars to the already existing problem of patch mismatch. To simulate such a scenario, we apply a relative magnitude (M_v) uncertainty to the simulated images. The magnitude uncertainty is varied from 0.0 to 0.4 in steps of 0.05. For example, an M_v uncertainty of 0.1 would suggest stars having M_v greater than 5.9 will be missed (normal magnitude threshold being 6.0 M_v). This simulation will have additional missing stars to the ones by normal patch mismatch (ideal case scenario). This variation of magnitude uncertainty is equivalent to the addition of random Gaussian noise to the magnitude of the stars which is adopted in Polestar [13]. In [13], the nominal standard deviation (σ) for the added Gaussian noise is 0.323 M_v , and the maximum value is 0.9 M_v with the magnitude threshold of 7.0 M_v . In our simulations, as our magnitude threshold is 6.0 M_v , we vary the magnitude uncertainty till 0.4 M_v , which is enough for performance evaluation. If the variation is further increased, the number of stars in the image will be less than 5, which would not correspond to a real image. The recognition reliability of the star pattern recognition techniques for the case of magnitude uncertainty is shown in Fig. 16.

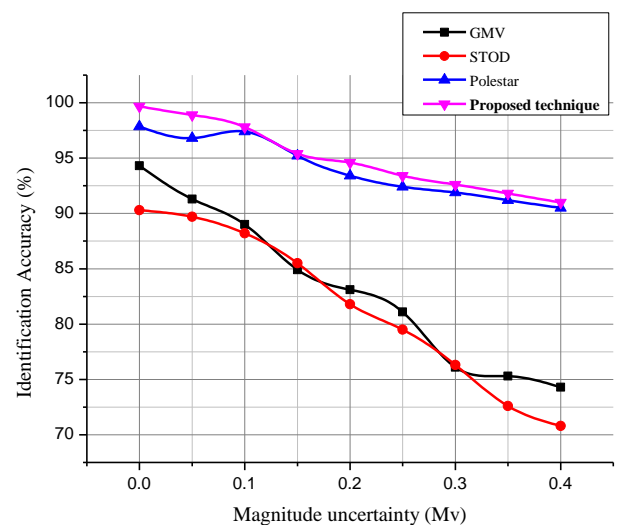


Fig. 16. Recognition reliability of star pattern recognition techniques for the scenario of magnitude uncertainty.

In the scenario of magnitude uncertainty, stars which are low in brightness are failed to be captured by the star tracker. As can be seen in Fig. 16, GMV and STOD fail drastically to identify stars correctly in this scenario. STOD compares the number of neighboring stars (in the FOV of the reference star) as the parameter for recognition and GMV counts the votes of the matched radial distances of the neighboring stars (in the FOV of the reference star). This approach leads to failure of both the techniques, as the low brightness neighboring stars are missed. Polestar maintains a decent recognition reliability in this case because it does not utilize the brightness of the stars as a feature for comparison and it also utilizes the various combination of neighboring stars in the FOV of the reference star for feature construction. Our technique maintains a high recognition reliability because of the very precise 512-sample signal construction and spearman correlation. The functioning of our proposed technique in this scenario was also explained in part B of section IV of this paper.

4) False stars

As was explained earlier in Section II of this paper, false stars in the image are also one of the problems for the star pattern recognition techniques. There can be as many as 1 to 4 false stars present in the image depending on the intensity threshold applied. Thus, 1 to 5 false stars with random brightness and position were added to the image simulated. This approach of adding false stars is adopted from the one described in STOD [16]. The recognition reliability of the star pattern recognition techniques for the case of false stars per image is shown in Fig. 17.

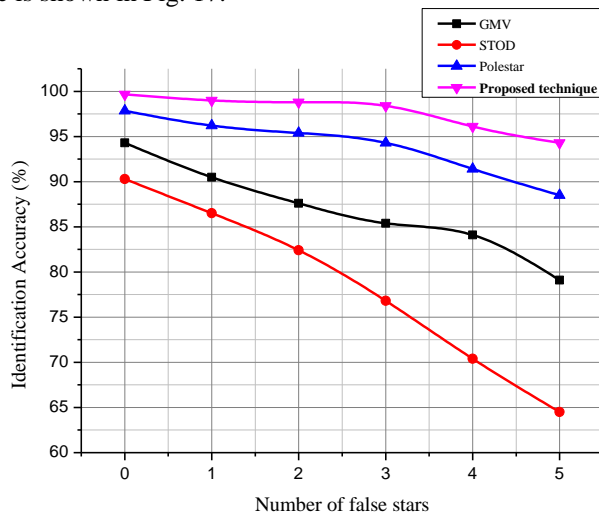


Fig. 17. Recognition reliability of the star pattern recognition techniques for the case of false stars.

As can be seen from Fig. 17, the GMV and STOD fails drastically for this case. The reason for the failure of GMV and STOD is the same for which they failed during the magnitude uncertainty of stars. Polestar's recognition reliability also drops, but not as drastic as that of the STOD and GMV. This is because it takes into consideration various combinations of the neighboring stars in the FOV for feature extraction. The reason for the high recognition reliability of the proposed technique for this case is the same as that for the magnitude uncertainty of the stars.

C. Performance on real images captured by star tracker

The proposed technique was implemented on Linux-based C platform (PC specifications: i7-3.4GHz and 4GB RAM) as well as on a Raspberry Pi 3 (RPi) (specs: 1.2 GHz quad-core ARM Cortex A53; 1GB RAM) and tested on 2384 real images captured by the star tracker SST-20S. The specifications of the star tracker were listed in TABLE I. The real images captured by the star tracker are continuous and thus, there are approximately 24 images per set (total of 100 sets of images).

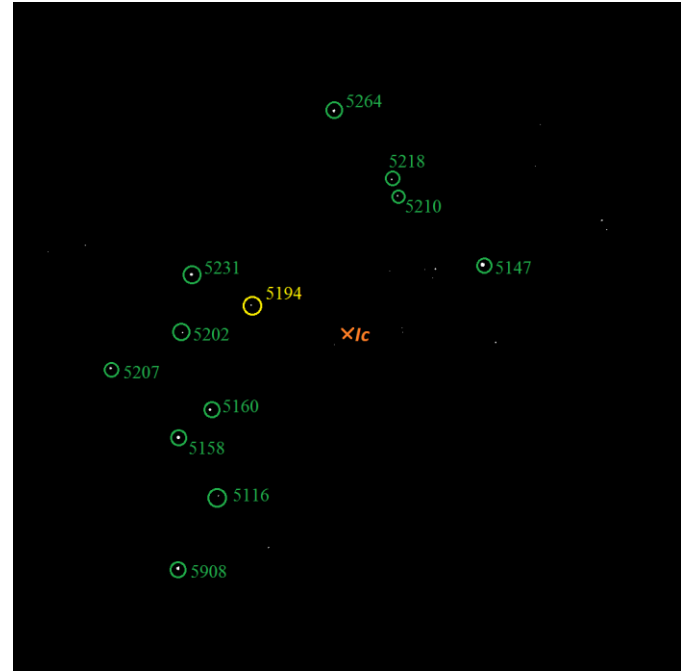


Fig. 18. Identified real image by the proposed technique.

The identification accuracy of the proposed technique was calculated by comparing the attitude obtained by applying QUEST [30] on the star IDs identified and the attitude obtained by the star tracker SST-20S. A correct identification was counted if the difference between the calculated and real attitude was less than 40 arc seconds (boresight accuracy of SST-20S). The proposed technique correctly identified 2310 out of 2384 real images, thus, achieving an identification accuracy of 96.9%. The technique failed to identify 3 out of 100 sets of images. The reason behind the failure of the technique was the maximum spearman-correlation co-efficient obtained for all these three sets was less than 0.5, thus, suggesting that the match between the signal acquired from the image and the signal stored in the SPD of the shortlisted entries is very low and cannot be trusted.

The proposed technique took only 79 ms (PC) and 394 ms (RPi) on an average to identify the stars of which reading the image data acquired, pre-processing and centroiding operation occupied 68 ms on PC and 342 ms on RPi. Implementation of the proposed idea and eventually to output the identified stars took only 11 ms on PC and 52 ms on RPi.

In Fig. 18, we show one of the identified real images by the proposed technique. The center of the image is marked with a cross (I_c). The star nearest to the center of the image is selected as the reference star to be identified and is marked with yellow. This reference star is identified as 5194 from the

SPD and the spearman correlation achieved for this image is 0.79. From the SPD sequence of the reference star ID 5194, rest of the stars in the image identified are marked with green.

VI. CONCLUSIONS

A novel approach to achieve both highly reliable and fast star identification for a Lost-In-Space mode star tracker is proposed in this paper. Firstly, a dedicated N_{LUT} (in our case 32) sample and N_{SPD} (in our case 512) sample signal is constructed from the features extracted in the acquired image. Later, the N_{LUT} sample signal is compared with the LUT using hamming distance and only a selected number of star IDs are shortlisted and passed for the star identification process. The very precise N_{SPD} sample signal is then compared with the ones stored in the SPD of the shortlisted entries, and the star ID signal which achieves the maximum spearman correlation is identified as the correct reference star. The simulation results depict that the proposed technique is highly reliable in the case of positional deviation, magnitude uncertainty and even false stars (maintaining the recognition reliability above 96%). This achievable performance in the simulations is also justified by the real-time testing, where the identification accuracy of 97% is achieved with the total run-time of only 79 ms on PC and 394 ms on RPi. The proposed technique achieves a significant improvement in identification accuracy and robustness along with a reduced run-time when compared to the existing approaches. Hence a highly robust and a fast star identification algorithm based on hamming distance and spearman correlation is developed, which is suitable for actual space applications.

REFERENCES

- [1] Griffin, M.D., & French, J.R., "Attitude Determination and Control," in *Space Vehicle Design* (pp. 325 - 350): 2nd ed., Reston, VA: American Institute of Aeronautics and Astronautics, 2004.
- [2] Davies, A., & Holt, A., "Use of autonomous star trackers in modern attitude and orbit control systems," *Spacecraft Guidance, Navigation and Control Systems*, p. 87, 2003.
- [3] Birnbaum, Morris M., "Spacecraft attitude control using star field trackers," *Acta Astronautica*, vol. 39(9), pp. 763-773, 1996.
- [4] Liebe, C. C., "Pattern recognition of star constellations for spacecraft applications," *IEEE Aerospace & Electronic Systems Magazine*, vol. 8, (1), pp. 31 - 39, 1993.
- [5] Mortari, D., Samaan, M. A., Bruccoleri, C., & Junkins, J. L., "The pyramid star identification technique," *Journal of the Institute of Navigation*, vol. 51(3), pp. 171-183, 2004.
- [6] Cole, C. L., & Crassidis, J. L., "Fast star-pattern recognition using planar triangles," *Journal of Guidance Control and Dynamics*, vol. 29(1), pp. 64-71, 2006.
- [7] Kolomenkin, M., Pollak, S., Shimshoni, I., & Lindenbaum, M., "Geometric voting algorithm for star trackers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44(2), pp. 441-456, 2008.
- [8] Li, J., Wei, X., & Zhang, G., "Iterative algorithm for autonomous star identification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51(1), pp. 536-547, 2015.
- [9] Padgett, C., & Kreutz-Delgado, K., "A grid algorithm for autonomous star identification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33(1), pp. 202-213, 1997.
- [10] Yoon, Youngwoo, "Autonomous star identification using pattern code," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49(3), pp. 2065-2072, 2013.
- [11] Na, M., Zheng, D. N., & Jia, P. F., "Modified Grid Algorithm for Noisy All-Sky Autonomous Star Identification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45(2), pp. 516-522, 2009.
- [12] Servidia, P. A., Pena, R.S., "New robust star identification algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42(3), pp. 1126, 6p, 2006.
- [13] Silani, E., & M. Lovera., "Star identification algorithms: Novel approach & comparison study," *IEEE Transactions on Aerospace and Electronic systems*, vol. 42(4), pp. 1275-1288, 2006.
- [14] Arani, M. Shayan, A. Toloei, & Z. Eghbaleh, "A geometric star identification algorithm based on triple triangle pattern," 7th IEEE International Conference on Recent Advances in Space Technologies (RAST), pp. 81-85, 2015.
- [15] Mortari, Daniele, & Beny Neta., "K-vector range searching techniques," *Adv. Astronaut. Sci.*, vol. 105, pp. 449-464, 2000.
- [16] Pham, M. D., Low, K. S., & Chen, S. S., "An Autonomous Star Recognition Algorithm with Optimized Database," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49(3), pp. 1467-1475, 2013.
- [17] Pham, M., Low, K. S., Shoushun, C., & Xing, Y., "A star pattern recognition algorithm for satellite attitude determination," IEEE Symposium on Industrial Electronics and Applications (ISIEA), 2012.
- [18] Li, Qi-Shen, Chang-Ming Zhu, & Jun Guan, "A fast star pattern recognition algorithm based on feature vector," IEEE International Conference on Computer Design and Applications (ICCD), vol. 1, pp. V1-256, 2010.
- [19] Xie, J., & Wang, X., "A robust autonomous star identification algorithm for ZY3 satellite," 1st IEEE International Conference on Agro-Geoinformatics (Agro-Geoinformatics), pp. 1-4, 2012.
- [20] Delabie, T., Durt, T., & Vandersteen, J., "Highly Robust Lost-in-Space Algorithm Based on the Shortest Distance Transform" *Journal of Guidance, Control, and Dynamics*, vol. 36(2), pp. 476-484, 2013.
- [21] Sohrabi, S. H., & Shirazi, A. B., "A novel, smart and fast searching method for star pattern recognition using star magnitudes," 25th IEEE International Conference Image and Vision Computing New Zealand (IVCNZ), pp. 1-6, 2010.
- [22] Spaceflight 101, "VELOX-CI Satellite", 2015. [Online]. Available: <http://spaceflight101.com/pslv-c29/velox-ci/>.
- [23] Myers, J. R., et al, "SKY2000-master star catalog-star catalog database," *Bulletin of the American Astronomical Society*, 191(128.12), 1997.
- [24] Padgett Curtis, Kenneth Kreutz-Delgado, & Suraphol Udomkesmalee, "Evaluation of star identification techniques," *Journal of Guidance, Control, and Dynamics*, vol. 20(2), pp. 259-267, 1997.
- [25] Accardo Domenico, & Giancarlo Rufino, "Brightness-independent start-up routine for star trackers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38(3), pp. 813-823, 2002.
- [26] Hamming, Richard W., "Error detecting and error correcting codes," *Bell System technical journal*, vol. 29(2), pp. 147-160, 1950.
- [27] Norouzi, Mohammad, David J. Fleet, & Ruslan R. Salakhutdinov., "Hamming distance metric learning," *Advances in neural information processing systems*, pp. 1061-1069, 2012.
- [28] Carroll, John B., "The nature of the data, or how to choose a correlation coefficient," *Psychometrika*, vol. 26(4), pp. 347-372, 1961.
- [29] Hauke, Jan, & Kossowski, T., "Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data," *Quaestiones geographicae*, vol. 30(2), pp. 87-93, 2011.
- [30] Shuster, Malcolm David, & S. D. Oh, "Three-axis attitude determination from vector observations," *Journal of Guidance, Control, and Dynamics*, vol. 4(1), pp. 70-77, 1981.
- [31] Jun, Z., Yuncui, H., Li, W., & Da, L., "Studies on dynamic motion compensation and positioning accuracy on star tracker," *Applied optics*, vol. 54(28), pp. 8417-8424, 2015.



Mehta Deval Samirbhai received his B.Tech. degree in electrical and electronics engineering from National Institute of Technology, Trichy, India in 2014. He started his Ph.D. programme in School of Electrical and Electronic Engineering at Nanyang Technological University, Singapore in 2014.

His Ph.D. thesis is on developing star trackers for nano-satellites and his main research interests include image and signal processing, and attitude determination for satellites.



Shoushun Chen (M'05) received his B.S. degree from Peking University, M.E. degree from Chinese Academy of Sciences and Ph.D. degree from Hong Kong University of Science and Technology in 2000, 2003, and 2007.

He held a post-doctoral research fellowship in the Department of Electronic & Computer Engineering, Hong Kong University of Science and Technology for one year after graduation. From February 2008 to May 2009 he was a post-doctoral research associate within the Department of Electrical Engineering, Yale University. In July 2009, he joined Nanyang Technological University as an assistant professor.

Dr. Chen serves as a technical committee member of Sensory Systems, IEEE Circuits and Systems Society (CASS); Associate Editor of IEEE Sensors Journal; Associate Editor of Journal of Low Power Electronics and Applications; Program Director (Smart Sensors) of VIRTUS, IC Design Centre of Excellence; regular reviewer for a number of international conferences and journals such as TVLSI, TCAS-I/II, TBioCAS, TPAMI, Sensors, TCSVT, etc.



Kay-Soon Low (M'88–SM'00) received his B.Eng. degree in electrical engineering from the National University of Singapore and his Ph.D. degree in electrical engineering from the University of New South Wales, Sydney, Australia.

He joined the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 1994, as a lecturer and was subsequently promoted to an Associate Professor in 1999. He also served as the Centre director for the Satellite Research Centre (SaRC) from April 2009 to October 2016. Presently he is a Professor at Department of Electrical and Computer Engineering, National University of Singapore, Singapore and also the Director of Satellite Technology and Research Centre (STAR). His funded projects are in the area of satellite system, wireless sensor network and clean energy system. Dr. Low has published more than 200 research papers and has 16 granted patents as the first inventor.