Proceedings of the 2021 IEEE International
Conference on Intelligence and Safety for Robotics
Nagoya, Japan, March 4-6, 2021

Cumulative Clustering Filter for MIMO Radar Detecting Human Hand Intrusion

Huazhe Wu[1], Eugene Kim[1], Yoji Yamada[1], Shogo Okamoto[1]

***Abstract*— MIMO radar sensors with a 3-dimension detection zone are recently used for safety-related sensors. They may possess the capability of detecting human body parts such as hands. When the reflectivity of a target is low and the sensitivity of the radar is therefore relatively high, it is necessary to eliminate cluttered noise out of detected points and increase the possibility of detecting the target at the same time. This paper presents a data processing method including a cumulative filter and clustering, for achieving the goals above. Taking the human hand intrusion scenario as an example, an experiment was conducted to investigate the improvement of the detection performance by implementing this cumulative clustering filter.**

I. INTRODUCTION

The international safety standard for industrial robots, ISO 10218-1, states safety requirements for implementing at least one of four suggested collaborative modes for a human operator to coexist with an industrial manipulator in the collaborative workspace, while two of them require fulfilling the safety integrity level for detecting the operator [1]. And Speed and Separation Monitoring (SSM) method stated in [2] requires the position or range information of the operator obtained by safety-related sensors (SRS).

In this respect, short-range radar sensors have gained a growing interest in the field of SRS, for their ability of resisting environmental conditions such as dust, sand, or moisture. By far, diverse types of protective equipment are used as SRS with corresponding standards (IEC 61496 series: general requirements are found in [3]). Among the radar applications, multiple-input and multiple-output (MIMO) radars can detect the 3-dimensional positions of targets. In addition, radars are promising enough to be easily attached even to moving robots due to the compactness and more stable performance in manufacturing environments, since their detection principle do not rely on movable mechanical structures [4]. MIMO radars have been proposed to detect humans existence [5], and their use as protective equipment is now more and more studied. Recent research trends of human detection by radars are the recognition and classification not only for body movements [6]–[8], but also for hand gestures [9].

However, there has been few study investigating the integrity of radars detecting human body parts [10]. Theoretically, a radar with an ultra-wide band has sufficient range resolution to distinguish between a torso and an outstretched hand. If the radar can detect the nearest body part such as a hand instead of the torso, with acceptable amount of errors, it will greatly reduce the intrusion distance, which is defined as the distance that a part of the body can intrude into the sensing field before it is detected [2]. A shorter intrusion distance leads to a shorter minimum distance (ISO 13855) [11] or a shorter protective separation distance (ISO/TS 15066) [2]. This can increase productivity due to closer collaboration [12]. Thus, this study focuses on detecting hand intrusion.

The capability of radars detecting a human hand has a limitation when the power of the reflected wave from the target is not sufficient to recognize it as it is. In general, the reflectivity of a human hand is a thousand times lower than that of a torso [13]. Therefore, such a low-reflectivity target presents too few detected points to be identified as a target of a human body, which can lead to miss-detection. On the contrary, when the sensitivity of radar is adjusted higher in order to detect low-reflectivity objects, clutters which originate in multi-path or other objects such as walls or floors also increase, resulting in false detection [14].

To deal with these problems, this study presents a data processing method at the stage of the target detection, which aims to exclude the cluttered noise by clustering the points, and increase the probability of a human hand getting detected by cumulating detection points. There are few studies on such a cumulative filter but for different usage [15]. An experiment was conducted to demonstrate the effect of the method on the sensing performance on the basis of requirements for the safety integrity level of SRS and safety of machinery.

II. METHOD

The raw data collected from the radar sensor includes time and position coordinates. The point cloud in the current frame is defined as a set

$$\boldsymbol{P}_k = \{\boldsymbol{p}_{i_k} | \boldsymbol{p}_{i_k} = (x_{i_k}, y_{i_k}, z_{i_k}) \in \boldsymbol{R}^3, i = 1, 2, \ldots, I_k\}, \quad (1)$$

where k is the current frame and I_k is the number of points at frame k. x, y and z are the position information of the point $\boldsymbol{p}$. In this section, such raw data is processed to obtain the final result R_k, representing the distance of a hand from the radar origin in frame k.

It should be noted that if the detected distance is longer than the actual distance between the hand and the radar, it is an error on the hazardous side. On the contrary, the error is considered to be on the safe side when the detected distance is shorter than the actual value. Therefore, the primary mission is to reduce hazardous-side error.

[1]Huazhe Wu is with Department of Mechanical Systems Engineering, , Nagoya University, Furo-Cho, Chikusa-ku, Nagoya, 464-8603, Japan wu.huazhe@b.mbox.nagoya-u.ac.jp

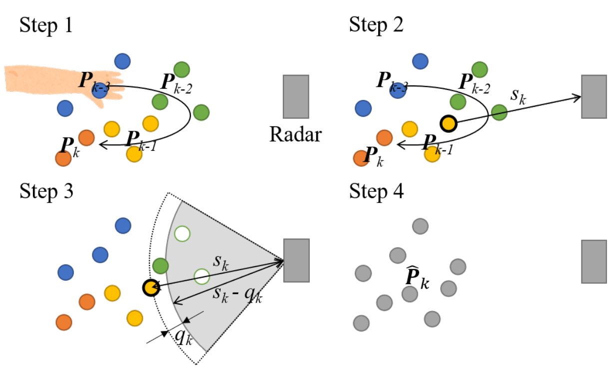

Fig. 1. Cumulative filter with selecting, when $n = 4$: step 1. collect all the points in recent 4 frames, points from each frames are in different colors and the curved arrow shows the flow of time; step 2. determine the the comparative distance from frame k and $k-1$; step 3. cumulate points with distance longer than $s_k - q_k$, points in the gray shaded area are excluded; step 4. the selectively cumulated point cloud remained.

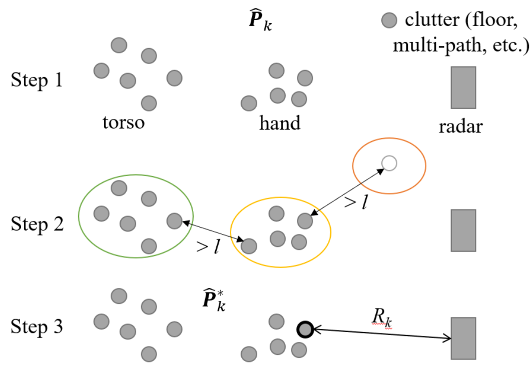

Fig. 2. Clustering and exclusion: step 1. the point cloud succeeded from cumulating phase; step 2. clustering stops when distance among clusters is over the criterion; step 3. the cluster contains only one point get excluded and the nearest of the remaining points leads to the result.

A. Cumulative Filter

Even though the clustering algorithm excludes the clutters which usually appear randomly in an isolated manner for only one frame (explained in II-B), the processing shall not accidentally exclude the point generated by a hand. The real target should remain as it is by cumulating points over a period of time, even the target point may appear alone only at some frames.

n frames of points, including those in the current frame, are cumulated first in the cumulative filter phase, as shown in Fig. 1.step 1. In order to decrease the time delay caused by the cumulating process when the target is moving away from the radar, the points in the past frames are selectively cumulated. Instead of collecting all the points for n frames in the past, the points which have shorter distance than $s_k - q_k$ are not be cumulated, such as the point in the gray shaded area in Fig. 1.step 3, namely the selecting process. q_k is a buffer distance in the selecting, which is for preserving valuable points so the point cloud is more easily clustered. As shown in Fig. 1.step 2, the comparative distance s_k in selecting is determined by the minimum distance in frame k and frame $k-1$,

$$s_k = min(dis(\boldsymbol{p}_{i_k}), dis(\boldsymbol{p}_{j_{k-1}})), \tag{2}$$

where $i = 1, 2, \ldots, I_k$ and $j = 1, 2, \ldots, I_{k-1}$, and dis is a function to calculate the euclidean distance of a certain point from the radar origin,

$$dis(\boldsymbol{p}_{i_k}) = \sqrt{x_{i_k}^2 + y_{i_k}^2 + z_{i_k}^2}. \tag{3}$$

Note that the purpose of keeping points in at least two adjacent frames k and $k-1$ is to avoid the target point becoming a single cluster. In conclusion, points cumulated selectively in frame k can be described as

$$\hat{\boldsymbol{P}}_k = \{\boldsymbol{p}_j, |dis(\boldsymbol{p}_j) \geq s_k - q_k, j = 1_t, 2_t, \ldots, I_t\}, \tag{4}$$

where $t = 1, 2, \ldots, n$.

The number n of cumulated frames is determined by the velocity v_k ,which is estimated from 6 past frames as

$$v_k = \frac{\sum_{i=k-3}^{3} R_i - \sum_{i=k-6}^{3} R_i}{9 \times T_s} \tag{5}$$

where R_i is the final distance result in frame i and T_s is the sampling time of the radar. In our preliminary experiments, a constant n was used and a disadvantage was observed. A long delay can be caused when n is large , but n should be sufficiently large for the case when targets move slowly. One of the characteristics of radars is they tend to detect objects with high velocity. The slower the target moves, the less detectable. Therefore, the number n is set to be adapted to the velocity of the target. In this study, $n = 8$ when $|v_k| \geq 1$ m/s and $n = 15$ when $|v_k| < 1$ m/s. This can be determined by experience and practical use.

The buffer distance q_k in the selecting process is also adapted to v_k. If the target is moving away from the radar ($v_k \geq 0$), $q_k = v_k T_s + CI$, where CI is the coverage interval of the radar, and $v_k T_s$ retraces the movement of the target in one frame before. If the target is approaching the radar ($v_k < 0$), $q_k = CI$. With the adaptive q_k, the selection functions more efficiently comparing to a constant $q = v_{\max} T_s + CI$, where v_{max} is the maximum speed of the target, which is 2 m/s in the case of human hands [11].

B. Clustering and Exclusion

After $\hat{\boldsymbol{P}}_k$ is gained in the cumulative filtering phase, these points need to be clustered to identify the clutter. Hierarchical Clustering is used in this method for following merits: Hierarchical Clustering does not need to specify the number of clusters, which is suitable for not knowing the number of targets; low-reflectivity targets with less points are hardly absorbed by big clusters.

To cluster all points representing the same target tempo-rally and spatially, the linkage criterion l should be determined as $l = v_{max} T_s + CI$, which is the shortest distance among individual clusters, as shown in Fig. 2.step 2. Such clustering procedure is achieved by 3 main steps: 1. compute the distances between all the points or clusters ; 2. link the

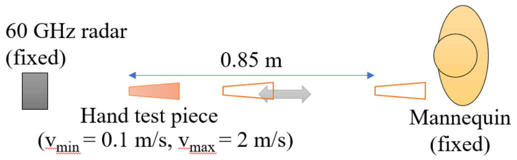

Fig. 3. Experimental setup

pair of points or clusters with the minimum distance; 3. repeat the two steps above until the distances among clusters are all longer than l. Usually, the point of cluttered noise appears individually. So after clustering, clusters including only one point, namely single cluster, are excluded, as shown in Fig. 2.step 2 to 3. The remaining points are described as a set $\hat{\boldsymbol{P}}_k^*$. Lastly, the final result is calculated from the point with minimum distance in $\hat{\boldsymbol{P}}_k^*$ as

$$R_k = min(dis(\boldsymbol{p}_j)), j = 1, 2, \dots J_k, \tag{6}$$

where $\boldsymbol{p}_j \in \hat{\boldsymbol{P}}_k^*$ and J_k is the number of points in $\hat{\boldsymbol{P}}_k^*$.

III. Experimental Design and Result

The experimental objective is to collect the data when a target moving in a mixture speed in order to investigate the effect of the cumulative filter and clustering, and the adaptability to changing speeds. The performance of detection should be improved especially when the target is moving slowly, due to an adequate number of n of cumulated frames. The efficiency was expected to be improved since the cluttered noises are excluded. Higher efficiency was also expected due to the selective cumulating in the case of target moving fast away from the radar.

A. Experimental Design

A 60GHz mmWave radar (IWR6843ISK, Texas instruments, America) was used and the parameter of the radar was adjusted for detecting low-reflectively targets with high range resolution. The target object was a test piece of ABS material with the shape of the human hand (IEC 61496-3). The test piece moved back and forth in a range of 0.85 m, which is the assumed length of the human arm, by using a linear actuator (LEFB25S2S-1000-S2A1, SMC, Japan). The approaching speed and the leaving speed of the test piece was set as 100 mm/s and 2000 mm/s, for the purpose of investigating the performance in both the miss-detection situation and the high-delay situation. Next, the human mannequin is placed behind the rail of the test piece to simulate the influence of a torso. The overall concept of the experimental setup is shown in Fig. 3. For comparison, a motion capture system (Motion Analysis Co., Santa Rosa, CA, USA) was used to monitor the position of the target object and the radar during the experiment, from which the data would be considered as true values. The detection period lasted for 15 minutes.

B. Experimental Result

As shown by (a) in Fig. 5, when the raw data was clustered without being cumulated, errors on the hazardous side increased because of accidentally excluding the target

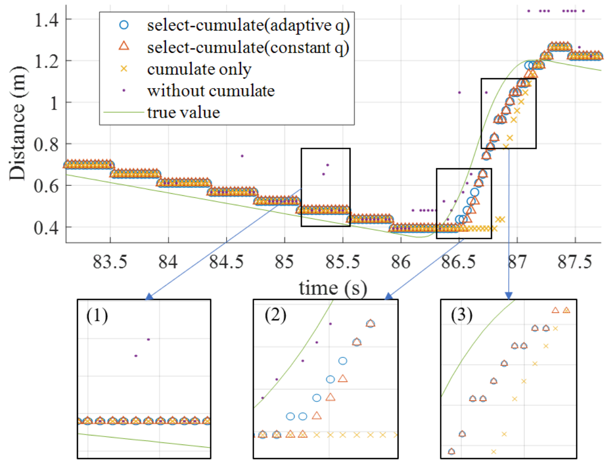

Fig. 4. Typical examples of distance detected by the radar, comparing between different data processing conditions together with data from motion capture. Each highlighted epoch shows : (1) long cumulating compensates the miss-detection when the target moved slowly; (2) applying adaptive q slightly improved the efficiency; (3) when the target moved away fast, the selecting strategy reduced the delay caused by cumulating

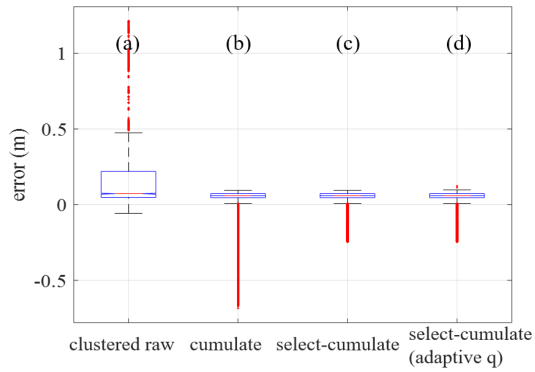

Fig. 5. Errors of the data of 4 conditions of processing. Hazardous side errors (> 0 m) are reduced by cumulating process (b), and safe side errors decreased from non-selective (b) to selective cumulating (c).

as a cluttered noise. As shown in (1) of Fig. 4, the cumulating strategy complemented errors, especially when the target moved slowly. However, (b) in Fig. 5 shows that the simple cumulating strategy generated the time delay, which is presented as errors on the safe side. On the contrary, the selective accumulating method reduced the delay by using adequate frame n as shown in (3) of Fig. 4. Comparison of selective and non-selective cumulating is shown by (b) and (c) in Fig. 5. In data processing, the effect of the adaptive buffer distance q_k is also observed by comparing to using a constant q, and the result showed that the efficiency improved slightly, typical samples of which can be observed in (2) of Fig. 4.

The errors larger than $\pm 5.16\ \sigma$ are considered as outliers in this study because the coverage interval which represents an acceptable error range of $PL_{\mathrm{r}} = d$ requires 5.16 σ [10], [16], which is set as 0.14 m for this radar detection. In addition, an outlier on the hazardous side means the detected distance is unacceptably longer than the true value, which

TABLE I

OUTLIER PROPORTION: THE OUTLIER PROPORTION ON HAZARDOUS SIDE REDUCED TO 0 USING CUMULATIVE CLUSTERING FILTER

	Hazardous side (%)	Safe side (%)
Raw	0.06	0
Clustered raw	49.73	0
Cumulate	0	5.95
Select-cumulate (constant q)	0	4.27
Select-cumulate (adaptive q)	0	3.61

may cause danger. As for the comparison between the experimental groups, the proportions of outliers were calculated for evaluating the effect of the proposed method. As shown in Table I, all outliers on the hazardous side in this experiment were removed after applying the method, as the proportion 0.06 % decreasing to 0 %. However, the effect of clustering to remove clutters near the radar could not be investigated, because the data did not contain such clutter.

IV. DISCUSSION

As mentioned, the data did not contain cluttered points near the radar. One of the causes of it is the test piece actuated by a linear slider moved much more smoothly than real humans. Also, the position of the radar during the experiment was static and it is observed in previous tests that such cluttered noises usually appear more in a dynamic condition, namely the radar be moving while detecting. However, the effect should still be investigated by experiments in dynamic conditions or other scenarios in which cluttered noise is prone to be generated. It should also be noted that, although the safe side outliers increased after the method was applied, in the situations of raw data containing more safe side outliers, it can be expected that the proposed method reduces them and thereby improves the efficiency.

Although all the outliers on the hazardous side were favorably removed in this experiment, which implies that there is no unacceptable error, a limitation of this method still exists. If the target is missed for a period due to an innate problem of radar performance, this method can not help predicting the current position. Therefore, the combination with a prediction process is expected in such situations. Another thing to note is that the clustering of detection points as well as exclusion of cluttered noise would exclude the isolated point of a real target if without accumulating. As shown in I, the hazardous side outliers increased severely from 0.06 % to 49.73 % by clustering without cumulating. So the clustering-exclusion algorithm must be combined with the cumulative filter.

V. CONCLUSION

This study was undertaken to complement the miss-detection issue when the reflectivity of a target is relatively low and the sensitivity of the radar threshold is relatively high. The proposed method of cumulating filter enables the radar to function in a manner that complementing the missed measurement which belongs to the target of human body parts. Further, it is evident from the experiment that the selective cumulating strategy which is the extension of naive cumulative filter, combining clustering-exclusion, has improved the integrity of the radar detection adaptively and efficiently. As a future study, it should be taken into account that more information in point clouds can be used to improve the performance of tracking a human hand, such as density, intensity, range-Doppler map, size, and shape of the point cloud.

ACKNOWLEDGMENT

The authors wish to thank Masatoshi Fujita, Takeshi Nomura, and Atsunori Hirano of FUJI corporation for their support of this project and technical assistance.

REFERENCES

[1] Robots and robotic devices–Safety requirements for industrial robots–Part 1: Robots, ISO 10218-1: 2011.

[2] Robots and robotic devices–Collaborative robots, ISO/TS 15066: 2015.

[3] Safety of machinery - Electro-sensitive protective equipment - Part 1: General requirements and tests, IEC 61469-1:2012.

[4] M. Zlatanski, P. Sommer, F. Zurfluh, and G. L. Madonna, "Radar sensor for fenceless machine guarding and collaborative robotics," *IEEE International Conference on Intelligence and Safety for Robotics*, pp. 19–25, 2018.

[5] G. Shingu, K. Takizawa, and T. Ikegami, "Human body detection using mimo-uwb radar sensor network in an indoor environment," *Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies*, pp. 437–442, 2008.

[6] D. Sasakawa, N. Honma, T. Nakayama, and S. Iizuka, "Evaluation of filtering technique for human activity identification using mimo radar," *International Symposium on Antennas and Propagation*, pp. 1–2, 2017.

[7] I. Alujaim, I. Park, and Y. Kim, "Human motion detection using planar array fmcw radar through 3d point clouds," *14th European Conference on Antennas and Propagation*, pp. 1–3, 2020.

[8] A. Sengupta, F. Jin, R. Zhang, and S. Cao, "mm-pose: Real-time human skeletal posture estimation using mmwave radars and cnns," *IEEE Sensors Journal*, 2020.

[9] P. Molchanov, S. Gupta, K. Kim, and K. Pulli, "Short-range fmcw monopulse radar for hand-gesture sensing," *2015 IEEE Radar Conference (RadarCon)*, pp. 1491–1496, 2015.

[10] E. Kim, Y. Yamada, and S. Okamoto, "Improvement of safety integrity level by multiplexing radio wave sensors," *System Integration, IEEE/SICE International Symposium*, pp. 942–947, 2017.

[11] Safety of machinery – Positioning of safeguards with respect to the approach speeds of parts of the human body, ISO 13855:2010.

[12] E. Kim, R. Kirschner, Y. Yamada, and S. Okamoto, "Estimating probability of human hand intrusion for speed and separation monitoring using interference theory," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101819, 2020.

[13] P. Hügler, M. Geiger, and C. Waldschmidt, "Rcs measurements of a human hand for radar-based gesture recognition at e-band," *German Microwave Conference*, pp. 259–262, 2016.

[14] X. Quan, J. W. Choi, and S. H. Cho, "A new thresholding method for ir-uwb radar-based detection applications," *Sensors*, vol. 20, no. 8, p. 2314, 2020.

[15] S. Lee, S. Cho, S. Sim, K. Kwak, Y. W. Park, and K. Cho, "A dynamic zone estimation method using cumulative voxels for autonomous driving," *International Journal of Advanced Robotic Systems*, vol. 14, no. 1, p. 1729881416687130, 2017.

[16] Safety of machinery - Safety-related sensors used for the protection of persons, IEC/TS 62998.