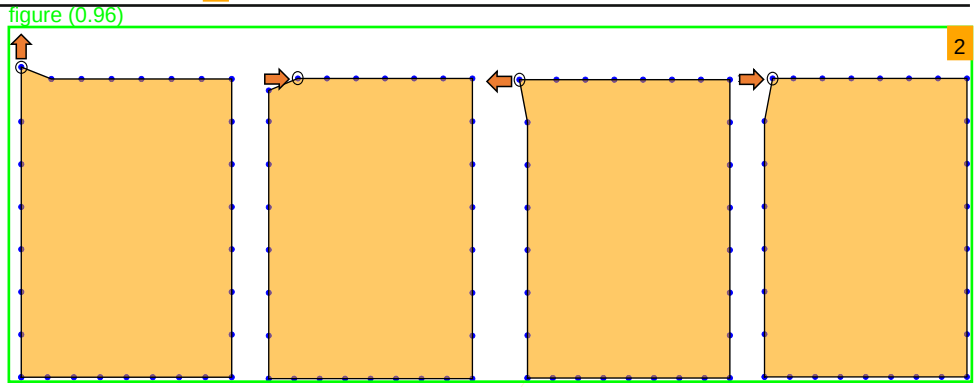

Fig. 2 Single grid char

3.3.1 Clusteri

(a) Circuit Grid Parameter Mat

The circuit grid parameter matrix consists of multiple gr points, each with coordinates. In this section, the parameter matrix describes a rectangle (Fig. 2) for the schematic. Three or more grid points are connected sequentially to construct a polygonal closure interval that describes the irregular structure. Each grid point has four directions of movement (up, down, left, and right). Polygon materials are usually conductors welded to a dielectric substrate.

(b) Characteristics Clusterin

The S-curve is an evaluation metric in filter design, and o model is based on the S-curve for learning. The performance of the transmission line circuit is studied based on the S-parameter, which can be equated to a two-port network. The S-parameter matrix of the two-port network is shown in Fig. 3, where S_{11} (reflected power/insertion power) is the reflection loss and S_{21} (output power/input power) is the insertion loss. A connection is created between the S-parameter and circuit grid parameter matrix, and the characteristics of the circuit are clustered based on the S-parameter.

As shown in Fig. 4, a single grid point has four shift dire tions. The S_{11} curve based on iterating each grid point shift 0.2 mm in four directions is calculated by invoking the *electromagnetic simulation-API* as a clustering dataset. The new circuit grid parameter matrix after a single grid point shift of 0.2 mm is shown in Fig. 4 d, g and j. The new and original S_{11} curves are shown in Fig. 4 e, h and k. The difference between the S_{11} curves(new curve and original curve without acti is shown in Fig. 4 f, i and l.

(c) Add Clustering Algorithm Reduce Explorat Space

The S_{11} curve is used as the evaluation index of the fi and calculated by invoking the *electromagnetic simulation-API* whenever the grid points move in any direction. As the number of grid points increases, the larger the number of grid points is, the greater the number of circuit solutions. Grid changes cause the filter to change and perform differently. For instance, the structure has 29 grid points, each of which can perform 4 operations (up, down, left, and right). Then, the agent action space is $116(29 * 4)$. It is difficult for the traditional reinforcement learning algorithm to converge within the task objective when there is an excessive amount of exploration space. Without clustering, the reinforcement learning algorithm has difficulty converging when there is too much action space. However, the action space can be effectively reduced by clustering the difference (result of the actions) of S_{11} curve. Details can be seen in Fig. 4.

Based on the K-means algorithm, the quality of the clu centre is essential to the clustering quality [43]. As shown in Fig. 5, the difference between the S_{11} curve is used as the initial value for clustering. When the grid points move in any direction, the *electromagnetic simulation-API* will be automatically invoking to calculate the S_{11} curve. Finally, four clustering results are obtained after computing all grid points (Fig. 4) based on the differential S_{11} curves and the initial values (Fig. 5). These results include S_{11} shifting downwards, leftwards, rightwards, and upwards. The filter design task requires the S_{11} curve to be shifted downwards within the specified frequency band without being shifted upwards. In other words, the upward shift of the S_{11} curve will lead to a decrease in filter performance. Therefore, the upward shift clustering results were excluded to ensure the performance of the filter.

In the reinforcement learning process, the state is defi as the grid parameter matrix, and the rewards can be seen in Section 3.3.2b. When in state $\{s_i, s_{i+1} \ldots, s_n\}$, any action by the policy model $\pi\ (a_i \mid s_i; \theta)$ makes the rewards

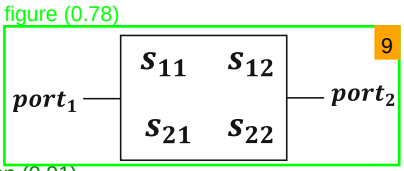

Fig. 3 S-parameter matrix of the two-port netw