Postprint

N.B. When citing this work, cite the original published paper.

2D spatial keystone transform for sub-pixel motion extraction from noisy occupancy grid map

Hongqi Fan, Dawei Lu
ATR Laboratory
National University of Defense Technology
Changsha, P. R. China
fanhongqi@nudt.edu.cn, davidloo.nudt@gmail.com

Tomasz P. Kucner, Martin Magnusson and Achim Lilienthal
Center of AASS
Örebro University
Örebro, Sweden
{Tomasz.Kucner, Martin.Magnusson, achim.lilienthal}@oru.se

Abstract—In this paper, we propose a novel sub-pixel motion extraction method, called as Two Dimensional Spatial Keystone Transform (2DS-KST), for the motion detection and estimation from successive noisy Occupancy Grid Maps (OGMs). It extends the KST in radar imaging or motion compensation to 2D real spatial case, based on multiple hypotheses about possible directions of moving obstacles. Simulation results show that 2DS-KST has a good performance on the extraction of sub-pixel motions in very noisy environment, especially for those slowly moving obstacles.

Index Terms—robotics, occupancy grid map, motion extraction, keystone transform, 2DS-KST, sub-pixel

I. INTRODUCTION

Efficient perception of and reasoning about environments is still a major challenge for mobile robots operating in dynamic, densely cluttered or highly populated environments [1]–[3], which usually includes three interleaved tasks, i.e., Simultaneous Localization And Mapping (SLAM) [4], [5], Detection And Tracking of Moving Objects (DATMO) [6] and Cell Transition Mapping (CTM) [7].

For the three tasks above, one of the most important things is the fast and reliable extraction of motion information from successive sensor observations, which affects the performance of localization, mapping, moving object tracking directly [8]. Meanwhile, as the most popular one among all environment representations, Occupancy Grid Map (OGM) proposed by Elfes [9], [10] maps the environment as an array of probabilistic cells and easily integrates scans from multiple sensors [11], so the motion extraction from successively OGMs drawn especially attentions in the mobile robotics.

Among all literatures so far, extracting motion information from OGMs have been done mainly under the framework of DATMO and Bayesian Occupancy Filter (BOF), which work well for motions across pixels, i.e. fast moving objects, when OGMs are not severely corrupted by noise, such as false alarms, occlusions, and miss detections. See [6] and the latest survey [12] for more details about DATMO, and see [1], [13]–[17] for different BOFs and their particle implementations. But when considering the case of slowly moving obstacles under very noisy circumstance, the performance of existing algorithms under DATMO and BOF will degrade severely, which commonly occurs many false alarms in order to detection the motion with a small velocity. In order to overcome the problem, this paper proposes a novel sub-pixel motion extraction method, called as Two Dimensional Spatial Keystone Transform (2DS-KST), which extends the KST in radar imaging or motion compensation [18] to 2D real spatial case, based on multiple hypotheses about possible directions of moving obstacles.

The rest of this paper is organized as follows. One Dimensional Spatial KST (1DS-KST) is derived first in terms of one dimensional OGM data in Section II for the sake of understanding, then 2DS-KST with multiple hypotheses is proposed in Section III in term of the sequence of 2D OGMs. Meanwhile, a simple motion detection and estimation method based on the result of 2DS-KST is given in this section. In Section IV, we demonstrate the validity of 1DS-KST and 2DS-KST through the simulation data of point object. Conclusions and the further work may be found in Section V.

II. ONE DIMENSIONAL SPATAIL KST

Occupancy grid maps model the environment as an array of cells. Typically, these are layered out in a two-dimensional grid. However, we first discuss the keystone transform for the one dimensional case, which is called as 1DS-KST hereinafter. One reason is that there has a prior straight line constraint with the motion of objects in many applications, for instance, cars moving on the highway or city roads. Another reason is that 1DS-KST, where the concept of fast time is instead by one dimensional spatial grid, has a more direct relationship with KST in radar signal processing. It is helpful to understand the principle of keystone transform.

For 1DS-KST, the main assumptions are the following:

- The velocities for all moving objects are constant during N successive frames.
- $R \geq 2 \cdot V_{\max} \cdot T$, i.e., $V_{\max} \leq R/(2T)$, where R, T, $V_{\max}$ denote the grid size, the sampling interval and the maximum speed of objects, respectively.
- The sensor is motionless or the motion of it has been compensated by SLAM or other methods.

The first assumption always holds as long as the total length of time window $N \cdot T$ is very short or the amount of velocity change is less than one velocity resolution cell of KST. The second assumption is borrowed from [7], and it is derived from

the Nyquist Sampling condition, which ensures the temporal signal of the occupancies in every grid cell are not aliased at a time sampling period T. Meanwhile it ensures the spatial continuity of the motion of objects, which is necessary to filter the non-continuous changes of OGMs. In fact, for typical applications and modern sensors, this condition is easy to meet. For example, if the size of grid cell R is equal to 2 meters and the time sampling period T is equal to 0.1 seconds, the maximum velocity of objects is 10 m/s, which is enough high for most dynamic environment monitoring applications involving pedestrians, industrial robots and vehicles. As for those applications having higher maximal speed, such as automotive application, we can use a larger R or smaller T in order to avoid objects "jumping over" adjacent cells. As the focus of this paper is to extract the motion information from OGMs, the third assumption is natural and it can make this problem isolated from other problems, such as registration and localization.

Let us first consider the case that there is only one occupied grid cell, called as an ideal point object blow, in the sensor field of view. Assume it is moving at a constant velocity V, $V \in [-V_{\max}/2, +V_{\max}/2)$, and has an initial position $l_0 R$, then for any given time instant t, the obtained OGM has the following form:

$$f_t(l) = \delta_{r_t}(l) \tag{1}$$

where δ denote a unit pulse function, which is defined as

$$\delta_{r_t}(l) = \begin{cases} 1 & \text{if } r_t \in [lR - \frac{R}{2}, lR + \frac{R}{2}) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

And in (1) r_t is the position of this object at time instant t, which can be written as

$$r_t = l_0 R + V t \tag{3}$$

The Discrete Fourier Transform $\mathcal{F}_t(i)$ of $f_t(l)$ can be given as follows

$$\mathcal{F}_t(i) \overset{\text{def.}}{=} \sum_{l=0}^{L-1} f_t(l) \cdot \exp\left(-\iota 2\pi \frac{l \cdot i}{L}\right) \tag{4}$$

$$\cong \exp\left(-\iota 2\pi \frac{r_t \cdot i}{LR}\right) \tag{5}$$

$$= \exp\left(-\iota 2\pi \frac{l_0 i}{L}\right) \cdot \exp\left(-\iota 2\pi \frac{V \cdot t \cdot i}{LR}\right) \tag{6}$$

where ι denotes the imaginary unit.

Since the signal $f_t(l)$ is a real signal in the case of OGM, $\mathcal{F}_t(i)$ always satisfies conjugate symmetry, i.e.,

$$\mathcal{F}_t(i) = \mathcal{F}_t^*(L - i) \tag{7}$$

Therefore, we only concern the non-negative spatial frequency cells, that is, cells of $i = 0, \ldots, L/2 - 1$.

To use KST, we need choose a fixed cell of spatial frequency i_c as a reference. In general, the center frequency within the effective bandwidth of signal is chosen for a reference in the application of radar signal processing. For our case of 1D-OGM, we can multiply $\mathcal{F}_t(i)$ by a window $w(i)$ in
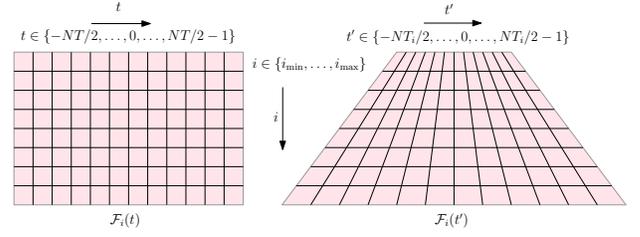

Fig. 1. Sampling pattern of Keystone transform.

the spatial frequency domain, which corresponds to a spatial filtering process and results in a blurred OGM. Without loss of generality, we denote this window $W(i)$ as the following:

$$W(i), \quad i_{\min} \leq i \leq i_{\max} \tag{8}$$

So $i_c = (i_{\min} + i_{\max})/2$ can be selected as the reference. Thus (6) can be expressed as

$$\mathcal{F}_t(i) = \exp\left(-\iota 2\pi \frac{l_0 i}{L}\right) \cdot W(i) \cdot \exp\left(-\iota 2\pi \frac{V i_c}{LR} \cdot \frac{i}{i_c} t\right),$$
$$i_{\min} \leq i \leq i_{\max}. \tag{9}$$

If let $t' = \frac{i}{i_c} \cdot t$, we can get

$$\mathcal{F}_t(i) = \exp\left(-\iota 2\pi \frac{l_0 i}{L}\right) \cdot W(i)$$
$$\cdot \exp\left(-\iota 2\pi \frac{V i_c}{LR} \cdot t'\right) \tag{10}$$

$$\overset{\text{def.}}{=} \mathcal{F}_i(t') \tag{11}$$

Now it is the time to consider a temporal variable t. During the time interval $[-NT/2, NT/2)$, we obtained N successive OGMs at some discrete time instants. Without loss of generality, we can assume that $t = -NT/2, \ldots, 0, T, \ldots, NT/2 - T$, that is to say, we obtain many signals $\mathcal{F}_i(t')$ at some time instants t'.

Since the scale factor i/i_c between t and t' is variable with i, the sampling period T_i and total time interval NT_i for t' are both different in terms of i. The sampling patterns of $\mathcal{F}_t(i)$ and $\mathcal{F}_i(t')$ are shown as Fig.1. In Fig.1, the sampling pattern of $\mathcal{F}_i(t')$ is like a keystone shape, that's why this corresponding transform is called as KST.

To correct the keystone effect in the sampling pattern of $\mathcal{F}_i(t')$, we need to compute the values at the discrete time $t' = -NT/2, \ldots, 0, T, \ldots, NT/2 - T$ for every i, which are usually obtained by an interpolate filter $h_i(n')$ in the context of KST. See [18] for more details of the interpolate filter $h_i(n')$. Thus, after the interpolate filtering,

$$\tilde{\mathcal{F}}_i(n) = \mathcal{F}_i(n') \otimes h_i(n') \tag{12}$$

$$\cong \exp\left(-\iota 2\pi \frac{l_0 i}{L}\right) \cdot W(i)$$
$$\cdot \exp\left(-\iota 2\pi \frac{V i_c}{LR} \cdot nT\right),$$
$$n = -N/2, \ldots, 0, 1, \ldots, N/2 - 1 \tag{13}$$

Then the IDFT transform of (12) in terms of i will give the following result:

$$\tilde{f}_n(l) = \delta(l - l_0) \otimes w(l) \cdot \exp\left(-\iota 2\pi \frac{V i_c}{LR} \cdot nT\right) \quad (14)$$

$$= w(l - l_0) \cdot \exp\left(-\iota 2\pi \frac{V i_c}{LR} \cdot nT\right) \quad (15)$$

$$= \begin{cases} w(0) \cdot \exp\left(-\iota 2\pi \frac{V i_c}{LR} \cdot nT\right) & \text{if } l = l_0 \\ w(l - l_0) \cdot \exp\left(-\iota 2\pi \frac{V i_c}{LR} \cdot nT\right) & \text{otherwise} \end{cases} \quad (16)$$

where, $w(l)$ is coefficients of spatial filter corresponding with the window $W(i)$.

In general, we must choose an appropriate window type and a suitable width of $W(i)$ so that $w(i)$ has a locally compact main-lobe and a side-lobe low enough. Thus, for those cells of $l \neq l_0$, the non-zero $\tilde{f}_n(l)$ will not affect the analysis of velocity as long as the distance of object is larger than the width of main-lobe of $w(i)$. This is easy to meet if the moving objects are not closely spaced pixel by pixel. Fortunately, this is the fact in OGM case. In fact, even for the superpositional target, if they have different velocities, the non-zero coefficients of $w(l)$ at $l \neq 0$ have no effect to velocity analysis as well. Therefore, let us consider the cell l_0 as the next step,

$$\tilde{f}_n(l_0) = w(0) \cdot \exp\left(-\iota 2\pi \frac{VT i_c}{LR} \cdot n\right) \quad (17)$$

After doing the DFT for $\tilde{f}_n(l_0)$ in terms of n, we can get:

$$\tilde{\mathcal{F}}_{l_0}(k) = \sum_{n=-N/2}^{N/2-1} w(0) \cdot \exp\left(-\iota 2\pi \frac{VT i_c}{LR} \cdot n\right)$$
$$\cdot \exp\left(-\iota 2\pi \frac{nk}{N}\right) \quad (18)$$

$$= \sum_{n=-N/2}^{N/2-1} w(0)$$
$$\cdot \exp\left(-\iota 2\pi n \left(\frac{VT i_c}{LR} + \frac{k}{N}\right)\right) \quad (19)$$

As (19) shown, different velocities V will be located at the different frequency grid cells k. The KST method, therefore, allow for different velocities in a single cell, which is similar to 4D-BOF [1] and BOFUM [14]. For an object having the velocity V, we have the following approximation:

$$\frac{k}{N} \cong -\frac{VT i_c}{LR} \Rightarrow V \cong -\frac{kLR}{NT i_c} \quad (20)$$

The corresponding resolution of the above velocity measurement is

$$\Delta V = \frac{LR}{NT i_c} \quad (21)$$

and the normalized one is

$$\Delta \mathsf{V} = \frac{\Delta V \cdot T}{R} = \frac{L}{N i_c} \quad (22)$$

We can choose i_c and N according to the dynamic characteristics of environment and the velocity resolution of interest which is related to the minimal detectable velocity.

III. Two Dimensional Spatial KST

A one-dimensional OGM is of limited practical use. For mobile robots, the two-dimensional OGM is the usual case. This section will develop a method of two dimensional spatial KST (2DS-KST) for the purpose of extracting motion information from successive OGMs. Besides those assumptions in section II, two additional ones needed here are as the following:

- The moving directions of all objects are kept constant but unknown during N frames.
- The unknown moving directions belong to a prior set with finite number of elements.

A. 2DS-KST with multiple hypotheses

Let us still consider only one ideal point object in the sensor field of view. Assume its initial position is $\mathbf{r}_0 = [l_0, m_0]^T$ and it has a constant velocity $\mathbf{V} = [V_x, V_y]^T = [V\cos(\theta), V\sin(\theta)]^T$, in which V_x, V_y satisfy the condition of $V_x, V_y \in [-V_{\max}/2, +V_{\max}/2)$. Then for a given time instant t, the OGM has the following form:

$$\delta_{\mathbf{r}_t}(l, m) = \begin{cases} 1 & \text{if } [r_{xt}, r_{yt}]^T \in \text{Rect}(l, m) \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

where $\text{Rect}(l, m)$ denote the region of the cell (l, m).

For the 2D spatial case, equation (5) becomes as follows:

$$\mathcal{F}_t(i, j) = \exp\left(-\iota 2\pi \frac{\mathbf{r}_0 \cdot \mathbf{i}}{L}\right) \cdot \exp\left(-\iota 2\pi \frac{V \mathbf{u}_\theta \cdot \mathbf{i}}{LR} t\right) \quad (24)$$

$$= \exp\left(-\iota 2\pi \frac{\mathbf{r}_0 \cdot \mathbf{i}}{L}\right) \cdot \exp\left(-\iota 2\pi \frac{V i_\theta}{LR} t\right) \quad (25)$$

where $\mathbf{i} = [i, j]^T$, $\mathbf{u}_\theta = [\cos\theta, \sin\theta]^T$ and $i_\theta = i\cos\theta + j\sin\theta$.

As the above described, we assume there are finite possible hypotheses for θ, and denote the pth hypothesis as θ_p ($p = 1, \ldots, \nu$). But the real case is we don't know the actually moving direction of the object, so we need to do KST for every possible hypothesis, then we can get:

$$\mathcal{F}_t^{\theta_p}(i, j) = \exp\left(-\iota 2\pi \frac{\mathbf{r}_0 \cdot \mathbf{i}}{L}\right) \cdot W_{\theta_p}(\mathbf{i})$$
$$\cdot \exp\left(-\iota 2\pi \frac{V_{\theta_p} i_c^{\theta_p}}{LR} \cdot \frac{i_{\theta_p}}{i_c^{\theta_p}} \cdot t\right)$$
$$\cdot \exp\left(-\iota 2\pi \frac{V_{\theta_p^\perp} i_{\theta_p^\perp}}{LR} \cdot t\right) \quad (26)$$

where

- V_{θ_p}, $i_c^{\theta_p}$ and i_{θ_p} are projections of vectors $\mathbf{V}$, $\mathbf{i}_c^{\theta_p}$ and $\mathbf{i}$ along the θ_p direction, respectively, while $\mathbf{i}_c^{\theta_p}$ is the reference spatial frequency vector for θ_p hypothesis when doing KST in next step;
- $V_{\theta_p^\perp}$, $i_{\theta_p^\perp}$ are projections of vectors $\mathbf{V}$, $\mathbf{i}$ perpendicular to the θ_p direction, respectively;
- $W_{\theta_p}(\mathbf{i})$ is the two dimensional window function for θ_p hypothesis which has the same meaning as the window (8) in one dimensional case.

(26) result from the fact that the item $V i_\theta$ in (25) can be rewritten as:

$$V i_\theta = V_{\theta_p} i_{\theta_p} + V_{\theta_p^\perp} i_{\theta_p^\perp} \qquad (27)$$

Furthermore, if the hypothesis θ_p is true, that is to say, θ is approximately parallel to θ_p, the last product item in (26) can be ignored further. In this case, for the θ_p hypothesis, (26) can be approximated as the following:

$$\mathcal{F}_t^{\theta_p}(i,j) \cong \exp\left(-\iota 2\pi \frac{\mathbf{r}_0 \cdot \mathbf{i}}{L}\right) \cdot W_{\theta_p}(\mathbf{i})$$
$$\cdot \exp\left(-\iota 2\pi \frac{V_{\theta_p} i_c^{\theta_p}}{LR} \cdot \frac{i_{\theta_p}}{i_c^{\theta_p}} \cdot t\right) \qquad (28)$$

$$\stackrel{\text{def.}}{=} \mathcal{F}_\mathbf{i}^{\theta_p}(t') \qquad (29)$$

where t' is defined as:

$$t' = \frac{i_{\theta_p}}{i_c^{\theta_p}} t \qquad (30)$$

The interpolate filter of Keystone transform is the same as 1D case, so we can get $\tilde{f}_n^{\theta_p}(l_0, m_0)$ as follows:

$$\tilde{f}_n^{\theta_p}(l_0, m_0) = w_{\theta_p}(0,0) \cdot \exp\left(-\iota 2\pi \frac{V_{\theta_p} T i_c^{\theta_p}}{LR} \cdot n\right) \qquad (31)$$

The counterpart of the above equation in 1D case is (17). After doing the DFT for $\tilde{f}_n^{\theta_p}(l_0, m_0)$ in terms of n for every occupied grid cell $[l_0, m_0]^T$, we can get:

$$\tilde{\mathcal{F}}_{l_0, m_0}^{\theta_p}(k) = \sum_{n=-N/2}^{N/2-1} w_{\theta_p}(0,0)$$
$$\cdot \exp\left(-\iota 2\pi n \left(\frac{V_{\theta_p} T i_c^{\theta_p}}{LR} + \frac{k}{N}\right)\right) \qquad (32)$$

Thus, similar to 1D case, V_{θ_p}, the magnitude of the velocity in the grid cell $[l_0, m_0]^T$ can be given as follows:

$$V_{\theta_p} \cong -\frac{kLR}{NT i_c^{\theta_p}} \qquad (33)$$

while the direction of the velocity in this cell is given by θ_p, which is the assumption given in the previous, once used in (28).

B. Merging multiple hypotheses

The velocity measurement given by (33) is in the case of the θ_p hypothesis is true for this cell. In practice, we don't know which hypothesis is true for any occupied grid cell $[l, m]^T$ in advance, so we must merge the results of those multiple hypotheses so that the correct motion information can be given.

In fact, by 2DS-KST with multi-hypothesis, we can get ν cuboid $\tilde{\mathcal{F}}^{\theta_p}(l, m, k)$, which is the general form of (32) for any arbitrary grid cell $[l, m]^T$ under the hypothesis θ_p. That is to say, we get the two dimensional matrix $\tilde{\mathcal{F}}_{l,m}(\theta_p, k)$ for every grid cell $[l, m]^T$, which represents the possible velocity in this cell, including the magnitude (denoted by k, which can be negative) and the direction (denoted by θ_p).

Here we use a MPD (Maximal Power Detector) based method for the purpose of extracting the motion information from $\tilde{\mathcal{F}}_{l,m}(\theta_p, k)$, which is based on the fact that the power item $|\tilde{\mathcal{F}}_{l,m}(\theta_p, k)|^2$ will be larger when θ_p and k are more matched with the true value of the velocity. It can be described as the following four steps:

- The first step is the maximal merging step:

$$\mathcal{P}(l, m) = \max_{\theta_p, k} |\tilde{\mathcal{F}}_{l,m}(\theta_p, k)|^2 \qquad (34)$$

- The second is the following power detector:

$$\mathcal{P}(l, m) \underset{H0}{\overset{H1}{\gtrless}} P_{\min} \qquad (35)$$

where $H1$ and $H0$ denote the event whether the grid cell to be determined is occupied or not, respectively, and where $P_{\min}$ denote the threshold of MPD.

- The third is the following estimator if $H1$ holds on:

$$(\hat{k}, \hat{\theta}) = \arg\max_{\theta_p, k} |\tilde{\mathcal{F}}_{l,m}(\theta_p, k)|^2 \qquad (36)$$

$$\hat{\mathbf{V}}_{l,m} = \hat{V}_{\hat{\theta}} \cdot \exp(\iota\hat{\theta}) \qquad (37)$$

where $\hat{V}_{\hat{\theta}}$ can be obtained according to (33) if $\hat{k}$ given.

- The last step is the separator as follows:

$$|\hat{\mathbf{V}}_{l,m}| \underset{S}{\overset{D}{\gtrless}} V_{\min} \qquad (38)$$

where D and S denote the event whether the grid cell to be determined is dynamic or not, respectively, and where $V_{\min}$ denote the threshold of the separator, which directly related to the minimal detectable velocity or ΔV in (21).

As the above MPD approach of merging multiple hypotheses only outputs the most likely point estimation for the velocity, it is obvious that it is only appropriate for the case when the occupancies in each grid cell only have one significant velocity during the time interval. Nevertheless, it does not affect the versatility of 2D-KST for extracting motion information. For instance, we can use the more complex GMM (Gaussian Mixture Model) instead of the point estimation in MPD method as the velocity model of occupancy, which would undoubtedly increase the robustness and the compatibility to the complex situations. In fact, which velocity model should be selected is a key problem in practice and it closely depends on the physical application.

For the more complex method of merging multiple hypotheses, it is already beyond the scope of this paper and may be an appropriate topic for the future work.

IV. SIMULATION RESULTS

In this section, we design an experiment to demonstrate the validity of our method through the simulation data of point object. Considering the sensor noise and the imperfections in the OGM building process, we add to the OGMs some Poisson noise, which is uniformly distributed in the grid cells and whose number obeys the Poisson distribution. The main parameters in the simulation is shown in the Table I, where

TABLE I
SIMULATION PARAMETERS IN THE EXPERIMENT I

	Parameter	1D case	2D case
OGMs	L	128	64
	N	100	40
Objects	#0	$20, 0$	$10, 10, 0, -$
State	#1	$40, -0.5$	$20, 15, 0.5, 0°$
$l_0, m_0, \mathsf{V}, \theta$	#2	$60, 0.05$	$30, 20, 0.1, 90°$
	#3	$80, -0.2$	$35, 30, 0.2, 45°$
	#4	$100, 0.1$	$40, 40, 0.3, 135°$
	#5	—	$45, 50, 0.4, 165°$
Poisson Rate	λ	16	64
Windows	θ_p	—	$0 : \pi/8 : 7\pi/8$
	$i_c^{\theta_p}$	$L/4$	$L/4\alpha$
	$i_{\min}$	$L/8$	$i_c^{\theta_p}/2$
	$i_{\max}$	$3L/8$	$3i_c^{\theta_p}/2$
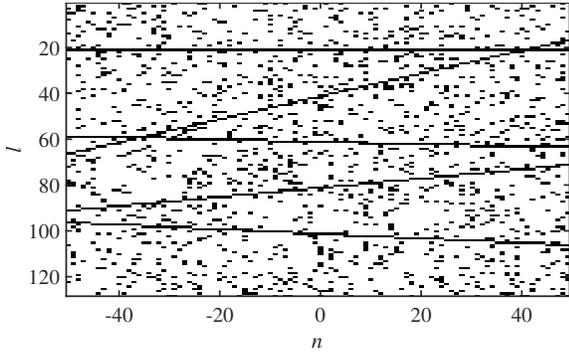
Fig. 2. Sequence of the simulated one dimensional OGMs: $f(l, n)$.

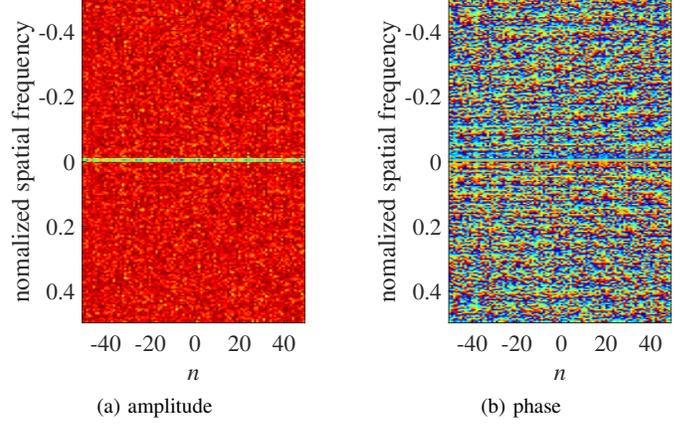

(a) amplitude   (b) phase

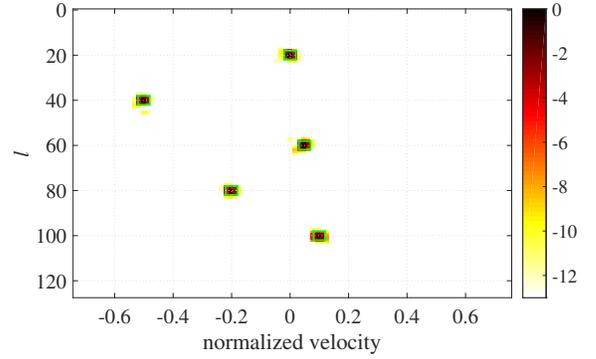Fig. 3. Result of OGM sequence by the spatial Fourier transform: $\mathcal{F}(i, n)$.

Fig. 4. Result of 1DS-KST: $\tilde{\mathcal{F}}(l, k)$. The color denotes the total power of the accumulated occupancies in this cell during the N time instants. The green box on every object indicates the corresponding resolution of 1D-KST for the position and the velocity.

the velocities V are represented in the normalized format, and the scale factor α in $i_c^{\theta_p}$ is defined as follows:

$$\alpha = \max(|\cos(\theta_p)|, |\sin(\theta_p)|) \tag{39}$$

The results of the one dimensional test are shown in Fig. 2 through 4. From Fig. 2, we can see five objects moving at different speeds along positive direction or negative direction. The speed of the fastest one is 0.5, which is the maximal feasible velocity given by the Nyquist sampling theorem, and the speed of the slowest moving object is 0.05, which is very close to the velocity resolution given by (22). Through this setting, we can validate 1DS-KST in terms of velocity measurement capability. Meanwhile, we set a stationary object at $l_0 = 20$, which can be used to check the performance on the separation of the dynamic grids from the static ones. After adding the Poisson noise, the OGMs look very noisy.

The result of OGM sequence by the spatial Fourier transform is shown in Fig. 3. As the noise in the OGMs, it is hardly to see any change information along the time axis from Fig. 3a. However, it can still be seen from the phase of $\mathcal{F}(i, n)$ in

Fig. 3b. Moreover, it is not hardly to see the symmetry with respect to the spatial frequency axis from Fig. 3 because our OGMs are all real numbers, thus from the information point of view we can only set the spatial frequency window in one side of the spatial frequency axis.

The final result of 1DS-KST is shown in Fig. 4. When looked together with Fig. 2, it is clearly shown that all objects are well located in the grid cells at the midpoint time instant and their velocities are also measured with a high precision even for these so noisy OGMs[1]. Moreover, the occupies of every are well focused in the green box determined by 1D-KST's resolution, which is determined by (22) and the width of the spatial frequency window:

$$\Delta l \cong L/(i_{\max} - i_{\min}) \tag{40}$$

From this result, we can conclude that 1DS-KST has a good performance in terms of OGM filtering and the extraction of motion information.

---

[1]If we want to obtain the OGM at any time instant $n = 0, 1, \ldots, N-1$, we only need multiply $\tilde{\mathcal{F}}(i, k)$ by a phase corrected item before the processing of spatial IFFT.
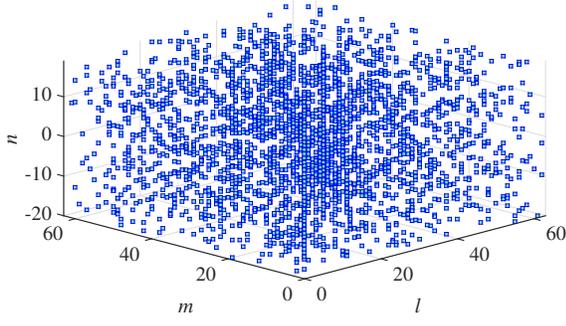
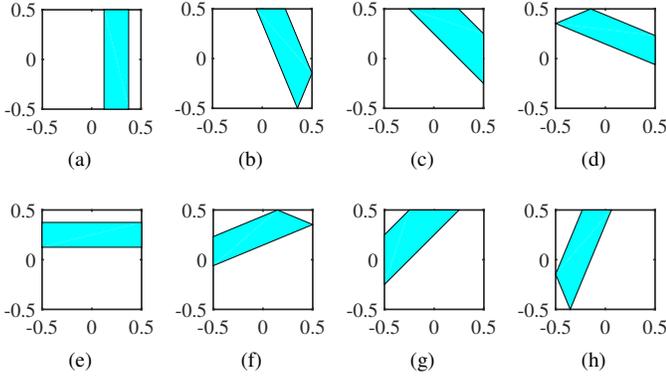Fig. 5. Sequence of the simulated two dimensional OGMs: $f(l, m, n)$.

Fig. 6. Spatial frequency window $W_{\theta_p}(i, j)$. (a)-(h) are the cases of $\theta_p = 0, \pi/8, \ldots, 7\pi/8$, respectively. The x-ticks and y-ticks are both the normalized values of spatial frequency. The value of the point inside the colorful quadrangle is equal to 1, while zero for the outside point.

The results of the two dimensional test are shown in Fig. 5 through 9. Similar to the one dimensional case, we choose the maximal feasible and minimal detectable speed to are set to evaluate the capability of 2DS-KST in terms of velocity measurement. Different from one dimensional case, we use eight hypotheses of direction to match the possible moving directions in OGMs. Among all five moving objects, four are moving along the direction in the hypothesis sets, while the last one (#5) is moving near the middle direction between the hypothesis $\theta_p = 7\pi/8$ and the reverse direction of $\theta_p = 0$, but slightly close to $\theta_p = 7\pi/8$. Through this setting, we can evaluate the performance when the true moving direction does not match any hypotheses. To see the result more clearly, we decrease the map size $L = 64$ in the two dimensional test. The detailed parameter setting used in this simulation can be seen in Table I.

Fig. 5 shows the noisy 2D-OGM sequence. It is very difficult to recognize the trajectory of each object by human eyes from this figure. To make the window $W_{\theta_p}(i, j)$ more intuitive to the reader, we show them in Fig. 6. In principle, the counterpart of the spatial frequency window here is the directional filter in optical signal processing.

After the first merging step of MPD, the result $\mathcal{P}(l, m)$ of 2DS-KST processing is shown in Fig. 7, which can be regard as the filtered OGM at the time instant $n = 0$.
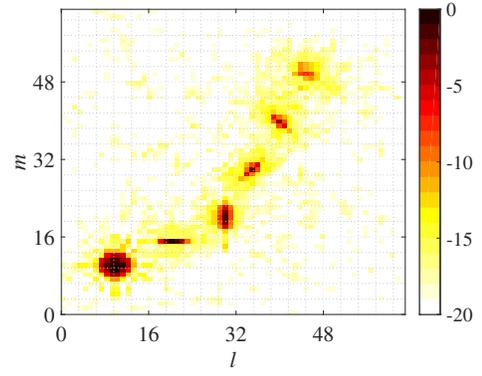

Fig. 7. Result of 2DS-KST after the first merging step of MPD: $\mathcal{P}(l, m)$. The color denotes the total power of the accumulated occupancies in this cell during the N time instants.
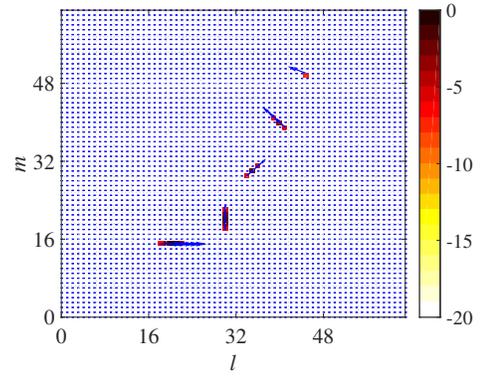

Fig. 8. Result of 2DS-KST after the processing of MPD. The color has the same meaning as Fig. 7. The blue arrows denote the velocities of the grid cells, whose lengths are proportional to the speed.
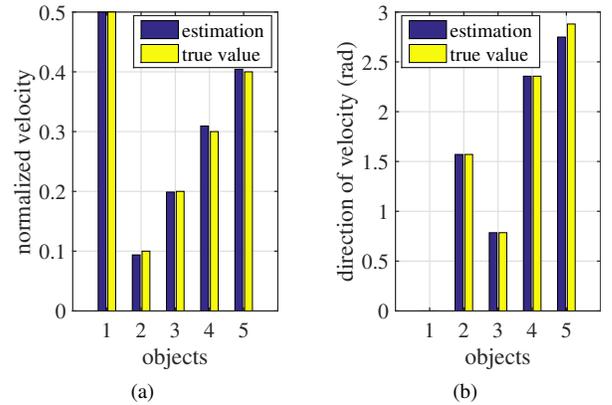

Fig. 9. Comparison of the measurement and the true value of the velocity.

From this figure, we can found that the noise in the original OGM is almost filtered, and the grid cells in the vicinity of objects have a significant occupancy value. The occupancies in the yellow grid cells around each object are the leakage of occupancy in the corresponding object grid cell, which is the consequence caused by the window $W_{\theta_p}(i, j)$. Among all the six objects, the stationary object at $(10, 10)$ behaves

obviously isotropic, while the other moving objects behave the obvious directionality, which means that they have the biggest extent along their moving direction. Moreover, we can found that the grid cells near the mismatched one at $(45, 50)$ have a relatively smaller value than those near the matched objects. Nonetheless, their values are at least twice as large as the values of those yellow grid cells. Thus, if we set the appropriate threshold $P_{\min}$, see (35), we can remove the leakage occupancies in the yellow cells but maintain those in the vicinity of the mismatched object grid cell. Furthermore, if our requirement is to extract the moving objects, we can filter the stationary one by setting the appropriate threshold $V_{\min}$, see (38).

The result of 2DS-KST after MPD processing is shown in Fig. 8, where $P_{\min} = 0.3981$ (-8dB) and $V_{\min} = 0.085$ are used, and where the blue arrows represent the velocity of the occupies in the corresponding cells, the length and pointing for the amplitude and the direction respectively. As can be seen from Fig. 8, the stationary object has been removed successfully, although it has the maximal occupancy, while the five moving objects all persist in existing. What's more, the velocity measurements are highly in accordance with the ground truth in Table I. To see it more obviously, the simple plot extractor is used, which extracts the local power maximum grid cells from the MPD results as the candidate detections and computes the occupancy weighted velocity as the velocity measurement of each candidate detection. The measurement results are shown in Fig. 9, which suggest that 2DS-KST has a good precision of velocity measurement and can be easily integrated with the plot extractor or other object clustering algorithm, such as FCTA [19].

V. Conclusions

This paper proposed a novel sub-pixel motion extraction method, called as 2DS-KST, for the motion detection and estimation from successive noisy OGMs. Simulation results show that our method can extract the sub-pixel motions effectively from the sequence of very noisy OGMs, which has a wide use, such as the industrial field, airport and other indoor environment.

Further evaluation by real data, hypothesis merging method based on more complex velocity model, and multi-resolution 2DS-KST for the case of across pixel and sub-pixel simultaneously existing are the work in the next step.

References

[1] C. Cou, C. Pradalier, C. Laugier, T. Fraichard, and P. Bessire, "Bayesian occupancy filtering for multitarget tracking: An automotive application," *The International Journal of Robotics Research*, vol. 25, no. 1, pp. 19–30, 2006. [Online]. Available: http://ijr.sagepub.com/content/25/1/19.abstract

[2] M. Tay, K. Mekhnacha, M. Yguel, C. Cou, C. Pradalier, C. Laugier, T. Fraichard, and P. Bessire, "The Bayesian occupation filter," in *Probabilistic Reasoning and Decision Making in Sensory-Motor Systems*, ser. Springer Tracts in Advanced Robotics, P. Bessire, C. Laugier, and R. Siegwart, Eds. Springer Berlin Heidelberg, 2008, vol. 46, pp. 77–98. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-79007-5_4

[3] P. Kondaxakis, S. Kasderidis, and P. Trahanias, "Tracking multiple targets from a mobile robot platform using a laser range scanner," in *Target Tracking and Data Fusion: Algorithms and Applications, 2008 IET Seminar on*, April 2008, pp. 177–184.

[4] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part i," *Robotics & Automation Magazine, IEEE*, vol. 13, no. 2, pp. 99–110, June 2006.

[5] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (slam): part ii," *Robotics & Automation Magazine, IEEE*, vol. 13, no. 3, pp. 108–117, 2006.

[6] A. Petrovskaya, M. Perrollaz, L. Oliveira, L. Spinello, R. Triebel, A. Makris, J.-D. Yoder, C. Laugier, U. Nunes, and P. Bessiere, "Awareness of road scene participants for autonomous driving," in *Handbook of Intelligent Vehicles*. Springer, 2012, pp. 1383–1432.

[7] T. Kucner, J. Saarinen, M. Magnusson, and A. J. Lilienthal, "Conditional transition maps: Learning motion patterns in dynamic environments," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 1196–1201.

[8] C.-C. Wang, C. Thorpe, S. Thrun, M. Hebert, and H. Durrant-Whyte, "Simultaneous localization, mapping and moving object tracking," *Int. J. Rob. Res.*, vol. 26, no. 9, pp. 889–916, Sep. 2007. [Online]. Available: http://dx.doi.org/10.1177/0278364907081229

[9] A. Elfes, "Occupancy grids: A probabilistic framework for robot perception and navigation," Ph.D. dissertation, Pittsburgh, PA, USA, 1989, aAI9006205.

[10] ——, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, June 1989.

[11] Q. Baig, M. Perrollaz, and C. Laugier, "Advances in the Bayesian occupancy filter framework using robust motion detection technique for dynamic environment monitoring," *IEEE Robotics and Automation Magazine*, Mar. 2014. [Online]. Available: https://hal.inria.fr/hal-00932691

[12] A. Pancham, N. Tlale, and G. Bright, "Literature review of SLAM and DATMO," in *4th Robotics and Mechatronics Conference of South Africa*, ser. 7484. CSIR International Conference Centre, Pretoria: RobMech 2011, Nov 2011.

[13] C. Chen, C. Tay, C. Laugier, and K. Mekhnacha, "Dynamic environment modeling with gridmap: A multiple-object tracking application," in *Control, Automation, Robotics and Vision, 2006. ICARCV '06. 9th International Conference on*, Dec 2006, pp. 1–6.

[14] T. Gindele, S. Brechtel, J. Schroder, and R. Dillmann, "Bayesian occupancy grid filter for dynamic environments using prior map knowledge," in *Intelligent Vehicles Symposium, 2009 IEEE*, June 2009, pp. 669–676.

[15] S. Brechtel, T. Gindele, and R. Dillmann, "Recursive importance sampling for efficient grid-based occupancy filtering in dynamic environments," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, May 2010, pp. 3932–3938.

[16] R. Danescu, F. Oniga, and S. Nedevschi, "Modeling and tracking the driving environment with a particle-based occupancy grid," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 4, pp. 1331–1342, Dec 2011.

[17] A. Negre, L. Rummelhard, and C. Laugier, "Hybrid sampling Bayesian occupancy filter," in *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*, June 2014, pp. 1307–1312.

[18] M. A. Richards, "The keystone transformation for correcting range migration in range-doppler processing," *pulse*, vol. 1000, p. 1, 2014.

[19] K. Mekhnacha, Y. Mao, D. Raulo, and C. Laugier, "Bayesian occupancy filter based "Fast Clustering-Tracking" algorithm," in *IROS 2008*, 2008.