Differential Mapping Spiking Neural Network for Sensor-Based Robot Control

Omar Zahra[a], Silvia Tolu[b], and David Navarro-Alarcon[a]

[a]The Hong Kong Polytechnic University, Hong Kong
[b]Technical University of Denmark, Denmark

Abstract

In this work, a spiking neural network (SNN) is proposed for approximating differential sensorimotor maps of robotic systems. The computed model is used as a local Jacobian-like projection that relates changes in sensor space to changes in motor space. The SNN consists of an input (sensory) layer and an output (motor) layer connected through plastic synapses, with inter-inhibitory connections at the output layer. Spiking neurons are modeled as Izhikevich neurons with a synaptic learning rule based on spike timing-dependent plasticity. Feedback data from proprioceptive and exteroceptive sensors are encoded and fed into the input layer through a motor babbling process. As the main challenge to building an efficient SNN is to tune its parameters, we present an intuitive tuning method that considerably reduces the number of neurons and the amount of data required for training. Our proposed architecture represents a biologically plausible neural controller that is capable of handling noisy sensor readings to guide robot movements in real-time. Experimental results are presented to validate the control methodology with a vision-guided robot.

Keywords: Robotics, spiking neural networks, sensor-based control, visual servoing.

1 INTRODUCTION

Sensor-guided object manipulation is one of the most fundamental tasks in robotics, with many possible approaches to perform it [1]. Conventional methods typically rely on mathematical modeling of the observed end-effector pose and its related joint configuration. These methods provide accurate solutions, however, they require an exact knowledge of the analytical sensor-motor relations (which might not be known); furthermore, these conventional methods are generally not provided with adaptation mechanisms to cope with uncertainties/changes in the sensor setup.

Among the many interesting cognitive abilities of humans (and animals, in general) is the motor babbling process that leads to the formation of sensorimotor maps [2]. By relying on such adaptive maps, humans can learn to perform many motion tasks in an efficient way. These advanced capabilities have motivated many research studies that attempt to artificially replicate such skills in robots and machines [3, 4]. Our aim in this work is to develop a bio-inspired adaptive computational method to guide the motion of robots with real-time sensory information and a limited amount of data.

Data-driven computational maps have been previously built for approximating unknown

sensorimotor relations [5–7]. One common limitation of these classical approaches is the demand for a high number of training data points and high computational power, which is impractical in many cases.

By drawing inspiration from the central nervous system, artificial neural networks (ANN) have been built and used for many decades. It started with the McCulloch Pitts model as the first generation of ANN by using binary computing units [8], followed by the second generation utilizing mainly the sigmoidal (or tanh) activation functions to make it more capable of approximating non-linear functions [9]. Taking one step further, spiking neuronal networks (SNN) (representing the third generation of ANN) are designed to incorporate most of the neuronal dynamics which provide them with more complex and realistic firing patterns based on spikes [10]. In this work, we have built an adaptive SNN-based model to guide robots with sensory feedback.

The main exclusive property of SNN is the incorporation of a temporal dimension; note that the relative timing of spikes (and spikes sequences) enables the encoding of useful information. As in real biological systems, SNNs hold an advantage for real-time processing (as concluded in [11] for the visual system) and multiplexing of information (such as amplitude and frequency in the auditory system [12]). For robotic applications, the SNN allows building computational models of brain regions to imitate intelligent behaviour in living organisms to a great extent, and in some cases even reveal the mysteries of the inner workings of the brain [13–15]. The currently rising neuromorphic chips allow real-time operation of such models while conserving power greatly compared to conventional systems [16]. While an SNN allows building a more biologically plausible system, its complexity makes it more difficult to predict and analyze the system behaviour. To this end, various methods can be used, e.g. simplifying the network's mathematical description [17] or applying techniques for tuning the parameters to achieve the desired performance [18]. The latter approach is the one addressed in this work.

Several studies have provided examples of the application of SNN in robotics [19]. However, in most studies, robots were only controlled in a simulation environment [20]. The main reason is due to the large size of the neural networks used in these works, which makes it impractical for real-time operations. In [18], a cognitive architecture is used for controlling a robotic hand to perform grasping motions. In [21], an SNN was used for learning motion primitives of a robot so to move in three axes (left-right, up-down and far-near). In [22], a self-organizing architecture was used to build an SNN representing spatio-motor transformations (i.e. kinematics) of a two degrees of freedom (DOF) robot. Despite its good performance and biologically plausibility, the network's size and limited scalability make it impractical for real-time control.

In [23], an SNN with learning based on spike timing-dependent plasticity (STDP) was used to build the kinematics of a robot. Although the synaptic connections illustrate the ability to approximate the kinematic relation, the approximation error is not evident. Additionally, an intermediate layer is required to scale up the dimensions of the input sensory data, and consequently, the relative computational power.

In this paper, we propose an SNN-based control architecture to guide robots with sensor feedback. Our proposed neural controller adaptively builds the differential map that correlates end-effector velocities (as measured by an external vision sensor) with joint angular velocities. In other words, it effectively works as a local Jacobian-like transformation between different (sensor-motor) spaces. This new method is characterized by inter-inhibitory connections at the output layer,

real-time operation and a fewer number of neurons, compared to previous works in literature [18, 21, 24]. The inter-inhibitory connections distinguish the proposed architecture from that introduced in [24].

These inter-inhibitory connections with a winner-take-all (WTA) effect is studied as well from a probabilistic perspective in [25]. Such effect is proven to be an essential component of a special setup, that allows the emergence of an approximation of a forward sampler [26] in a Hidden Markov Model (HMM) [27,28] capable of online learning. Moreover, the study in [24] suggests to divide the input into separate bins to improve the learning but both the formulation and the explanation of how this is achieved is lacking, and the method to set and tune the network parameters is not mentioned. To the best of the authors' knowledge, this is the first study to report an SNN-based method capable of forming the sensor-motor differential map in a computationally efficient way (resulting from an intuitive guideline to adjust its parameters). Additionally, the network relies on biologically plausible models of neurons and synapses, which are more complex compared to other relatively simpler models, to both preserve features of the biological nervous system as much as possible and serve as a future building block for simulating the whole hierarchy of the biological motor system. An example of this is the simple Izhikevich neuron model [29] used in the proposed network, instead of the commonly used Leaky Integrate and Fire model (LIF) [30]. To validate the theory, we present a detailed experimental study with a robotic platform performing vision-guided manipulation tasks.

The rest of this paper is structured as follows: Sec. II describes the developed spiking neural network; Sec. III presents the verification of the method proposed through a dummy test; Sec. IV presents the experimental results; Sec. V discusses the methods and results; Sec. VI gives the conclusions.

2 METHODS

Many studies have suggested that humans use internal models to represent perception and action [31–33]. Some researchers have built computational models of brain areas (e.g the cerebellum) responsible for the generation and coordination of fine motor actions [34, 35]. Unlike in traditional robot control (where sensory transformations to motion commands are solved analytically), in the human brain (motor cortex) the sensorimotor relations are encoded by specific neural circuits. To carry out a typical (eye-to-hand) visual servoing task, we must first establish the kinematic relation between the joint velocities and the measured end-effector motion [36]. This model can be formulated as $\dot{x} = J(\theta)\dot{\theta}$, such that:

$$J(\theta) = \begin{bmatrix} \frac{\partial x_1}{\partial \theta_1} & \cdots & \frac{\partial x_1}{\partial \theta_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial \theta_1} & \cdots & \frac{\partial x_n}{\partial \theta_m} \end{bmatrix} \tag{1}$$

where $\dot{x} \in \mathbb{R}^n$, $\theta \in \mathbb{R}^m$, $\dot{\theta} \in \mathbb{R}^m$ and $J(\theta) \in \mathbb{R}^{n \times m}$ are the (observed) spatial velocity, joint angles, joint velocities and the Jacobian matrix, respectively (without loss of generality, we assume that $m \geq n$). For velocity-based control it is necessary to estimate the joint velocities to achieve a certain spatial velocity. This expression is denoted as:

$$\dot{\theta} = J^{\#}(\theta)\dot{x} \tag{2}$$

where $J^{\#}$ is the pseudo-inverse of the J. In this work, a differential mapping spiking neural network (DMSNN) is proposed to build a computational model analogous to the Jacobian transformation relating changes in 1 joint-space DoF to 1 task-space DoF.

As shown in Fig. 1, the network is first subject to the training phase in which both sensory

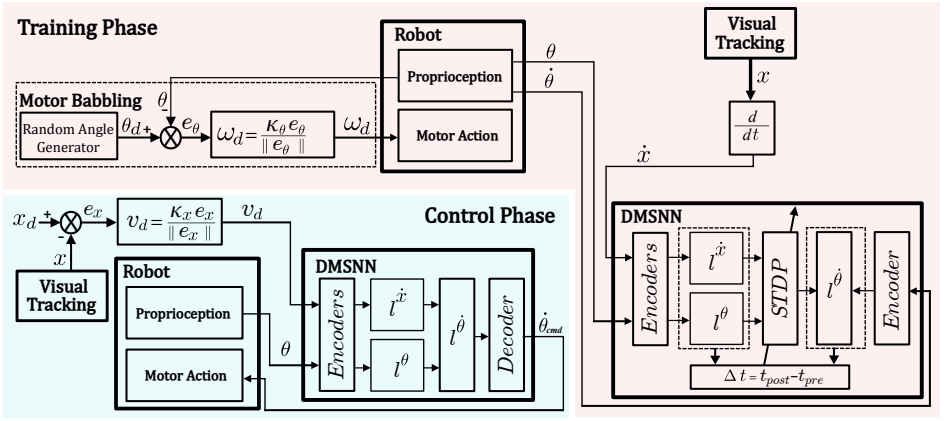


Figure 1: A schematic diagram for both the training and control phases for the DMSNN. The signals introduced to each neuron bundle are depicted. During the training phase, the robot motion is guided by motor babbling in joint space. During the control phase, the robot motion is guided by motor commands decoded from the activity of motor neurons at the output layer.

readings and motor commands are fed to the network through motor babbling by executing random motions. After training the network for several iterations, the differential mapping is formed by the modulation of the connection between the input and output layers. Then, the network can be used during the control phase to guide the robot through the estimation of the required motor command to reach the desired target by feeding the sensory information only to the input layer. This can be seen as an approximation of the motor cortex which is responsible for converting the desired motion into a motor command to be executed by other regions in the brain [37]. Details of the proposed network are described in the rest of this section.

2.1 Network Layout

The network layout of the proposed neural controller is shown in Fig. 2, where each dimension of sensory input (joint angles and spatial velocity) and motor output (joint velocity) is represented by a one-dimensional array (bundle) of neurons. The complete network consists of $n + 2m$ bundles of neurons. The input sensory layer consists of bundles $l^{\theta_{1:m}}$ encoding the joint angles and $l^{\dot{x}_{1:n}}$ encoding the spatial velocity, while the output motor layer consists of $l^{\dot{\theta}_{1:m}}$ encoding the joint velocity controls. Each sensory neuron is connected through an excitatory and inhibitory synapse (which is omitted in Fig. 2 for clarity) to each motor neuron. This acts as a substitute for adding an inhibitory interneuron and allows for stable learning dynamics while avoiding an unbounded increase in connection strength and

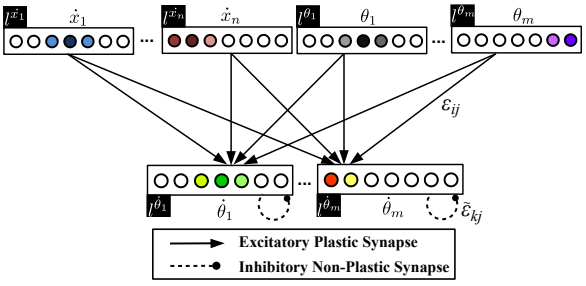


Figure 2: The layout of the proposed spiking neural network. Input (sensory) neurons are connected to output (motor) neurons through excitatory and inhibitory plastic synapses ε_{ij}; neurons in each motor bundle are interconnected through inhibitory non-plastic synapses $\tilde{\varepsilon}_{ij}$.

neuron activity. Additionally, each of the neurons encoding the output at each bundle (dimension) $\dot{\theta}_{1:m}$ is connected to the neurons of the same bundle through non-plastic inhibitory connections:

$$\tilde{\varepsilon}_{kj} = \exp\left(\frac{-(k-j)^2}{(\sigma_n N_l)^2}\right) - 1 \tag{3}$$

where $\tilde{\varepsilon}_{kj}$ denotes the strength of the *non-plastic* connection between neurons k and j, σ_n is the standard deviation, and N_l is the number of neurons in the bundle. Therefore, the further the neuron is, the stronger the inhibition activity is. This approximates the behavior of a winner-take-all effect (WTA), but with the change in the inhibitory value depending on the proximity to the winner neuron (ensuring a continuous and more robust output).

2.2 Training Phase and Control Phase

For the proposed network to form the desired differential map, the information needs to be input/encoded into the network and extracted/decoded in a proper way. The input to the sensory layers (during training and control phases) and motor layers (during the training phase only) is calculated for each neuron based on its preferred (central) value ψ_c. Thus, this Gaussian distribution models the network's tuning curve. The firing rate for a certain input can be formulated as:

$$\alpha_i(t) = \exp\left(\frac{-\|\psi - \psi_c\|^2}{2\sigma^2}\right) \tag{4}$$

where ψ is the input value, and σ is calculated based on the number of neurons per layer N_l, and the range of change of the variable to be encoded from Ψ_{min} to Ψ_{max}. This leads to the contribution of the whole layer to encode a particular value (a process that can be interpreted as "population coding" [38]).

To get the estimated output from the network, a proper decoding function has to be chosen. Among the various decoding methods, the central neuron voting scheme is selected to calculate the decoded value corresponding to the firing rate in all the neurons of a layer. This can be modeled, for a specific time window, as follows:

$$\psi_{est} = \frac{\Sigma \psi_i . \alpha_i}{\Sigma \alpha_i} \tag{5}$$

where ψ_i is the central value of neuron i in the bundle l^{Ψ}, and ψ_{est} is the estimated (decoded) value of the output [38].

Fig. 1 depicts a schematic diagram of the proposed SNN-based method. Firstly, the motor babbling process initiates the training phases by providing the motor commands (joint velocity command (ω_d) for the robot to move linearly in the joint space through numerous random targets, as it is formulated in:

$$\omega_d = \kappa_\theta \frac{e_\theta}{\|e_\theta\|} \tag{6}$$

where $e_\theta = \theta_d - \theta$ is the error between the randomly generated desired joint angles (θ_d) and the current joint angles (θ).

This joint velocity command is scaled by the gain $\kappa_\theta > 0$, which is also varied randomly, within a certain range, during the babbling, to generate richer training data. During babbling motions, sensory information and motor commands are fed into the network to guide the modulation of the plastic synapses between the input and output layers through STDP. Sensory information is introduced from proprioception (θ and $\dot{\theta}$) and the external sensor (the visual velocity $\dot{x}$), which are then encoded. The difference in time of spikes generated in sensory and motor neurons $\Delta t = t_{post} - t_{pre}$ controls the amount of change in the synapses' strength. After a sufficient number of iterations (decided depending on the size of work space, learning rate and desired precision as

explained is subsection 2.5 , the robot is ready to perform a sensor-guided motion task.

$$v_d = \kappa_x \frac{e_x}{\|e_x\|} \tag{7}$$

where $e_x = x_d - x$ denotes the feedback spatial error and $\kappa_x > 0$ a variable gain that regulates the velocity by which the robot is driven towards the target x_d from the current position x. Finally, the motor command $\dot{\theta}_{cmd}$ is decoded from the spikes generated at the output bundles $l^{\dot{\theta}_{1:m}}$ as follows:

$$\dot{\theta}_{cmd} = \psi_{est} \tag{8}$$

This value is then fed into the robot servo controller to guide its motion towards x_d.

2.3 Neuron Model

Among the models available for spiking neurons is the Hodgkin and Huxley model that explains the mechanism of triggering an action potential and its propagation [39]. The mathematical model is composed of a set of nonlinear differential equations, which makes it computationally intense. However, this model is the most biologically plausible among all available models.

Another model that is widely used is the Leaky Integrate and Fire model (LIF). In this model, the behavior of a neuron is approximated as a simple RC circuit with a low-pass filter and a switch with a thresholding effect [30]. The RC circuit is first charged by an input current, which makes the voltage at the capacitor to increase until it reaches the threshold value. The switch opens and lets a pulse (spike) to be generated; the capacitor then starts to build up again. This model is simple and has a low computational cost, however, compared to Hodgkin-Huxley's model, it is not very biologically plausible.

In 2003, Izhikevich developed a model that can reproduce the various firing patterns recorded by different neurons in different brain regions [29]. This model is chosen for our study as it holds a balance between a reasonable computational cost while preserving the biological plausibility. The model can be described by the following set of differential equations:

$$\dot{v} = f(v,u) = 0.04v^2 + 5v + 140 - u + I \tag{9}$$

$$\dot{u} = g(v,u) = a(bv - u) \tag{10}$$

After a spike occurs, the membrane potential is reset as:

$$\text{if } v \geq 30 \text{ mV}, \quad \text{then } v \leftarrow c,\ u \leftarrow (u+d) \tag{11}$$

where v is the membrane potential and u is the membrane recovery variable as shown in Fig.4b. The parameter a determines the time constant for recovery, b determines the sensitivity to fluctuations below the threshold value, c gives the value of the membrane potential after a spike is triggered, and d gives the value of the recovery variable after a spike is triggered. The term I represents the summation of external currents introduced.

2.4 Synaptic Connections

Synapses are the connections that transmit signals between two neurons. Let us denote by ε_{ij} the weight/strength of the synapse. The transmission acts only in one direction, such that signals are carried from the ith presynaptic to the jth postsynaptic neuron. The information transmitted through synapses is usually encoded in the form of spikes (or action potentials). Different theories have been presented for the way of encoding and decoding such information in our brains [38]. The synaptic connections can either be excitatory or inhibitory, and plastic or non-plastic. The excitatory synapses are the connections that are more likely to increase the activity of postsynaptic neurons with the increase in the activity of the presynaptic neuron, while the

inhibitory synapses decrease that likelihood. The synapses connecting the input layer to the output layer are plastic (which means that its strength is subject to change). Let us denote by $\Delta\varepsilon_{ij}[t]$ the synapse's change of strength at the time instance t, which satisfies the following discrete update rule:

$$\varepsilon_{ij}[t+1] = \varepsilon_{ij}[t] + \Delta\varepsilon_{ij}[t] \tag{12}$$

One of the very first learning rules to update the synaptic weights is the Hebbian Learning rule [40], whose basic formulation is:

$$\Delta\varepsilon_{ij} = \eta a_i a_j \tag{13}$$

where the scalar η is the learning rate, a_i and a_j are the activities (or average firing rates) of the pre and postsynaptic neurons, respectively. This rule strengthens the connections between strongly correlated variables, and has been shown to perform principal component analysis (PCA) [41]. However, this type of learning does not take into consideration the *time difference* between spikes, which is the main feature of SNNs.

Another learning rule that is more appealing from a biological perspective is STDP, where potentiation (increase) or depression (decrease) in the strength of the connections is dependent upon the relative timing of spikes that occur in presynaptic and postsynaptic neurons [42]. STDP is considered to be one temporal form of Hebbian learning [43], e.g. in [44], it is shown that it can perform 'kernel spectral component analysis' (kSCA), which resembles PCA. This attribute makes it suitable for mapping two spaces while successfully updating the synaptic weights.

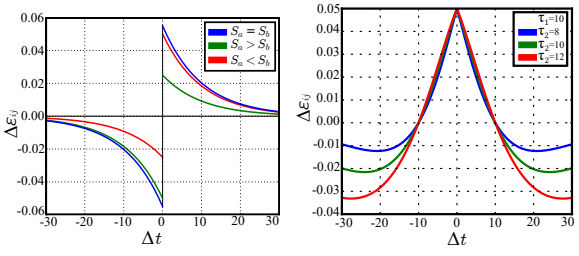
In the literature [45], two common patterns for STDP is either as symmetric [46] or antisymmetric [47], as depicted in Fig. 3. The antisymmetric model can be formulated as:

$$\Delta\varepsilon_{ij} = \begin{cases} -S_a \exp\left(-\Delta t/\tau_a\right) & \Delta t \leq 0 \\ \\ S_b \exp\left(-\Delta t/\tau_b\right) & \Delta t > 0 \end{cases} \tag{14}$$

where S_a and S_b are coefficients that control the magnitude of the synaptic depression and potentiation, respectively, where τ_a and τ_b determine the time window through which depression and potentiation occur. The asymmetric STDP is thus suitable for a learning process whenever the sequence of signals matters, e.g. if a spike arrives from the presynaptic neuron before the spike from a postsynaptic neuron (which results in a positive value for Δt, and thus the synapse's weight is potentiated).

The symmetric learning model is the most appropriate in this case given the continuous firing at both input and output layers. Furthermore, a symmetric STDP rule (with different reward modulated versions) was reported to be observed in hippocampus and prefrontal cortex in several studies [48–50] and was studied in [51]. The symmetric STDP rule can be described by:

$$\Delta\varepsilon_{ij} = S\left(1 - \left(\frac{\Delta t}{\tau_1}\right)^2\right)\exp\left(\frac{|\Delta t|}{\tau_2}\right) \tag{15}$$



(a) (b)

Figure 3: (a) Asymmetric and (b) Symmetric learning STDP rules. For asymmetric learning, S_a and S_b are plotted at different values while keeping τ_a and τ_b constant. For the symmetric one, τ_2 is plotted at different values while keeping τ_1 constant.

where S is a coefficient that controls the magnitude of the synaptic change, the ratio between τ_1 and τ_2 decides the time window through which potentiation and depression occurs, and Δt is the difference between the timing of spikes at post t_{post} and pre t_{pre} synaptic neurons. In this study, we use $S = 0.05$, $\tau_1 = 20$ms and $\tau_2 = 18$ms. In this learning model, the change in synapse's weight is controlled by the absolute value of Δt, but not the sign, i.e. the sequence of firing. The chosen time window for the pre and postsynaptic neurons for STDP is 30ms, which means as long as $-30 \leq \Delta t \leq 30$, it will still contribute to the modification of the synaptic strength.

With the absence of the hidden layer(s), nonlinearity in the neuronal units, as well as the features of STDP, make it possible to learn the differential map. Such approach is supported by previous studies (see [24] and [18]), as illustrated in section 4.

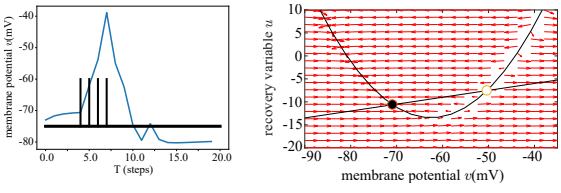
Additionally, the inter-inhibitory connections establish the WTA effect along with the STDP and the slight contribution from adjacent neurons (approximating the effect of lateral excitatory connections) develop an approximation of a Hidden Markov Model (HMM) online learning as described in [25]. For a Markov process, the current state depends only on the current state and independent of the past states [52]. This property is useful in our case as it guarantees the continuity of the output and avoiding jerky motions, but still capable of changing the output based on the current state in case of inaccurate estimations. Moreover, the HMM developed in that case proven to perform *forward sampling* [26] in the WTA circuit, and achieve an online stochastic approximation to *expectation-maximization (EM)* parameter learning thanks to the STDP. So, from a probabilistic perspective, connections from the input layer provide the *observations* for the HMM, while the WTA circuit provides an approximation for the *E-step*, based on forward sampling. To study the proposed architecture and compare it to that in [25], it needs an extensive separate study to fully analyze the DMSNN from a statistical perspective.

To perform the vision-guided motion task with the proposed network, a careful setting of its parameters is needed, as explained in the next section.

2.5 Tuning the Network Parameters

Tuning the various parameters for the neurons and synapses is a difficult task. Some basic rules can be used to guide the trial and error approach to choose a set of appropriate values. For example, as u is the membrane recovery variable, it is responsible for the delivery of negative feedback to v, such that it resists change of the value of v. The network's parameters a, b, c and d tune how v and u change and interact together over time.

Fig. 4b shows the phase portrait of the output neurons, as it plots the recovery potential versus the membrane potential. The two black curves represent the *nullclines* for both v and u, i.e. the line along which partial derivative equals zero which separates the planes of variation of v and u. To obtain these nullclines and analyze the system, equations 9 and 10 are set equal to zero to obtain lines along which there



(a) (b)

Figure 4: (a) Computational and (b) analytical analysis of a spiking neuron. In (a) an integrator neuron receives spikes from four neurons to cross a threshold value and trigger a spike, While in (b) a phase portrait of a spiking neuron is plotted.

is no change in v and u, respectively. The intersection of these nullclines forms attractor (stable) points or repeller (unstable) points [53]. From the location of attractors and repellers, the stability regions can be concluded. The equilibrium points (v^*, u^*) are obtained by solving the equations of the parabola and line obtained from equating $f(v,u)$ and $g(v,u)$ together. The stability of these points is determined by the eigenvalues of the Linearization matrix L:

$$L(v^*,u^*) = \begin{bmatrix} \frac{\partial f}{\partial v}(v^*,u^*) & \frac{\partial f}{\partial u}(v^*,u^*) \\ \frac{\partial g}{\partial v}(v^*,u^*) & \frac{\partial g}{\partial u}(v^*,u^*) \end{bmatrix} = \begin{bmatrix} 0.08v^*+5 & -1 \\ ab & -a \end{bmatrix} \tag{16}$$

As the v-nullcline shifts upward, the attractor and repeller annihilate each other and merge into a saddle; any further upward shift leads to the disappearance of the saddle point.

As shown in Algorithm 1, to use an SNN to model a certain system, initial values should be assigned for neuron parameters a, b, c, d based on their role within the neuronal layer. In this work, the motor neurons are modeled as *integrator neurons* as it acts as a coincidence detector, such that it triggers a spike by accumulating closely timed signals as illustrated in Fig. 4a. The initial values for the neuron parameters are $a = 0.02$, $b = -0.1$, $c = -55$, and $d = 6$. The sensory neurons are modeled as *fast spiking* neurons providing high frequency spikes and initialized with $a = 0.1$, $b = 0.2$, $c = -65$, and $d = 2$. The above-mentioned initial values are suggested in [54].

The parameters' variation produces specific properties in the system, which can be used as a guideline to adjust the neuron's firing pattern: (i) The value of a controls the decay rate of u. This can be clearly noticed when comparing the firing behavior of low-threshold spiking (LTS) neurons with resonator (RZ) neurons. RZ neurons have a higher value of a (to set sub-threshold oscillations and resonate at a narrow band of frequencies) than LTS neurons, but same values of b,c and d. (ii) The parameter b controls the sensitivity of u to changes in the membrane potential v below the threshold value. The integrator neuron has a low value of b which is why many spikes with small time interval in between are needed to trigger a spike. (iii) The parameter c describes the value to which v is reset after firing. Therefore, decreasing its value enables to create spikes bursts since it makes the recovery to the original threshold value faster. (iv) The parameter d describes the reset in the value of u after a

Algorithm 1 Network Parameters Tuning

Input

n,m = Number of dimensions of spaces to be encoded

N_l = Number of neurons to encode each dimension

Itr^* = Number of iterations to build the map

Output

a, b, c, d = Neurons parameters

I^* = Minimum input to have continuous spikes

C_I, C_E = Maximum strength of Inhibitory and Excitatory connections, respectively

S = Learning rate for STDP based synaptic connections

Routine

1: Assign initial values for neuron parameters a, b, c, d
2: **while** $\xi \geq e_{th}$ **do**
3: Build the neuron model $f(v,u)$ and $g(v,u)$
4: Solve for nullclines at $f(v,u) = 0$ and $g(v,u) = 0$
5: Get v^* and u^* where $f(v^*,u^*) = g(v^*,u^*) = 0$
6: Evaluate $L(v^*,u^*)$ to check stability
7: Obtain I^* and hence C_E
8: Set S for given Itr^*
9: Update values for a, b, c, d
10: **end while**

spike occurs, thus, lowering it creates a higher firing frequency for the same input current.

After setting the neuron parameters, the value for I^* and C_E can be concluded by analyzing the stability at v^* and u^* such that a specific firing rate at the output neurons is obtained for a certain input from multiple input neurons. I^* is even more critical to estimate in our study, as selective disinhibition has to be achieved and the neuron has to be maintained at the verge of firing to avoid excessive firing [55]. This ensures that only the correct motor neurons fire upon excitation of the corresponding sensory neurons.

To obtain I^*, we first equate $g(v^*, u^*) = a(bv^* - u^*) = 0$, and solve for $u^* = bv^*$. Then, substitute u^* into $f(v^*, u^*) = 0$ such that $0.04v^{*2} + (5-b)v^* + 140 + I = 0$. The intersection of the u and v nullclines at one point is when the neuron starts to give continuous spikes, which means there is only one solution for the quadratic equation. The equilibrium points are given by $v^* = -(5-b)/(2 \times 0.04)$ and $u^* = bv^*$. The value of I^* is finally computed by solving $f(v^*, u^*) = 0$. The parameter C_E must be chosen such that at the end of the training phase the selected firing behavior is still maintained. After setting C_E, the spiking behavior is tested. The chosen values for the motor neurons must not allow evoking spikes at low spiking frequency from sensory neurons. Note that increasing the frequency makes the learning process more susceptible to noise. Therefore, the motor neuron parameters are to be modified instead. The parameter b is then incremented slightly until a satisfactory performance is obtained.

Depending on the size of the data set and the number of neurons in each bundle, the number of iterations Itr^* must be defined to build the map, which in turn determines the value of the learning rate gain S. If Itr^* is relatively small, it will lead to a large S, which often leads to instability and noise sensitivity. A large Itr^* results in an exhaustive (computationally demanding) training process.

To quantify the accuracy of the computed differential map approximated by the network, we define the following metric for the accuracy of estimations as introduced in Algorithm 1:

$$\xi = \frac{1}{N} \sum_{n=1}^{N} \sqrt{(\vec{v}_d - \vec{v}_{est})^2} \tag{17}$$

which is simply the difference between the desired spatial velocity $\vec{v}_d$ and the spatial velocity $\vec{v}_{est}$ obtained upon execution of the estimated motor command $\dot{\theta}_{cmd}$ as shown in Fig. 5 (calculated from the decoding equation 8); the scalar $N > 0$ is the number of points in the workspace over which the difference is measured to obtain a mean error value. The tuning process continues until the value of ξ is below some threshold value e_{th}, indicating that the network's performance is acceptable. The defined metric can be used for future improvements of the network by searching for the optimal network parameters using evolutionary algorithms [56], or building a confidence map to guide the babbling process [57].

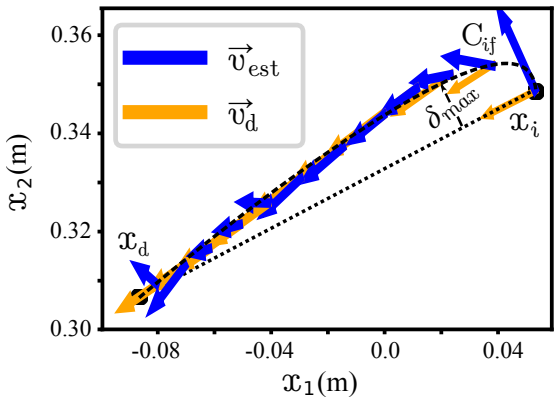


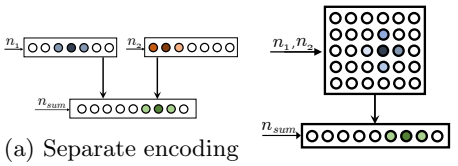
Figure 5: A planar robot driven by the DMSNN, initially at x_i reaching to a target at x_d and at each time step both the actual velocity ($\vec{v}_{est}$) and the desired velocity vector ($\vec{v}_d$) are calculated.

3 SIMULATION RESULTS

3.1 Network Layout Selection

To test the proposed network and tuning method, the summation of two variables ($n_1 + n_2 = n_{sum}$) is carried out in different approaches. In Fig. 6a, the sum is estimated using two 1D-layers that encode the two variables ("n_1" and "n_2") connected through all plastic connections to a 1D-layer encoding the result ("n_{sum}"). Random numbers are generated within a range, and introduced to the input layers, with their summation introduced into the output layer. The network parameters are tuned as described in Sec. 2.5, such that during the training phase firings in both input and output layers allow the plastic connections to represent the desired summation function. After training, the sum is estimated from the decoded values at the output layer. The approach gives a mean error of 7.5%.

Fig. 6b shows another network layout used to test our method. A 2D input layer is used instead of two separate 1D layers. The difference from the previous layout is that the neurons' activity is estimated by multiplying the normally distributed activity of both variables. The mean error for this layout is around 2%. To build the proposed DMSNN, we use the former approach. The rationale behind this choice is elaborated in Sec. 5.



(a) Separate encoding (b) Joint encoding

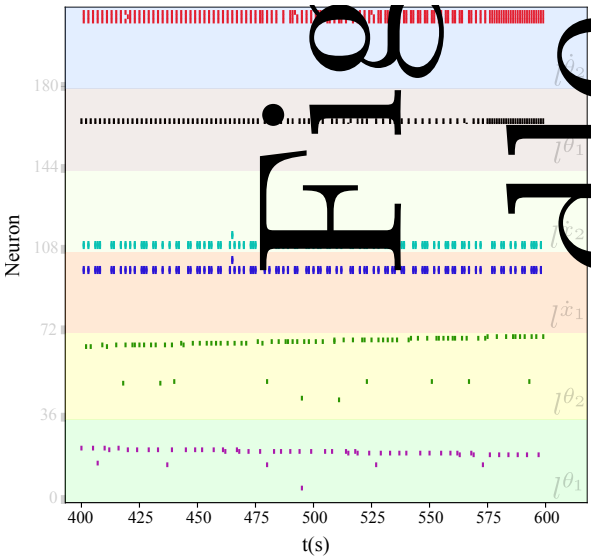
Figure 6: A schematic of the two possible ways to perform summation of two variables (n_1 and n_2). In this paper, (a) is the chosen approach.

3.2 2DOF Planar Robot

To verify the effectiveness of our method for visual servoing tasks, a simulation model of a 2DOF planar robot with revolute joints is built. The differential kinematics of this system are:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}-l_1\sin(\theta_1) - l_2\sin(\theta_{12}) & -l_2\sin(\theta_{12})\\ l_1\cos(\theta_1) + l_2\cos(\theta_{12}) & l_2\cos(\theta_{12})\end{bmatrix}\begin{bmatrix}\dot{\theta}_1\\ \dot{\theta}_2\end{bmatrix} \tag{18}$$

where l_i denotes the link length, and $\theta_{12} = \theta_1 + \theta_2$. With the above definition of the Jacobian matrix, we can derive the inverse differential mapping (based on equation 2), which is then used to generate training data for the SNN. During the training process, normalized spatial velocity vectors are fed at random joint angles to the corresponding bundles in the input layer along with the corresponding angular velocities to the bundles of the output layer. The firing activity of the neurons is monitored, as shown in Fig. 7, to make sure the network parameters are correctly set to obtain the desired firing frequencies and patterns of the neurons. For the accuracy of the robot sim-



ulation, around the singular configurations, a small value (10^{-5}) is added to the determinant of the Jacobian matrix to avoid the unbounded values of joint velocities.

After training, the network is tested by providing a random target in Cartesian workspace x_d for a random initial position x_i. By generating a normalized spatial velocity vector $\vec{v}_d$ from the current position x to reach the target, the estimated angular velocities ($\dot{\theta}_{cmd}$) at each joint can be decoded from the bundles of the output layer.

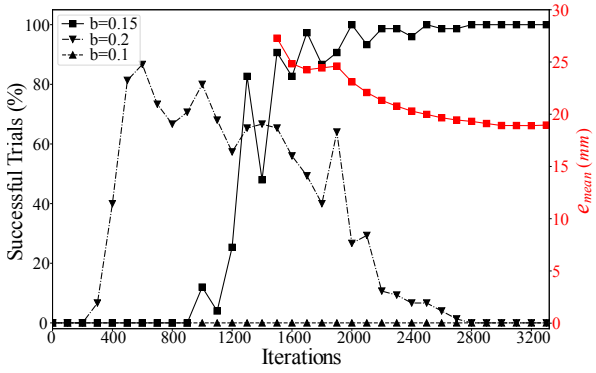
To quantify the performance, we calculate the minimum distance δ from the current position to the line segment $\overline{x_i x_d}$ as shown in Fig. 5. The target path is a straight line, thus, the shortest distance at each point is defined by the normal to $\overline{x_i x_d}$. But in other cases, a more complex path may be required. For a robot moving along the path C, divided into N_C points, given a reference target path Ω, composed of N_Ω points, the mean error e_{mean} is calculated by averaging the maximum deviation δ_{max} over N_{trials} by applying Algorithm 2. The training is done over approximately 3000 iterations and tested over 5 times repetition of servoing to 15 different targets (generated randomly in the workspace), to obtain the number of successful trials (in which the final end-effector error e_x is below the threshold value, that is chosen to be 1mm here) and the standard deviation of the maximum deviation δ_{max} for approaching chosen targets is given by:

$$\sigma_{servoing} = \sqrt{\frac{\sum_{n=1}^{N_{trials}} (\delta_{max}(n) - e_{mean})^2}{N_{trials} - 1}} \tag{19}$$

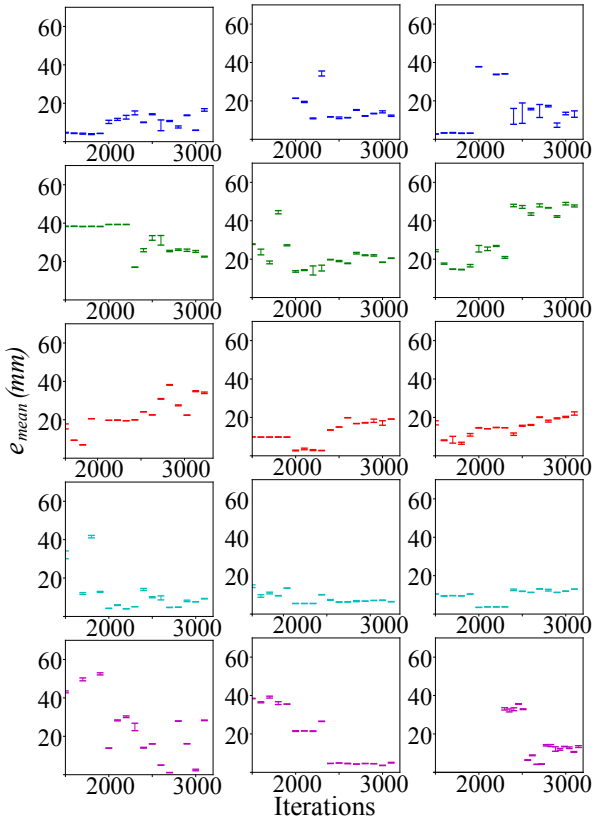
Fig. 8a shows the plot of the percentage of successful trials for the chosen parameters against the number of training iterations. It can be concluded that the learning in that case reaches a stable state after around 2700 iterations. Also, the figure includes the plot of the network performance with a slight change of the key parameters (in this case parameter b for the output neurons as discussed earlier) with a noticeable degradation in the learning ability and stability of the network. Thus, it can be concluded the importance of fine tuning of the network parameters. Additionally, Fig. 8b shows the development of the performance of servoing to different points in the workspace. It shall be noted that the performance of the network is degraded in some areas while the training proceeds, thanks to the generalization of the learning process. Thus, instead of training at a local area, the whole defined workspace is instead learnt. This means that initially specific examples are learnt and all the neurons act to represent these examples, but as the training proceeds the representation includes a wider set of examples.

Algorithm 2 Calculating Mean Error

$\delta_{max} = 0$
for $i = 1$ to N_{trials} **do**
  $e_j = 0$
  for $j = 1$ to N_C **do**
    $e_k = \|C(j) - \Omega(1)\|$
    for $k = 2$ to N_Ω **do**
      $e_{ij} = \|C(j) - \Omega(k)\|$
      if $e_{ij} \leq e_k$ **then**
        $e_k = e_{ij}$
      end if
    end for
    if $e_k \geq e_j$ **then**
      $e_j = e_k$
    end if
  end for
  Stack error $\delta_{max}(i) \leftarrow e_j$
end for
$e_{mean} = \frac{1}{N_{trials}} \sum_{n=1}^{N_{trials}} \delta_{max}(n)$



(a)



(b)

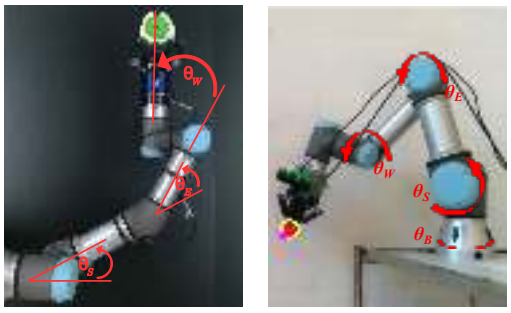
Figure 8: (a) Plot of both the number of successful trials (in black), and the mean error (in red) during the control phase versus the number of training iterations. The solid, dotted, and dashed black lines are related to the output neurons with the parameter $b = 0.15, 0.1$ and 0.2, respectively. The mean error is only plotted for the case in which successful trials are above 80%. (b) Plot of the mean error and standard deviation for each individual target approach (unsuccessful trials are discarded from calculations and plot).

4 EXPERIMENTS

4.1 Experimental Setup

The proposed SNN is simulated using NeMo library, a tool developed to simulate SNN [58, 59]. A UR3 robot is set with an Intel RealSense D415 camera in setup as shown in Fig. 9a. The D415 is an RGB-D camera providing real-time color frames and depth maps. Image registration is carried out by aligning the color and depth frames, then, the end-effector position can be measured through color filtering of the manipulated object. For effective color filtering, the image is blurred to remove high-frequency noise, then the color frame is converted to HSV color space for more robust performance independent of the light intensity.

A mask, in the desired color range, is applied to obtain a binary image, which is then subjected to some morphology operators to remove the small blobs. These operations allow to filter the image and make it easier to discriminate the colored end-effector as the biggest contour. Afterwards, the moment of this contour is calculated such that $M_{ij} = \sigma\phi(x_1, x_2){x_1}^{(i)}{x_2}^{(j)}$. Then, the centroid (ρ_x, ρ_y) is calculated in pixel units based on the moment such that $\rho_x = M_{10}/M_{00}$ and



(a) UR3 Planar configuration (b) UR3 Spatial configuration

Figure 9: The UR3 moves (a) planar motion while controlling 3DOFs and (b) spatial motion while controlling 4DOFs.

$\rho_y = M_{01}/M_{00}$, see [60].

The location at the center of the object is then converted to world coordinates by using the camera's intrinsic parameters [61]:

$$x_1 = (\rho_x - c_x)\frac{x_3}{F}, \qquad x_2 = (\rho_y - c_y)\frac{x_3}{F} \quad (20)$$

where ρ_{xy} and ρ_y denote to the end-effector position in pixels, x_3 is the end-effector's depth, c_x and c_y denote the principal point and F is the focal length. The following first-order filter is used to remove noise from these visual measurements:

$$s_f(t+1) = s_f(t) - \lambda(s_f(t) - s(t+1)) \quad (21)$$

where s and s_f are the variable before and after filtering, respectively, and $\lambda > 0$ denotes the filter's gain.

To train the network, the robot is driven linearly in joint space between randomly generated angles. Data is collected at a constant sampling rate of 25Hz, then filtered. Fig. 10c shows the position obtained from the camera before and after filtering. Each neuron bundle is fed with the corresponding data to let the plastic connections develop to form the required differential mapping. Once the training phase ends (dependent upon the number of neurons and size of workspace), the strength of the synaptic connections is kept constant. Random targets are given across the robot workspace as shown in Fig. 10a and Fig. 10b. The current joint angles and the desired spatial velocity are updated every 20ms.

4.2 3DOF Planar Robot

For this case, three joints are controlled to guide the UR3 robot in planar motion, as shown in Fig. 11. The network consists of five input bundles ($l^{\theta_{1:3}}$ and $l^{\dot{x}_{1:2}}$) and three output bundles ($l^{\dot{\theta}_{1:3}}$). The network parameters, as well as the number of neurons in each bundle, are shown in Table 1. These three joints

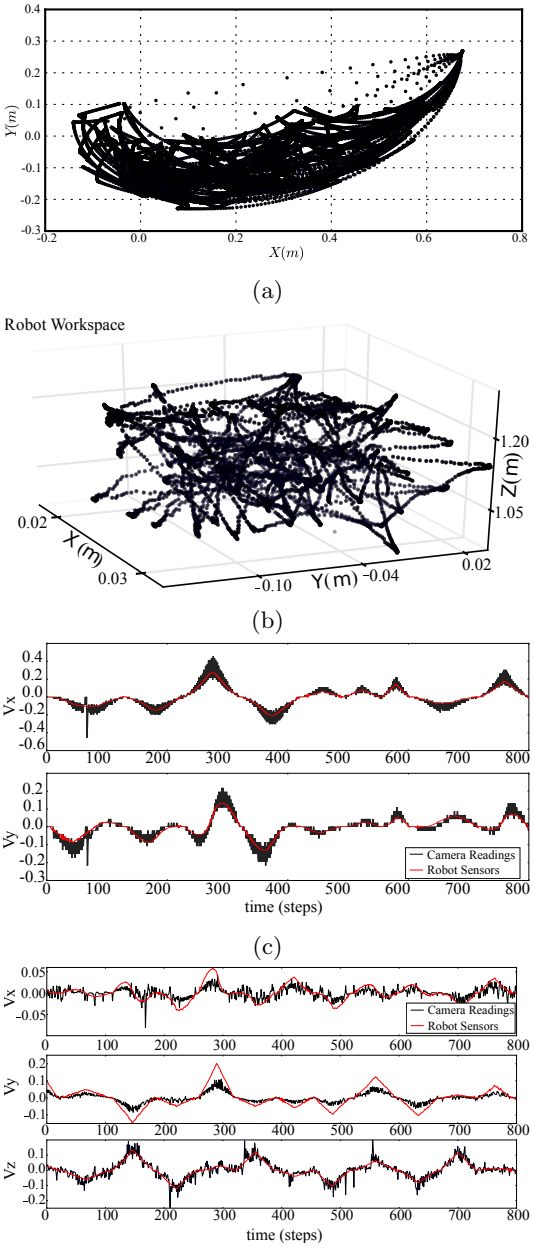


Figure 10: Plot of data collected for training the DMSNN for (a) 3DOFs and (b) 4DOFs, and velocities as measured from the camera (after filtering) against that estimated by the internal sensors of the robot for (c) 3DOFs and (d) 4DOFs.

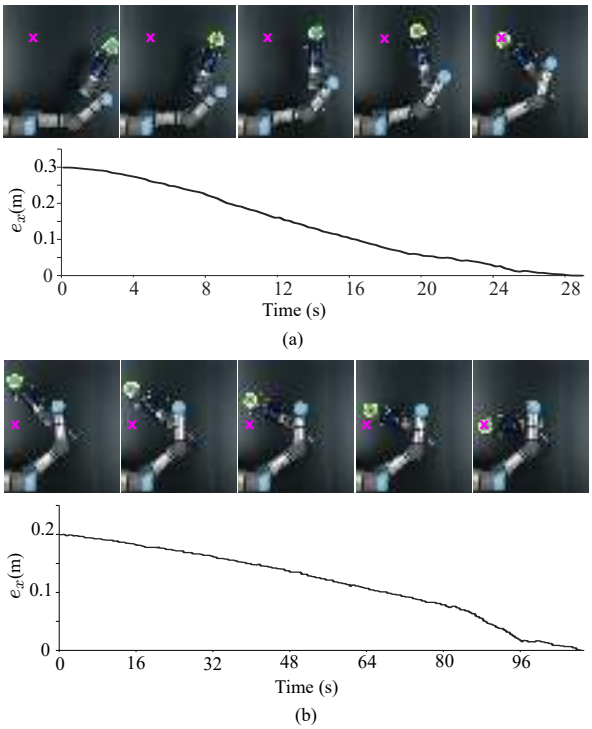


Figure 11: The UR3 while performing planar motions in (a) and (b) with the norm of the error e_x plotted during motion.

have ranges of: $[-180°, -90°], [-45°, 0°]$ and $[90°, 180°]$.

From the motor babbling data collected, the maximum and minimum values for both the spatial and angular values are obtained. For each layer with $\mathcal{N}_{3D}$ neurons, the central value ψ_c encoded by each neuron is assigned by dividing the range of each variable evenly over the whole layer. The update of the synaptic strength is shown as heatmaps in Fig. 12a and Fig. 12b. Each heatmap depicts the relation between one bundle from the sensory layer to one bundle from the motor layer, where each pixel gives the strength of an excitatory synapse connecting one sensory neuron to one motor neuron. After running the simulation for 6000 iterations, the strength of synapses between a motor layer to the sensory layer is modulated to represent the required differential map. The robot is then given random points to reach by its end-effector through a visual servoing process, and considered successful at reaching if the end-effector is less than 3 pixel away from the target in each coordinate. As shown in Fig. 11, a target point is provided and the robot automatically moves towards it by using decoded motion command $\dot{\theta}_{cmd}$. The distance between the end-effector and the target is represented as the norm $\|e_x\| = \|x - x_d\|$.

4.3 4-DOF Spatial Robot

In this case, four joints are controlled to guide the UR3 robot as shown in Fig. 9b.

The network consists of seven input groups ($l^{\theta_{1:4}}$ and $l^{\dot{x}_{1:3}}$) and four output groups ($l^{\dot{\theta}_{1:4}}$). The network parameters and the number of neurons in each layer are shown in table 1. The four joints ranges are: $[-200°, -170°], [-75°, -45°], [90°, 110°]$ and $[-200°, -160°]$. The heatmaps depicting the update of weights at the 4000 and 9000 iterations are shown in Fig. 12c and Fig. 12d, respectively. Similar to the previous setup, several end-effector targets (randomly given across the work space) are given to the robot in

Table 1: Network Parameters

Neuronal layers					
Param / Layer	a	b	c	d	$\mathcal{N}_{3D/4D}$
l^{θ_i}	0.1	0.2	-65	2	68/136
$l^{\dot{x}_i}$	0.1	0.2	-65	2	68/136
$l^{\dot{\theta}_i}$	0.02	0.15	-55	6	68/136
Synaptic connections					
Param / DOF	S	τ_1/τ_2	C_I/C_E	Itr^*	
3 (Planar)	0.05	20/18	-4/4	6000	
4 (Spatial)	0.03	20/18	-5/5	9000	

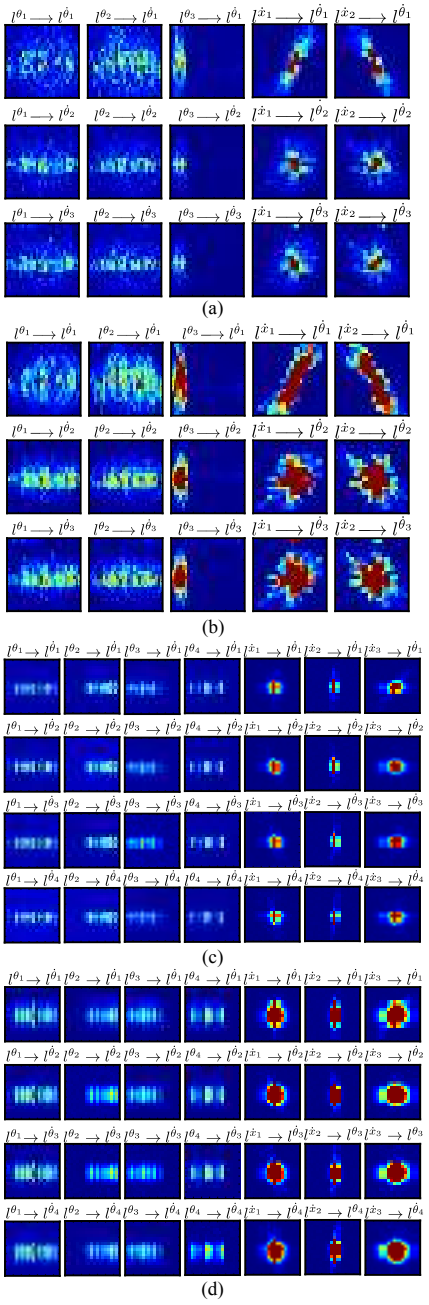


Figure 12: Heatmap of the weights update process for the 3DOF case of study at (a) 3000 and (b) 6000 iterations, and for 4DOF case at (c) 4000 and (d) 9000 iterations. The dark blue color corresponds to zero weight while dark red color corresponds to the maximum weight.

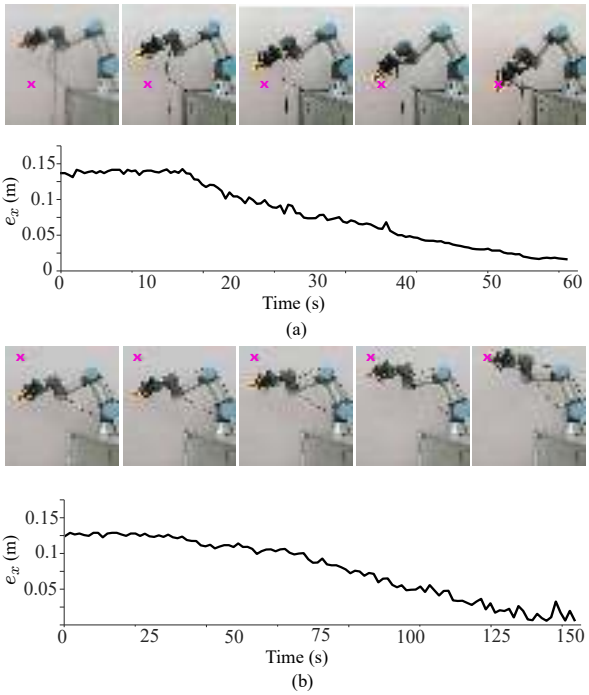


Figure 13: The UR3 while performing spatial motions in (a) and (b) with the norm of the error in the 3 axes e_x plotted during motion.

sensor space. The target reaching in this case is considered successful when the end-effector is less than 3 pixel away from the target in x and y coordinates and 5mm in depth (z coordinate). Representative results for these 4-DOF visual servoing tasks are depicted in Fig. 13, and it shall be noted that the ripples in the plot are due to the noise in the data collected by the depth sensor, which can be judged by comparison to the plots of velocities based on the readings collected by the motor encoders as shown in Fig. 10d. The accompanying multimedia video demonstrates the performance of our new method with many experimental results.

5 DISCUSSION

As shown in section 3.1, the accuracy of the summation of two numbers by the network in the case of multiple one-dimensional arrays is lower than that obtained in the case of a multi-dimensional array. However, this result comes on the expense of the network size where the 1D layers require only $2\mathcal{N}$ neurons, for $\mathcal{N}$ is the number of neurons per layer, while the 2D layer needs $\mathcal{N}^2$ neurons. This means that for an O-dimensional case, $\mathcal{N}^O$ neurons would be needed. However, note that the explicit use of visual feedback in our formulation provides a valuable *rectifying* property to the network (i.e. it corrects for small errors in an otherwise open-loop controller) while demanding moderate computational power.

Moreover, in this study the SNN is used to guide the robot for visual servoing process, where it requires fast state updates, hence the proposed network is to be used instead as it compromises between real-time operation and sufficient accuracy. However, in case of accurate mapping of robot states and commands the second approach is to be used, or additional hidden layers are to be added to the network for higher accuracy and better estimates. Another feature, which is not evident from the summation test studied, is that SNN better approximates real-life operations where the values to be encoded/decoded are continuous, i.e. with no discontinuity or jumps. For the neurons in the network to evoke a spike, it needs first to let the membrane potential build up over a series of time steps. While incorporating the Gaussian activity, this leads to sharing excitation with neurons in the neighborhood, and it becomes more likely for the adjacent neurons to be triggered in next cycles as well. As the adjacent neurons encode values close to each other, this achieves a continuous smooth operation and avoid sudden changes and jerky motions. So, despite the fact that in the asynchronous SNN, as in our case, neurons that release a spike first tend to spike again, the fine tuning of the parameters acts to facilitate more accurate approximations and the neighboring neurons activity acts to correct the estimations during the servoing process.

This work also exploits the potential of an SNN based on STDP, following some procedures for fine-tuning of the network parameters as described in subsection 2.5, where only two layers (input and output) and a low number of neurons are needed to develop a differential map for multi-DOF robot. Additionally, the noisy data is well handled thanks to the inhibition that occurs for distal firing of pre and post-synaptic neurons, which weakens the synapses connecting uncorrelated neurons, aside to the inter-inhibitory connections that inhibits distal neurons to avoid undesired firing.

The developed network can then be compared to the work from [23], where a SNN is used to build a map for static transformation of coordinates for a 2DOF robot (let us denote it as *CTSNN*), and the work from [62], where an SNN modeling the sensorimotor cortex is used to guide a 2DOF robot in a reaching task (and denoted as *SMCSNN*). The comparison is concluded in Table 2 to show the time and number of iterations needed for training the networks, the network size, and the final error e_x.

It can be concluded that both the simulation time and the real time required for training and running the network is less for the proposed network, thanks to the fine tuning of the network parameters and maximizing the benefit from the spiking nature of the network.

Following the proposed tuning method, each neuron bundle consists of around 136 neurons, in case of the spatial configuration, compared to around 1000 neurons in [24]. In [18] each bundle is constructed of around 100 neurons only, however, a hidden layer is needed which

increases the size of the network and complexity of the tuning process. Moreover, the robot in this work needed to approach only 100 points, and data recorded while moving to these target points was sufficient for the training process compared to 6426 target points in [21]. The robot succeeded in all trials to reach the destined targets through visual servoing.

The time required to reach the target varies depending on the distance from the current position and the richness of data trials available along the path between the current and target pose. This can be seen in the two examples in Fig. 11, where in Fig. 11a the robot takes less time to cover a bigger distance compared to Fig. 11b. This can be explained by having the training data collected from motor babbling in joint space with more examples of the robot moving in waving-like trajectories and less abundant data for linear motion in task space. Similarly, the trial in Fig. 13a less time is needed compared to the trial in Fig.13b as the target in the latter is close to the boundary of the studied section of the workspace which has less training examples. Richer data collected from motions in both joint space and task space can be used in a future study to role out the effect of varying the data on different motion types, and the possibility of emerging of motion primitives from such behavior.

It can be noticed, from Fig. 10c and Fig. 10d, the training data is filtered using only first-order filter and contains more noise in case of the spatial motion, due to the noise in depth readings, however, the DMSNN is still able to approximate the differential relationship and build the desired map. This can be justified by the depression in the strength of uncorrelated neurons due to the chosen STDP rule.

Table 2: Networks Comparison

Param / Net	Itr^*	$T^*(s)$	e_x	$\mathcal{N}_{net}$
CTSNN	≈4000	800	NA	2000
SMCSNN	200ep	3000	≈1mm	704
DMSNN	3000	30	≤1mm	216

* The number of iterations for *CTSNN* is estimated based on 8000s training divided into 1600 time steps (each time step 0.125ms) for each stimulus introduced. The number of training steps is not mentioned for the *SMCSNN*, instead the number of epochs is mentioned (which in this case is the motion from one starting point to a target point) and each epoch takes around 15s. Each iteration in the DMSNN is 80ms of simulation time (which is only 10ms of real time), through which a certain stimulus is introduced to the network.

6 CONCLUSIONS

Based on an intuitive architecture and tuning guideline, the proposed DMSNN provides a way to build a differential map to drive robots with many DOFs through multi-modal servoing tasks. The network is featured by real-time operation and a small amount of data, compared to similar methods in literature, is needed for training. The current limitations of this method are related to the limited resolution of the network output. Future work will focus on improving the output resolution, as well as deriving a mathematical formula to conclude the optimal parameters for neurons firing and STDP learning. We also plan to use this method for conducting vision-guided shape control tasks with deformable objects, as in [63], and apply a cerebellar controller to enhance the performance and reduce the deviation from the path, as in [20].

Funding

This research work is supported in part by the Research Grants Council (RGC) of Hong Kong under grant number 14203917, in part by PROCORE-France/Hong Kong Joint Research Scheme sponsored by the RGC and the Consulate General of France in Hong Kong under grant F-PolyU503/18, in part by the Chinese National Engineering Research Centre for Steel Construction (Hong Kong Branch) at PolyU under grant BBV8, in part by the Key-Area Research and Development Program of Guangdong Province 2020 under project 76 and in part by The Hong Kong Polytechnic University under grant G-YBYT and 4-ZZHJ.

References

[1] D. Navarro-Alarcon, A. Cherubini, and X. Li, "On model adaptation for sensorimotor control of robots," in *2019 Chinese Control Conf. (CCC)*. IEEE, 2019, pp. 2548–2552.

[2] J. Fagard, R. Esseily, L. Jacquey, K. O'Regan, and E. Somogyi, "Fetal origin of sensorimotor behavior," *Frontiers in Neurorobotics*, vol. 12, p. 23, 2018. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnbot.2018.00023

[3] T. Aoki, T. Nakamura, and T. Nagai, "Learning of motor control from motor babbling," *IFAC-PapersOnLine*, vol. 49, no. 19, pp. 154–158, 2016.

[4] X. Xiong, F. Wörgötter, and P. Manoonpong, "Adaptive and energy efficient walking in a hexapod robot under neuromechanical control and sensorimotor learning," *IEEE Trans. on cybernetics*, vol. 46, no. 11, pp. 2521–2534, 2015.

[5] L. Pinto and A. Gupta, "Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours," in *2016 IEEE Int. Conf. on Robot. and Automat. (ICRA)*. IEEE, 2016, pp. 3406–3413.

[6] H. A. Pierson and M. S. Gashler, "Deep learning in robotics: a review of recent research," *Advanced Robot.*, vol. 31, no. 16, pp. 821–835, 2017.

[7] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Mag.*, vol. 34, no. 6, pp. 26–38, 2017.

[8] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bull. of Math. Biophys.*, vol. 5, no. 4, pp. 115–133, 1943.

[9] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. USA: Prentice-Hall, Inc., 1985.

[10] W. Maass, "Networks of spiking neurons: the third generation of neural network models," *Neural networks*, vol. 10, no. 9, pp. 1659–1671, 1997.

[11] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *nature*, vol. 381, no. 6582, pp. 520–522, 1996.

[12] G. Wang and M. Pavel, "A spiking neuron representation of auditory signals," in *Proceedings. 2005 IEEE Int. Joint Conf. on Neural Networks, 2005.*, vol. 1. IEEE, 2005, pp. 416–421.

[13] J. L. Krichmar, "Neurorobotics—a thriving community and a promising pathway toward intelligent cognitive robots," *Frontiers in neurorobotics*, vol. 12, p. 42, 2018.

[14] N. R. Luque, J. A. Garrido, R. R. Carrillo, S. Tolu, and E. Ros, "Adaptive cerebellar spiking model embedded in the control loop: Context switching and robustness against noise," *International Journal of Neural Systems*, vol. 21, no. 05, pp. 385–401, 2011.

[15] I. B. Ojeda, S. Tolu, and H. H. Lund, "A scalable neuro-inspired robot controller integrating a machine learning algorithm and a spiking cerebellar-like network," in *Conference on Biomimetic and Biohybrid Systems*. Springer, 2017, pp. 375–386.

[16] S. Furber, "Large-scale neuromorphic computing systems," *Journal of neural engineering*, vol. 13, no. 5, p. 051001, 2016.

[17] G. Schöner, *Dynamic thinking: A primer on dynamic field theory*. Oxford University Press, 2016.

[18] J. C. V. Tieck, H. Donat, J. Kaiser, I. Peric, S. Ulbrich, A. Roennau, M. Zöllner, and R. Dillmann, "Towards grasping with spiking neural networks for anthropomorphic robot hands," in *Int. Conf. on Artif. Neural Networks*. Springer, 2017, pp. 43–51.

[19] Z. Bing, C. Meschede, F. Röhrbein, K. Huang, and A. C. Knoll, "A survey of robotics control based on learning-inspired spiking neural networks," *Frontiers in neurorobotics*, vol. 12, p. 35, 2018.

[20] C. Corchado, A. Antonietti, M. C. Capolei, C. Casellato, and S. Tolu, "Integration of paired spiking cerebellar models for voluntary movement adaptation in a closed-loop neuro-robotic experiment. a simulation study," in *2019 IEEE Int. Conf. on Cyborg and Bionic Syst.* IEEE, 2019.

[21] J. C. V. Tieck, L. Steffen, J. Kaiser, A. Roennau, and R. Dillmann, "Controlling a robot arm for target reaching without planning using spiking neurons," in *2018 IEEE 17th Int. Conf. on Cogn. Inform. & Cogn. Computing (ICCI* CC)*. IEEE, 2018, pp. 111–116.

[22] N. Srinivasa and Y. Cho, "Self-organizing spiking neural model for learning fault-tolerant spatio-motor transformations," *IEEE Trans. on neural networks and Learn. Syst.*, vol. 23, no. 10, pp. 1526–1538, 2012.

[23] Q. Wu, T. M. McGinnity, L. Maguire, A. Belatreche, and B. Glackin, "2d co-ordinate transformation based on a spike timing-dependent plasticity learning mechanism," *Neural Networks*, vol. 21, no. 9, pp. 1318–1327, 2008.

[24] A. Bouganis and M. Shanahan, "Training a spiking neural network to control a 4-dof robotic arm based on spike timing-dependent plasticity," in *The 2010 Int. Joint Conf. on Neural Networks (IJCNN)*. IEEE, 2010, pp. 1–8.

[25] D. Kappel, B. Nessler, and W. Maass, "Stdp installs in winner-take-all circuits an online approximation to hidden markov model learning," *PLoS Comput Biol*, vol. 10, no. 3, p. e1003511, 2014.

[26] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[27] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[28] M. N. Murty and V. S. Devi, "Hidden markov models," in *Pattern Recognition*. Springer, 2011, pp. 103–122.

[29] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Trans. on neural networks*, vol. 14, no. 6, pp. 1569–1572, 2003.

[30] L. F. Abbott, "Lapicque's introduction of the integrate-and-fire model neuron (1907)," *Brain research Bull.*, vol. 50, no. 5-6, pp. 303–304, 1999.

[31] D. M. Wolpert and M. Kawato, "Multiple paired forward and inverse models for motor control," *Neural networks*, vol. 11, no. 7-8, pp. 1317–1329, 1998.

[32] S.-J. Blakemore, D. Wolpert, and C. Frith, "Why can't you tickle yourself?" *Neuroreport*, vol. 11, no. 11, pp. R11–R16, 2000.

[33] A. Maravita and A. Iriki, "Tools for the body (schema)," *Trends in Cogn. sciences*, vol. 8, no. 2, pp. 79–86, 2004.

[34] I. Abadía, F. Naveros, J. A. Garrido, E. Ros, and N. R. Luque, "On robot compliance: A cerebellar control approach," *IEEE Trans. on cybernetics*, 2019.

[35] L. Vannucci, E. Falotico, S. Tolu, V. Cacucciolo, P. Dario, H. H. Lund, and C. Laschi, "A comprehensive gaze stabilization controller based on cerebellar internal models," *Bioinspiration & biomimetics*, vol. 12, no. 6, p. 065001, 2017.

[36] F. Chaumette and S. Hutchinson, "Visual servo control. Part I: Basic approaches," *IEEE Robot. Autom. Mag.*, vol. 13, no. 4, pp. 82–90, 2006.

[37] J. F. Kalaska, "From intention to action: motor cortex and the control of reaching movements," in *Progress in Motor Control*. Springer, 2009, pp. 139–178.

[38] S. Amari *et al.*, *The handbook of brain theory and neural networks*. MIT press, 2003.

[39] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, pp. 500–544, 1952.

[40] D. Hebb, "The organization of behavior. 1949," *New York Wiely*, vol. 2, no. 7, p. 8, 2002.

[41] E. Oja, "Simplified neuron model as a principal component analyzer," *Journal of Math. Biol.*, vol. 15, no. 3, pp. 267–273, 1982.

[42] D. Buonomano and T. Carvalho, "Spike-timing-dependent plasticity (stdp)," in *Encyclopedia of Neuroscience*, L. R. Squire, Ed. Oxford: Academic Press, 2009, pp. 265 – 268.

[43] N. Caporale and Y. Dan, "Spike timing–dependent plasticity: a hebbian learning rule," *Annu. Rev. Neurosci.*, vol. 31, pp. 25–46, 2008.

[44] M. Gilson, T. Fukai, and A. N. Burkitt, "Spectral analysis of input spike trains by spike-timing-dependent plasticity," *PLoS computational Biol.*, vol. 8, no. 7, 2012.

[45] V. Cutsuridis, S. Cobb, and B. P. Graham, "A ca 2+ dynamics model of the stdp symmetry-to-asymmetry transition in the ca1 pyramidal cell of the hippocampus," in *Int. Conf. on Artif. Neural Networks*. Springer, 2008, pp. 627–635.

[46] M. A. Woodin, K. Ganguly, and M.-m. Poo, "Coincident pre-and postsynaptic activity modifies gabaergic synapses by

postsynaptic changes in cl- transporter activity," *Neuron*, vol. 39, no. 5, pp. 807–820, 2003.

[47] K. Buchanan and J. Mellor, "The activity requirements for spike timing-dependent plasticity in the hippocampus," *Frontiers in Synaptic Neuroscience*, vol. 2, p. 11, 2010. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnsyn.2010.00011

[48] H. Ruan, T. Saur, and W.-D. Yao, "Dopamine-enabled anti-hebbian timing-dependent plasticity in prefrontal circuitry," *Frontiers in neural circuits*, vol. 8, p. 38, 2014.

[49] J.-C. Zhang, P.-M. Lau, and G.-Q. Bi, "Gain in sensitivity and loss in temporal contrast of stdp by dopaminergic modulation at hippocampal synapses," *Proceedings of the National Academy of Sciences*, vol. 106, no. 31, pp. 13 028–13 033, 2009.

[50] Z. Brzosko, W. Schultz, and O. Paulsen, "Retroactive modulation of spike timing-dependent plasticity by dopamine," *Elife*, vol. 4, p. e09685, 2015.

[51] Y. Hao, X. Huang, M. Dong, and B. Xu, "A biologically plausible supervised learning method for spiking neural networks using the symmetric stdp rule," *Neural Networks*, vol. 121, pp. 387–395, 2020.

[52] P. A. Gagniuc, *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons, 2017.

[53] E. M. Izhikevich and J. Moehlis, "Dynamical systems in neuroscience: The geometry of excitability and bursting," *SIAM review*, vol. 50, no. 2, p. 397, 2008.

[54] E. M. Izhikevich, "Which model to use for cortical spiking neurons?" *IEEE Trans. on neural networks*, vol. 15, no. 5, pp. 1063–1070, 2004.

[55] D. Sridharan and E. I. Knudsen, "Selective disinhibition: a unified neural mechanism for predictive and post hoc attentional selection," *Vision research*, vol. 116, pp. 194–209, 2015.

[56] K. D. Carlson, J. M. Nageswaran, N. Dutt, and J. L. Krichmar, "An efficient automated parameter tuning framework for spiking neural networks," *Frontiers in neuroscience*, vol. 8, p. 10, 2014.

[57] R. Saegusa, G. Metta, G. Sandini, and S. Sakka, "Active motor babbling for sensorimotor learning," in *Int. Conf. Robotics and Biomimetics*, Feb 2009, pp. 794–799.

[58] A. K. Fidjeland, E. B. Roesch, M. P. Shanahan, and W. Luk, "Nemo: a platform for neural modelling of spiking neurons using gpus," in *2009 20th IEEE Int. Conf. on Application-specific Syst., Architectures and Processors*. IEEE, 2009, pp. 137–144.

[59] D. Gamez, A. K. Fidjeland, and E. Lazdins, "ispike: a spiking neural interface for the icub robot," *Bioinspiration & biomimetics*, vol. 7, no. 2, p. 025008, 2012.

[60] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. on information theory*, vol. 8, no. 2, pp. 179–187, 1962.

[61] P. Sturm, *Pinhole Camera Model*. Boston, MA: Springer US, 2014, pp. 610–613.

[62] S. A. Neymotin, G. L. Chadderdon, C. C. Kerr, J. T. Francis, and W. W. Lytton, "Reinforcement learning of two-joint virtual arm reaching in a computer model

of sensorimotor cortex," *Neural computation*, vol. 25, no. 12, pp. 3263–3293, 2013.

[63] D. Navarro-Alarcon and Y.-H. Liu, "Fourier-based shape servoing: A new feedback method to actively deform soft objects into desired 2D image shapes," *IEEE Trans. Robot.*, vol. 34, no. 1, pp. 272–1279, 2018.