J Vis (2023) 26:977–998
https://doi.org/10.1007/s12650-023-00913-7

REGULAR PAPER

Fuqiu Chen · Jizhuo Li · Fengjie Wang · Shangsong Liu · XiaoLin Wen · Pengyuan Li · Min Zhu

WarehouseLens: visualizing and exploring turnover events of digital warehouse

Received: 6 November 2022 / Revised: 7 January 2023 / Accepted: 1 February 2023 / Published online: 26 February 2023

Abstract Goods turnover is the core of digital warehouse operation, including many processes, such as receiving, picking, and packing of goods. Analyzing goods turnover data can generate valuable insights for optimizing warehouse management, thereby improving operation efficiency. However, most existing methods focus on partial processes, making it hard for warehouse managers to understand the operation state and the goods turnover patterns, which often require the analysis of the interrelated processes of goods turnover. In this paper, we abstract six types of goods turnover events to describe the warehouse operation workflow and present WarehouseLens, a visual analytics system to analyze goods turnover from an overall perspective. To understand the warehouse operation state, we propose a temporal visualization method consisting of a novel state calendar view and an improved circular heat map to reflect the trend and periodicity pattern of the operation state. To explore the goods turnover patterns, we provide an improved parallel coordinate plot for users to view the attribute distribution of goods to filter key goods and a tailored mode circle view to discover the frequent outbound mode of goods. Three case studies and expert interviews on a real-world warehouse dataset demonstrate the usefulness and effectiveness of WarehouseLens in revealing the warehouse operation state and goods turnover patterns.

Keywords Digital warehouse · Turnover event sequence · Visual analytics · Warehouse management

F. Chen · J. Li · F. Wang · S. Liu · X. Wen · M. Zhu (✉)
College of Computer Science, Sichuan University, Wuhou District, Chengdu 610065, China
E-mail: zhumin@scu.edu.cn

F. Chen
E-mail: chenfuqiu@stu.scu.edu.cn

J. Li
E-mail: lijizhuo@stu.scu.edu.cn

F. Wang
E-mail: wangfengjie@stu.scu.edu.cn

S. Liu
E-mail: piny@stu.scu.edu.cn

X. Wen
E-mail: wenxiaolin@stu.scu.edu.cn

P. Li
Sichuanwulianyida Technology Co., Ltd., Chuanshan District, Suining 629000, China
E-mail: lipengyuan@50yc.com

1 Introduction

Warehouses are an essential component of supply chains, whose major role is to consolidate products from various suppliers for combined delivery to customers (Gu et al 2007). Goods are the basic unit of warehouse management, and the flow of goods in the warehouse operation is called goods turnover (Fig. 1), during which the goods turnover data are generated to record the state changes of goods. Analyzing goods turnover data can provide effective insights for optimal warehouse management.

Through collaborating with domain experts, it is concluded that warehouse management generally involves six stages of goods turnover (Fig. 1), i.e., six types of goods turnover events. (1) *Arriving*. Different goods owners deliver their goods to the partner warehouse for management. They regularly create replenishment orders and ship the goods to the warehouse at a specified time. (2) *Receiving*. Warehouse operators check the quantity and quality of goods and then, receive these goods. (3) *Storing*. According to the type of goods, operators place the goods on the corresponding shelves. (4) *Allocating*. Each customer order is composed of one or more stock-keeping units (SKU), i.e., goods. After the warehouse receives customer orders, the warehouse system creates outbound batches for all the goods included in orders and distributes the list of outbound batches to specific operators. (5) *Picking*. According to the outbound list, operators go to the corresponding shelf and pick target goods. (6) *Packing*. Operators check the quality of goods and pack up the goods.

Prior warehouse management optimization work mainly focuses on the individual stage in warehouse operation, such as storage location assignment in *Storing* stage (Yang et al 2015; Pan et al 2015; Ramtin and Pazour 2015; Guo et al 2016) and routing and order batching in *Picking* stage (De Koster et al 2007; Tappia et al 2019; Pinto and Nagano 2019; Jaghbeer et al 2020). However, warehouse operation management in various stages is logically interrelated and the independent analysis of individual stage is hard to improve the overall operation efficiency of the warehouse (Zhen and Li 2022). Therefore, the integrated analysis of goods turnover should be implemented for further improvement of warehouse performance. It can help warehouse managers to recognize the impact of each stage on the warehouse operation state and make integrated decision-making in warehouse management. For example, storage location assignment in *Storing* stage and order batching in *Picking* stage together determine the path and distance for goods picking. It is necessary to consider both stages to improve overall order picking efficiency and shorten order completion time.

In addition to the integrated analysis of the whole process, warehouse management optimization should also focus on outbound area activities in the goods turnover process, which turns out to be the most complex, time-consuming, and costly in a warehouse (Lee et al 2018). Consequently, an efficient goods outbound process arises to be crucial for warehouse management. However, there are many hidden but potentially valuable patterns in outbound area activities, e.g., goods with large inventory fluctuations, high outbound frequency, or abnormal picking time, the whole of which are considered to be the key goods in the outbound process. It is challenging to uncover such valuable information manually to achieve effective outbound management. In the real scenario, some general views, like pie, bar, and line charts, are used to show operation parameters of outbound area activities, e.g., outbound order quantity, goods turnover rate, etc. However, these approaches are only a simple presentation of the data, lacking effective visualization and interaction to guide warehouse managers to adopt more accurate management strategies.

To address the above issues, we present WarehouseLens (Fig. 2), an interactive visual analytics system that enables warehouse managers to gain insights from the comprehensive analysis of goods turnover data. In our work, we integrate six stages of goods turnover data and abstract them into time-tagged event sequences for further analysis. To understand the warehouse operation state, a novel state calendar view is

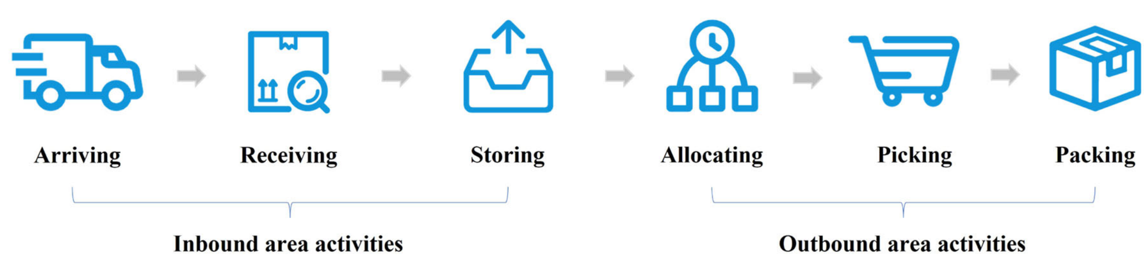

Fig. 1 The workflow of goods turnover in the digital warehouse, which includes the inbound area activities and outbound area activities (Lee et al 2018)

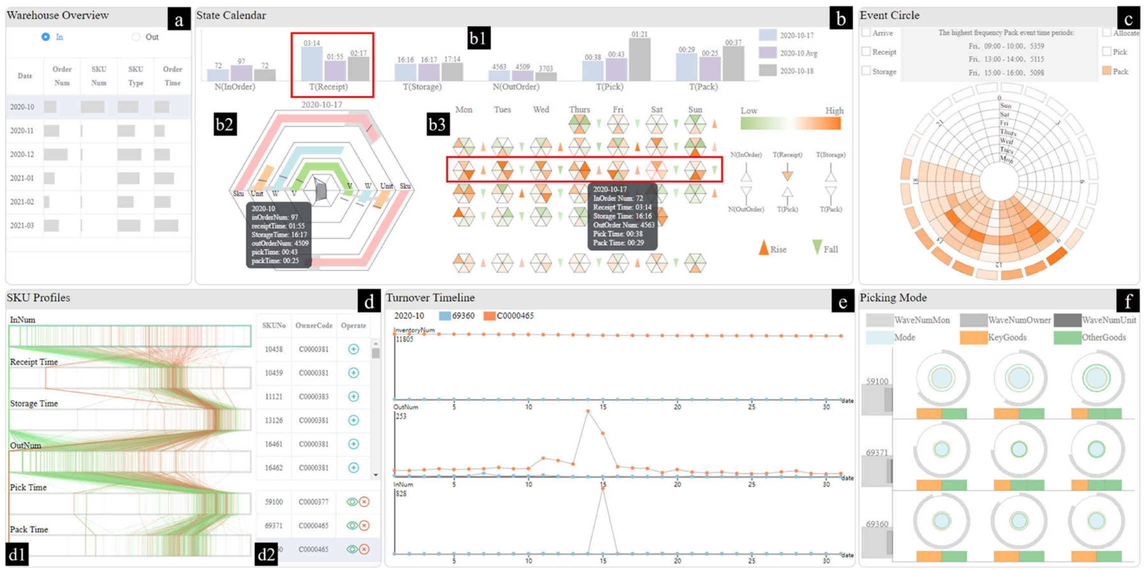

Fig. 2 System user interface. **a** *Warehouse Overview* provides an overview of inbound and outbound activities. **b** *State Calendar View* supports analysis of daily warehouse operation state. **c** *Event Circle View* depicts the periodicity pattern of six goods turnover events. **d** *SKU Profiles View* shows distribution patterns of goods attribute. **e** *Turnover Timeline View* displays daily inventory changes of goods. **f** *Picking Mode View* presents the frequent outbound mode of key goods

proposed to reflect temporal changes of the warehouse operation state. An improved circular heat map aims to present the frequency of six goods turnover events to support the discovery of periodic patterns. To explore the goods turnover patterns, we adopt an improved parallel coordinate plot to reflect the distribution pattern of goods attributes to filter key goods. And a customed mode circle view is designed to compare and analyze frequent outbound modes of key goods, which are exploited by the frequent pattern mining algorithm. Finally, we evaluate the usefulness and effectiveness of our work through three case studies and expert interviews.

The contributions of this study are as follows:

- We design an interactive system WarehouseLens to facilitate the visual analysis of goods turnover events. It integrates the whole process of goods turnover supporting the understanding of the warehouse operation state and focuses on the outbound area activities for the exploration of goods turnover patterns.
- We propose two tailored visualization designs to further understand goods turnover data, including the state calendar view reflecting changing state of warehouse operation and the mode circle view showing frequent outbound modes of key goods.
- We conduct three case studies on a warehouse dataset and expert interviews to evaluate the usefulness and effectiveness of WarehouseLens.

2 Related work

In this section, we discuss studies that are most relevant to WarehouseLens, including warehouse data visualization and event sequence visualization.

2.1 Warehouse data visualization

Warehouse data visualization research mainly focuses on operation process monitoring and warehouse scene simulation. Zhong et al (2015) processed workers' trajectory data to visualize their work patterns, providing suggestions for warehouse efficiency improvement. Tarigonda et al (2018) designed a dashboard to analyze workers' operational processes to reduce the risk of warehouse operations. Xu et al (2017) tracked assembly line performance for real-time scenarios and support the historical data exploration to identify inefficiencies and locate anomalies. Jo et al (2014) improved the conventional Gantt chart for visualizing manufacturing schedules with better scalability. Bräuer and Mazarakis (2020) visualized the turnover rate of goods through augmented reality (AR) to guide the allocation of storage space. Fang et al (2019) applied AR to warehouse order picking for tracking low-cost goods. Cogo et al (2020) proposed a 3D visualization method to analyze goods' volume and storage space, helping users quickly retrieve target goods.

Most traditional approaches focus on the individual analysis of warehouse operation processes, without considering the integrated analysis of the whole process, which is insufficient to reveal the warehouse operation state. In addition, existing studies concentrate on tracking and retrieving goods at the goods level, ignoring the analysis of outbound patterns for specific goods. Tang et al (2022) abstracted the order processing workflow and presented an interactive system OrderMonitor to facilitate real-time order monitoring, whose analysis tasks were the detection and handling of delayed orders. Different from OrderMonitor, our work describes the workflow of goods turnover for a better understanding of warehouse operation state and goods turnover patterns.

2.2 Event sequence visualization

There is much visual analytics for event sequence data, like pattern discovery (Wongsuphasawat and Shneiderman 2009; Perer and Wang 2014), sequence inference (Mei and Eisner 2017; Xu et al 2016), and sequence modeling (Jin et al 2020; Guo et al 2019). Pattern discovery aims to find frequently occurring patterns and statistically significant associations of data samples (Guo et al 2022). And our work is most relevant to pattern discovery of event sequences.

To reveal the patterns hidden behind the event sequence, one most straightforward visualization is the timeline-based visualization. This method organizes events of raw sequences successively along a temporal axis, as done by LifeLines (Plaisant et al 1998), Cloudlines (Krstajic et al 2011), Timeslice (Zhao et al 2012) and Vasabi (Nguyen et al 2020). However, these techniques can lead to substantial cognitive load and hinder the identification of patterns as the number of sequences increases (Wang et al 2022a). To tackle this challenge, one solution is the tree-based visualization. This method generates an aggregated overview of multiple records. For example, ActiviTree (Vrotsou et al 2009) used a node-link tree for the systematic identification of sequences in social science activity diary data; Google+ Ripples (Viégas et al 2013) proposed a treemap metaphor to show how public posts were shared on Google+. Another solution is the Sankey-based visualization. Outflow (Wongsuphasawat and Gotz 2012) and CareFlow (Perer and Gotz 2013) condensed sequences into transition graphs and visualized the sequences using a Sankey diagram.

Several authors have proposed the matrix-based visualization to reduce visual confusion caused by the dense edges in the tree-based and Sankey-based visualization. Zhao et al (2015) adopted a matrix-based visualization to provide an overview of the differences in click traffic patterns. Du et al (2016) designed an event matrix to present the frequency of events in different time intervals. Mu et al (2019) combined a flow-based approach with a matrix-based approach to present users' learning patterns at different stages. Zhang et al (2022) employed a 45-degree-rotated matrix to show the events involving different historical figures, with color encoding the event quantity.

The tree-based visualizations emphasize the hierarchical structure of the event sequence and the Sankey-based visualizations focus more on providing an overview of transitions between different types of events. There is little hierarchical structure or transition in the goods turnover event sequence. Thus, the above two visualizations are not applicable. The matrix-based visualizations are typically applied to demonstrate a summary of event frequency or frequent patterns. However, the frequent outbound mode of goods contains a lot of information, e.g., the outbound frequency, combination of outbound goods, and picking time of the combination. Traditional matrix-based visualizations can only present limited information, like outbound frequency. Thus, tailored visual design is required for comprehensive reflection of the frequent outbound mode of goods.

3 Background and system overview

In this section, we first give a detailed description of the data used in our study. Then, we summarize requirements and analytical tasks based on iterative interviews with domain experts. Finally, we give a system overview to demonstrate the whole pipeline.

3.1 Data description

The goods turnover data studied in this paper is provided by a warehouse company from October 2020 to March 2021, including inbound, outbound, and inventory data. Inbound data describe goods inbound activities, including 80,000 inbound orders and 250,000 inbound sub-orders. Outbound data depict goods outbound activities, including 2.74 million outbound orders and 10.33 million outbound sub-orders. Inventory data describe the daily inventory quantity of various goods containing 110,000 inventory records.

For the original dataset, we filter the incorrect records and fill in the missing records. For example, delete the record whose reception time is earlier than the arrival time; use the order completion time to replace the packing time of goods. Further, we convert and integrate the cleaned data to obtain valuable information about goods. Based on the time stamps on the inbound and outbound orders, we calculate the daily operation time of different goods turnover events. After that, we count the number of daily orders to finally obtain the *Daily State Data*. It includes six state indicators, i.e., the number of inbound and outbound orders, and the time of receiving, storing, picking, and packing goods every day. These state indicators are defined through discussions with our collaborating experts. To identify the periodic patterns in the warehouse, we calculate the frequency of six goods turnover events from different time granularities (i.e., each month, week, day, and hour) to finally get the *Event Frequency Data*. Combined with the inbound and outbound sub-orders and inventory data, we calculate the daily inbound and outbound quantity and storage quantity of goods to get the *Goods inventory data*, which shows daily inventory changes of goods.

To cope with the exploration of goods turnover patterns mentioned in Sect. 1, we focus on the outbound mode of goods after an in-depth discussion of the requirements with domain experts. Experts point out that in the outbound process, orders belonging to the same owner are grouped into the same outbound batch and are picked uniformly. Therefore, goods in the same batch have the same picking time. By exploring the outbound combinations of goods with high frequency in the same batches and analyzing the average picking time of these combinations, we can judge whether an outbound mode performs abnormally in the picking event, assisting warehouse managers in adjusting goods storage location. In this paper, we employ the frequent pattern mining algorithm FP-Growth (Han et al 2000) to find high-frequency outbound combinations of goods. The algorithm uses a tree structure for higher computational speed and its workflow is as follows. Firstly, we count the batches of each type of outbound goods and build the transaction dataset *itemModeList* of goods. Secondly, scan *itemModeList* and count the frequency of different goods. Goods with a frequency greater than the minimum support A are added to the frequent item list L. Thirdly, iterate over L to construct the FP tree. Finally, find the frequent sets recursively from the leaf nodes of the FP tree and then, integrate them into the *Outbound Mode Data*. It presents detailed information on the frequent outbound mode of goods. Take goods A as an example, the *Outbound Mode Data* contains a mode list of goods A: [mode1, mode2, mode3,...]. Regarding the mode1, it contains goods composition (i.e., goods A, B, and C), picking time (i.e., the individual picking time of goods A, B, and C; and the total picking time of this combination), outbound frequency of goods A and its owner (the owner indicates which supplier the goods belong to).

3.2 Requirement analysis

Following the user-centered design study methodology proposed by Sedlmair (Sedlmair et al 2012), we collaborate with a warehouse company's three domain experts (denoted as E1, E2, and E3) to integrate domain knowledge and expertise into the analysis loop (Deng et al. 2023). E1 (male, age: 27) is a product manager with 5 years of working experience. Both E2 (female, age: 32) and E3 (male, age: 35) are warehouse managers with at least eight-year working experience. At the early stage of the collaboration, we hold weekly online meetings with our experts to analyze their requirements when managing the warehouse, which include several stages. Firstly, experts explain to us the business scenario and collected data (e.g., data volume, type, and attributes). Secondly, we ask them to illustrate the problems they face when managing the warehouse and confirm with them the problems we find from the literature review. Finally, we distill two design requirements *(R1 and R2)* from several rounds of iterative communication with domain experts to guide the system development.

R1: Understand the operation state of the warehouse at the overall level. Warehouse managers wish to establish an overall understanding of the warehouse operation state. On the one hand, they need to summarize warehouse operation's trend and periodicity pattern, which is helpful to the rational allocation of production resources. On the other hand, experts need to discover and analyze abnormal conditions during the operation process, providing a reference for disposing of anomalies.

R2: Explore the turnover patterns of goods at the individual level. Goods are the basic unit of warehouse management. Managers need to find key goods and conduct an in-depth analysis of their turnover patterns, like outbound modes and inventory changes, which benefit storage space allocation and reasonable stocking.

3.3 Design tasks

Based on requirements *(R1 and R2)* in *Section 3.2*, we identify two categories of six analysis tasks *(T1-T6)* to guide our visual design. The first type of task *(T1-T3)* focuses on the understanding of the warehouse operation state. And the second type of task *(T4-T6)* is related to the exploration of goods turnover patterns.

T1: Understand the warehouse operation state from multipletime granularity (R1). Warehouse managers want to comprehensively compare and understand the warehouse operation state from different time granularities. For example, the number of customer orders per month, the number of goods inbound and outbound per week, the execution time of turnover events per day, and the frequency of turnover events per hour.

T2: Analyze the trend and periodicity pattern of the warehouseoperation state (R1). Based on our discussion with experts, the analysis of the trend and periodicity pattern of goods turnover events can help to identify critical operation time points, which is useful for experts to adjust production and human resources. For example, it is found that the picking event is often scheduled in the morning, so warehouse managers need to assign more workers during the period to improve picking efficiency.

T3: Reveal the abnormal operation state of the warehouse (R1). As experts point out, anomalies in warehouse operation may result in delayed orders and failure to provide timely delivery to customers. It is necessary to discover and analyze the abnormal warehouse operation state, which can guide warehouse managers to prepare for and deal with anomalies, ensuring timely delivery to customers. The visual design should provide a detail-on-demand comparison of the operation state on different dates, helping warehouse managers quickly check and locate anomalies for decision making.

T4: Filter key goods based on warehouse state indicators (R2). Different goods have different receiving time, picking time, packing time, etc. In other words, the distribution of goods on warehouse state indicators varies. The visual design needs to present the distribution pattern of goods, helping managers filter key goods quickly for further exploration.

T5: Summarize the frequent outbound mode of goods (R2). The visual design should help warehouse managers analyze the frequent outbound mode of goods while allowing them to find the abnormal mode, thus guiding storage location adjustment for goods. For example, if there is an outbound combination of goods with a long picking time, managers need to consider whether to adjust the storage location of goods to shorten the picking time.

T6: Show the changing pattern of goods inventory (R2). Warehouse managers wish to know the changes in goods inventory to facilitate reasonable stocking. For example, it is found that the number of goods inbound and outbound fluctuates wildly, meaning that such goods need to be prepared in sufficient quantities.

According to the feedback from the experts, the above two categories of analytical tasks are interrelated to support knowledge discovery and decision making. For example, when analyzing the daily operation state of the warehouse, when warehouse managers find abnormal picking events, they can try to check the day's goods turnover pattern to locate the cause of the anomaly.

3.4 System overview

According to the analysis tasks in Sect. 3.3, we implement an interactive visual analysis system, WarehouseLens (Fig. 2), to assist warehouse managers in understanding the warehouse operation state and exploring the turnover pattern of goods as well. The system overview is illustrated in Fig. 3. We integrate three types of data: inbound, outbound, and inventory data. The data processing module includes data cleaning, filtering, and frequent pattern mining. And then, the processed data are stored structurally. The visualization module has a prototype system with multiple coordinated views and interactions.

The workflow of WarehouseLens is as follows. Users begin to check warehouse inbound and outbound activities in different months from *Warehouse Overview* and select a specific month to start the exploration.

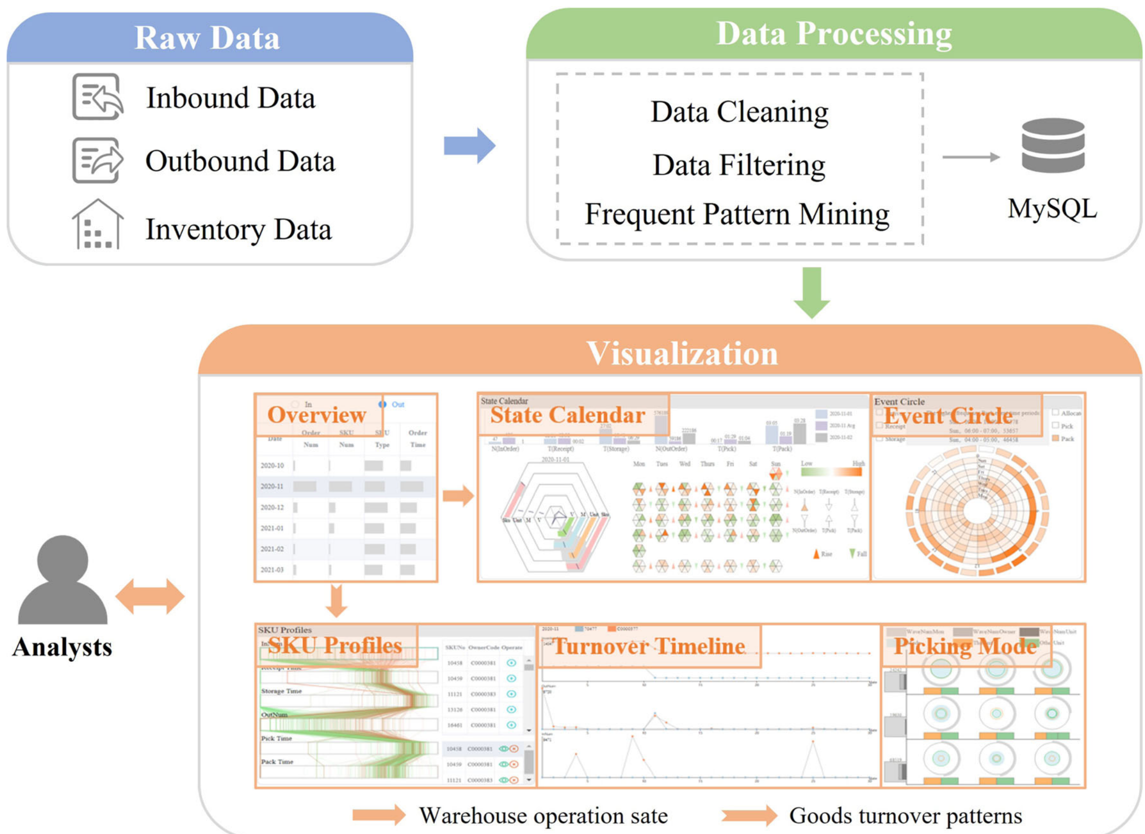

Fig. 3 System overview. WarehouseLens contains three modules: raw data, data processing, and visualization

Then, they can focus on discovering trend changes and periodic patterns in warehouse operation through the *State Calendar View* and *Event Circle View (T1 and T2)*. These two views help users analyze turnover events from different levels of details and discover potential operation anomalies *(T3)*. Next, users can also better understand the details of warehouse operation by exploring goods turnover patterns. They can view the distribution pattern of goods attributes and then filter key goods in the *SKU Profiles View (T4)*. Finally, they can add those key goods to the *Picking Mode View* and *Turnover Timeline View* for the analysis of frequent outbound modes and inventory changes *(T5 and T6)*. We introduce the design and implementation of these views in the following section.

4 Visualization design

As illustrated in Fig. 2, the visual interface consists of six collaborative views, including the *Warehouse Overview*, *State Calendar View*, *Event Circle View*, *SKU Profiles View*, *Turnover Timeline View*, and *Picking Mode View*. In this section, we present a detailed description of the visual designs and user interactions for each view.

4.1 State calendar view

The *State Calendar View* (Fig. 2b) consists of three components: the calendar overview (b3), hexagon analysis view (b2), and tri-bar comparison view (b1). This view can present the daily operation state of the warehouse *(T1)* and reflect the changing trend in the warehouse state indicators *(T2)*, giving support to the discovery of warehouse anomalies *(T3)*.

The calendar overview (Fig. 2b3) is an overview of the daily operation state of the warehouse, composed of the state hexagon and the trend triangle. As shown in Fig. 4a, six small triangles in a state hexagon represent six state indicators (see Sect. 3.1). From top to bottom, from left to right, the six small triangles

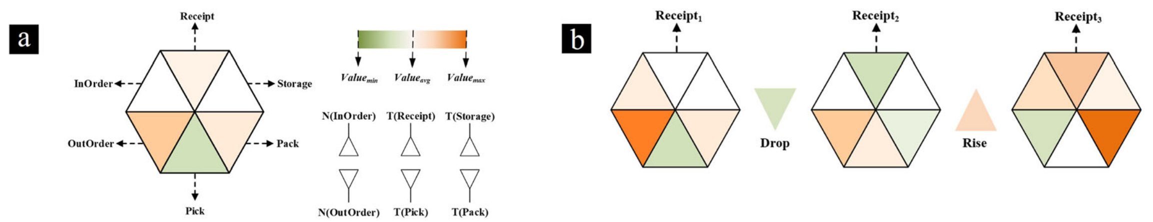

Fig. 4 **a** A state hexagon contains six small triangles (i.e., six state indicators), with color encoding value of indicators. Two rows of horizontally aligned funnel legends at the bottom right are legend identification for six small triangles. **b** A trend triangle between two state hexagons reflects the rising or dropping of the indicator value. The downward green triangle encodes a dropping value, while the upward orange triangle encodes a rising value

are, respectively, the number of inbound and outbound orders, and the time of receiving, storing, picking, and packing goods every day. Users can also identify these triangles by checking the relative positions of two rows of horizontally aligned funnel legends on the rightmost side (e.g., the legend identification *N*(*InOrder*), *T*(*Receipt*), etc.). The color of triangles encodes the value of state indicators, with white representing the monthly average value, green indicating the lower value, and orange indicating the higher value. All state hexagons are arranged according to the calendar layout in Fig. 2b3, and the hexagons in the last row represent the overview of the operation state from Monday to Sunday. The trend triangle between state hexagons shows how the state indicator (i.e., one of the six state indicators, which is selected by users) changes in adjacent days, whose direction indicates the rise or drop of the value of state indicators. Its color coding is the same as the state hexagon. As Fig. 4b shows, the downward green triangle encodes a dropping value, while the upward orange triangle encodes a rising value.

The hexagon analysis view (Fig. 2b2) is used to analyze the influence of goods attributes on the warehouse operation state, which comprises the outer stacked bullet chart and the inner radar chart. As displayed in Fig. 5a, with the horizontal axis as the dividing line, the upper side of the hexagon analysis view represents inbound goods, and the red, orange, blue, and green bands (i.e., SKU_{in}, $Unit_{in}$, $Weight_{in}$, $Volume_{in}$) correspond to the values of the four goods attributes (i.e., the type, quantity, weight and volume of inbound goods) on the day. The black tick marks (i.e., SKU_{avgIn}, $Unit_{avgIn}$, $Weight_{avgIn}$, $Volume_{avgIn}$) correspond to the average values of the four goods attributes in the current month. The gray ribbons (i.e., SKU_{Qin}, $Unit_{Qin}$, $Weight_{Qin}$, $Volume_{Qin}$) correspond to the upper and lower quartile of the attribute values in the current month. The lower side of the hexagon analysis view represents the outbound goods. The angle in Fig. 5b encodes the attribute value. Taking the inbound goods as an example $\theta_{SKUQinL}$ encodes the lower quartile value, while $\theta_{SKUQinU}$ is the upper quartile value. The larger the value, the larger the angle. Similarly, for outbound orders, $\theta_{volumeOut}$ represents the total volume of outbound goods on the day, and $\theta_{volumeAvgOut}$ indicates the average volume of outbound goods in the current month. The radar chart in Fig. 5a is used to compare the state indicators of the current month and the day. The hexagon $State_{mon}$ indicates an overview of the current month, while the hexagon $State_{date}$ shows the details of the day.

The tri-bar comparison view compares the warehouse state on different days with the current month. The comparison of two consecutive days is helpful to infer whether an anomaly is incidental or persistent. In addition, the aim of comparing with the current month is to determine the degree of anomaly by deviation from the average. All comparisons are made for more accurate judgments and better decision making. Figure 2b1 shows the operation state comparison between October 17, October 18, and October as a whole.

Design alternatives. During the iterative process of our system, two candidate designs are discussed to show six state indicators, including a bar chart and a radar chart. Although the bar chart can combine the above six data, it is difficult to set a scale standard for them due to the different ranges between order quantity and event time. The radar chart is useful in comparing multiple quantitative variables. However, given that we need to know the warehouse operation state for each day of any month, it is challenging to perform inter-comparisons Lin et al (2021). Finally, after discussions with domain experts, we propose the design of the state hexagon, which could be divided into six identical small triangles to represent six state indicators. Experts believed that they could perceive changes in state indicators through color differences of these state hexagons. In addition, organizing the state hexagons in a calendar layout is helpful for a quick comparison of operation states across different days.

To visualize the goods attributes of each day, our initial consideration was a vertically aligned bullet chart in Fig. 5c. In each bullet unit, there are multiple scales (i.e., the upper quartile, lower quartile and mean) are added to depict the attribute value. In addition, our experts also pointed out that they wanted to

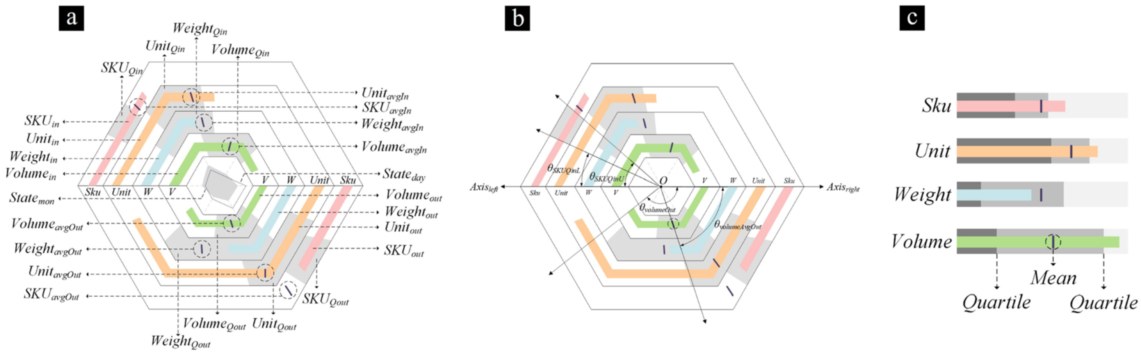

Fig. 5 **a** A hexagon analysis unit is composed of stacked bullet charts, which can be divided into upper and lower parts, representing information on inbound and outbound goods, respectively. **b** Angle in the hexagon analysis unit encodes goods attribute value. **c** An alternative solution to the hexagon analysis unit

check the attributes of both inbound and outbound goods. However, it is not space efficient to arrange eight bullet units vertically, and not convenient to compare different goods attributes. Finally, with the advice of experts, we came up with the design of a stacked bullet chart and integrated it into the hexagon, which could maintain visual consistency with the state hexagon. Also, encoding attribute values with angles in hexagons allows for clearer comparisons.

Interaction. In the rightmost of the calendar overview Fig. 2b3, there are two rows of horizontally aligned funnel legends, which represent the six state indicators. Users can click on any funnel legend to update trend triangles when aiming to analyze the temporal changes of a specific state indicator. After that, users can click on a state hexagon to update the hexagon analysis view Fig. 2b2 to check operation details (i.e., different attributes of inbound and outbound goods) on any date. At the same time, the *SKU Profiles View* is updated immediately to display the detailed information of outbound goods on the selected date, which is convenient for subsequent joint analysis. To compare the operation state of different dates, users can double-click on two hexagons to add the two days into the tri-bar comparison view Fig. 2b1 for further comparison. We use highlighting, bolding, and tooltips to stand out specific visual elements and provide additional information to improve users' cognitive efficiency. For example, by hovering the mouse over a state hexagon, users can view the date and the complete information of the six state indicators, i.e., the inbound and outbound order quantity, the receipt, storage, picking, and packing time of goods on the selected day.

4.2 Event circle view

The *Event Circle View* (Fig. 2c) utilizes a circular heat map to present the frequency of six types of goods turnover events per week, day, and hour. It can reveal the periodicity pattern of the warehouse operation state comprehensively *(T1 and T2)*.

In this view, we use circular grids to encode different periods. As shown in Fig. 6b, starting from the position *Start*, the 24 columns of grids arranged clockwise encode 24 h from 0 o'clock to 23 o'clock, respectively. There are eight rings from outside to inside in total. The first ring encodes the selected month, and the second to eighth rings encode from Sunday to Monday. For example, $Frequency_{0-1}$ represents the frequency of *Arrive* events from 0 o'clock to 1 o'clock in October (October is currently the selected month in the *Warehouse Overview*). $Frequency_{Wed-19-20}$ encodes the frequency of *Arrive* events each Wednesday from 19 o'clock to 20 o'clock in October. The grid color encodes the event frequency; the higher the frequency, the darker the color. Compared with the traditional circular heat map, we add a ring at the outermost layer to observe the event frequency at any hour in the selected month. It takes full advantage of the time granularity of turnover event sequences, and the gap between adjacent grids reduces visual confusion.

In the circular heat map, we distinguish the different periods by successive adjacent circular grids with a similar color in the circumferential direction (e.g., circular grids with blue outline) and the radial direction (e.g., circular grids with red outline). For example, in the single ring labeled *Fri* in Fig. 6b, there are three successive circular grids from $Frequency_{Fri-13-14}$ to $Frequency_{Fri-15-16}$ with a similar darker color, which indicates that *Arrive* events continue to occur during this period. It is the same with the radial direction. For

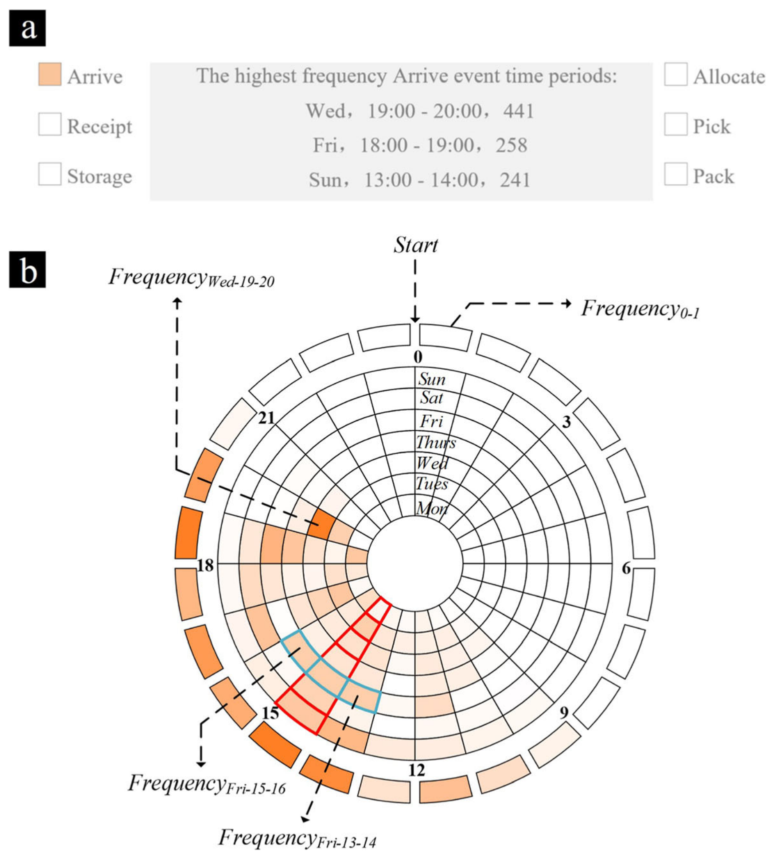

Fig. 6 The visual design of the *Event Circle View*. **a** Six event buttons on the left and right and a text description in the middle. **b** In the circular heat map, grids encode different periods, and the color of the grids encodes the event frequency. Grids with a blue outline indicate the circumferential direction, and grids with a red outline indicate the radial direction

example, in the fifteenth column of circular grids in the clockwise direction, there are seven successive circular grids with a similar darker color, which shows that there are more *Arrive* events from 14 o'clock to 15 o'clock on any day from Monday to Sunday. For different periods, we judge whether it is a longer period or a shorter period by the number of successive circular grids. For example, in the ring labeled *Wed*, there is only one circular grid with a darker color (i.e., $Frequency_{Wed-19-20}$), which indicates a large number of *Arrive* events occur only during this period. While in the ring labeled *Fri*, there are three successive circular grids (i.e., $Frequency_{Fri-13-14}$ to $Frequency_{Fri-15-16}$) showing continuous *Arrive* events. Therefore, the former can be called a short period and the latter a longer period. Warehouse managers need to schedule a varying number of workers at different periods to ensure the operation efficiency of the warehouse.

Design alternatives. During the design iteration, initially, we decided to use the line chart to reflect the event frequency. The line chart can analyze the trend of data over time, while it is space-wasting. After discussions with domain experts, we knew that goods turnover events in the warehouse had a periodicity dependency. Therefore, we finally replaced the line chart with a circular heat map, which could better convey the periodicity pattern.

Interaction. The left and right sides in Fig. 6a are the toggle buttons of the six events. Users can inspect any event by clicking these buttons (*Arrive* is currently clicked). Then, the text description in the middle is also updated to reveal the three periods with the highest event frequency. In the circular heat map Fig. 6b, we provide highlights and tooltips to help users know the specific event information in the selected period. Users can view detailed information (e.g., the event type, event frequency, and period) by hovering the mouse over a grid.

4.3 SKU profiles view

The *SKU Profiles View* (Fig. 2d) contains two components: the parallel rectangular view to reveal potential clusters and anomalies of goods and the goods list view to give a detailed description of filtered goods. This view can show the distribution pattern of goods on different state indicators and support filtering goods worthy of further analysis quickly *(T4)*.

Inspired by the work (Sun et al 2020; Hou et al 2022), we extend the axis in the traditional parallel coordinate plot to a rectangle to clearly show the goods' distribution on different state indicators. Six parallel rectangles in the parallel rectangular view (Fig. 2d1) encode six types of goods attributes, including the inbound number, receipt time, storing time, outbound number, picking time, and packing time. Each line in the rectangular represents a SKU (i.e., the goods included in customer orders) whose position depicts the attribute value. For example, $PackTime_{SKU}$ in Fig. 7a indicates the packing time of the goods *SKU*. Users can select different attributes to sort goods and then, analyze the distribution pattern to filter the key goods. The goods list view (Fig. 7b) includes a general list above and a detailed list below, and each row of the two lists represents a SKU. The general list shows the whole goods filtered from the parallel rectangular view, which consists of *SKUNo*, *OwnerCode*, and *Operate* (i.e., *Add SKU*). The detailed list shows the key goods selected from the general list, composed of *SKUNo*, *OwnerCode*, and *Operate* (i.e., *See SKU* and *Delete SKU*).

Design alternatives. Two alternative designs are considered to display the distribution and clustering of goods. The first design is a scatter plot with the layout generated by the multidimensional scaling (MDS) technique Kruskal (1964). However, the domain experts reported that they could not identify the distribution of goods on the six attributes (i.e., the inbound number, receipt time, storing time, outbound number, picking time, and packing time). Another alternative design is the traditional parallel coordinates plot. After discussing with experts, they thought that it was not intuitive to find clusters with points on the coordinate axis. Thus, we extended the axes to rectangles as discussed above, and used lines in the plot to represent goods.

Interaction. Users can click any parallel rectangle to select the sorting attribute, with the chosen rectangle marked green. As shown in Fig. 2d1, when the inbound number (i.e., the uppermost parallel rectangle in the *SKU Profiles View*) is selected as the sorting attribute, the position and color of all lines will update, where the green line indicates the number of inbound goods is small, while the red indicates the number is large. Users can brush lines in any rectangle to add goods in a specific attribute range to the general list. After that, users can click the *Add SKU* icon to add the target goods to the detailed list, and the outbound modes of these target goods will be shown in the *Picking Mode View*. Users can also click the *See*

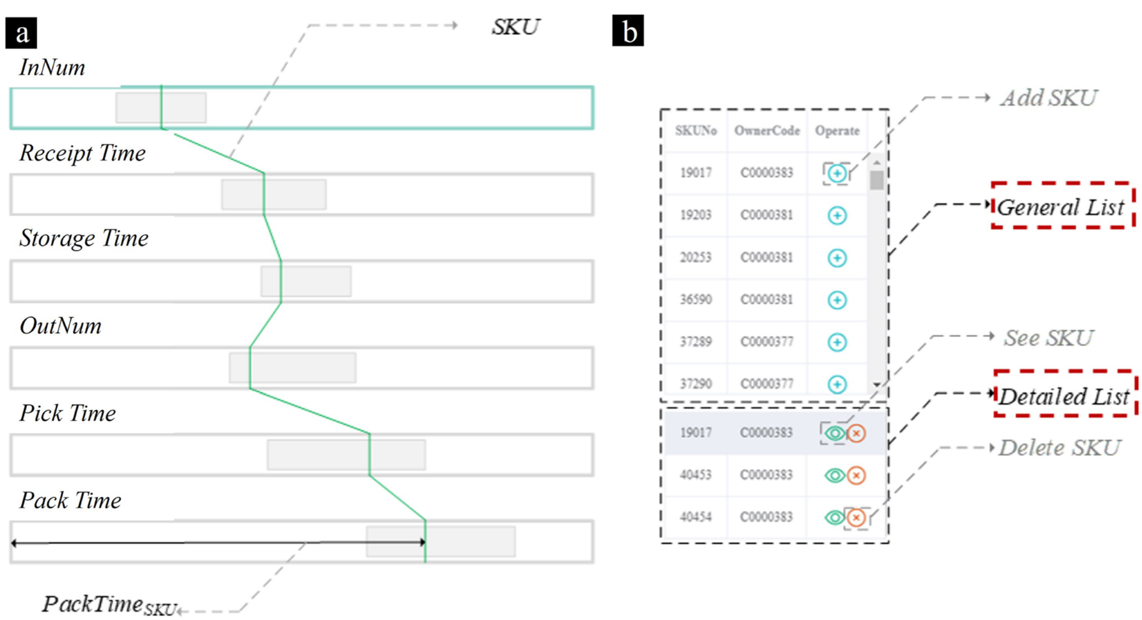

Fig. 7 The layout and visual design of the *SKU Profiles View*. **a** The parallel rectangular view shows the distribution pattern of goods attributes. **b** The goods list view consists of a general list and a detailed list

SKU icon to view the inventory changes of goods in the *Turnover Timeline View* and remove the target goods from the detailed list by clicking the *See SKU* icon.

4.4 Picking mode view

The *Picking Mode View* (Fig. 2f) is composed of two components: the mode circle view to reveal the frequent outbound mode of key goods (i.e., goods with large inventory fluctuations, high outbound frequency, or abnormal picking time, etc.) and the nested rectangle view to show the outbound frequency of key goods, owner, and warehouse. This view gives support to the discovery of abnormal outbound modes of goods, providing a reference for storage location adjustment *(T5)*.

The mode circle view consists of three rows of mode units (Fig. 8a), which can present the outbound mode of three goods at the same time. For example, in Fig. 2f, from the first row to the third row it shows the outbound mode of goods 11121, 21270, and 27049, respectively. Take goods 11121 as an example, its outbound mode refers to a frequent outbound combination containing goods 11121 (denoted as *KeyGoods*) and other goods (denoted as *OtherGoods*). As shown in Fig. 8a, the upper of the mode unit consists of the inner rings, solid circle, and the outer arc. The radius of the inner rings (i.e., $PickTime_{KeyGoods}$ and $PickTime_{OtherGoods}$) shows the individual picking time of different goods in the combination and the radius of solid circle $PickTime_{Mode}$ represents the picking time of the combination. For example, suppose an outbound combination of goods A is A, B, and C in Fig. 8b. The radius R_A, R_B, and R_C are the individual picking time of goods A, B, and C, respectively, while radius R_{Mode} is the picking time of the outbound combination. The radian of the outer arc $\theta_{frequency}$ in Fig. 8c encodes the outbound frequency of an outbound combination. The lower part of the mode unit is rectangles arranged horizontally in Fig. 8a, with a rectangle representing a type of goods in the outbound combination. The orange rectangle indicates the key goods (i.e., denoted as *KeyGoods*) selected from the detailed list in the *SKU Profiles View* for outbound mode analysis, and the green ones represent the remaining goods (i.e., denoted as *OtherGoods*) delivered with the key goods in the combination.

The nested rectangle view (Fig. 8d) displays the outbound frequency proportion of the key goods. The three nested rectangles from inside to outside $WaveNum_{unit}$, $WaveNum_{owner}$, and $WaveNum_{mon}$ are, respectively, the outbound frequency of the key goods, owner, and warehouse. At the top of Fig. 2f, there is the legend description of the mode circle view and the nested rectangle view.

Design alternatives. During the design iteration, initially, we used the radius of different rings to represent the picking time of the outbound combination and the individual picking time of goods in the combination. The color of these rings indicated the combination and different goods. However, our experts pointed out that too many colors caused severe visual clutters. Thus, we used the solid circle to represent the outbound combination while the ring for the goods in the combination. All rings had only two colors to represent the individual picking time of *KeyGoods* and *OtherGoods*, respectively. Users could hover over any rectangle below the mode unit to know the specific picking time of *KeyGoods* and *OtherGoods*.

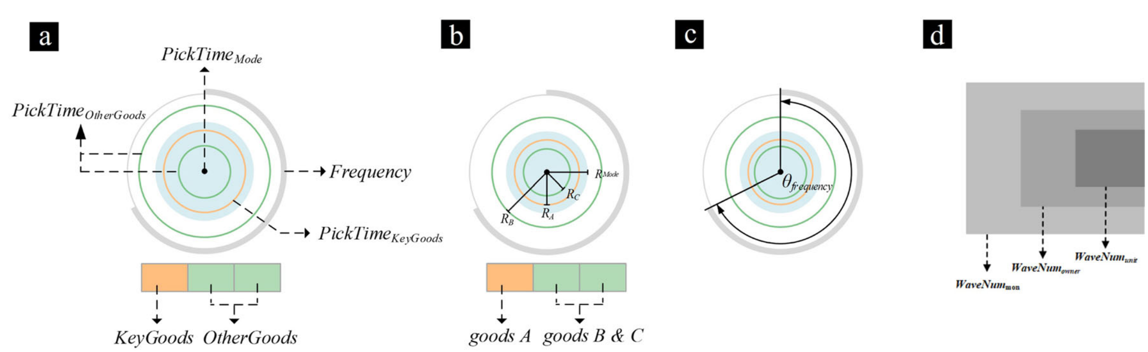

Fig. 8 The visual design of the *Picking Mode View*. All black auxiliary lines and text are additional descriptions to explain the visual design. **a** The mode unit contains the outer arc, the inner rings, and a solid circle, reflecting the outbound frequency and the picking time of the outbound combination. **b** An outbound combination of goods A, containing goods A, B, and C. The radius R_{Mode} is the picking time of the combination, while radius R_A, R_B, and R_C are the individual picking time of goods A, B, and C, respectively. **c** The radian of the outer arc encodes the outbound frequency of the outbound combination. **d** The nested rectangles from inside to outside represent the outbound frequency of goods, owner and warehouse, respectively

Interaction. User can hover over any model unit to view the detailed information of the outbound combination, including goods composition and picking time for the outbound combination. In addition, when hovering over any rectangle below the mode unit, the tooltip shows the individual picking time of *KeyGoods* and *OtherGoods* in the combination. Users can also click the green rectangle to choose *OtherGoods* as the key goods. After that, users can view these new key goods' information in the detailed list in the *SKU Profiles View*. When users hover over the nested rectangle, a tooltip will pop up showing detailed information about the date, and the respective outbound frequency of the key goods, owner, and warehouse.

4.5 Other views

The system includes two auxiliary visualizations, the *Warehouse Overview* (Fig. 2a) and the *Turnover Timeline View* (Fig. 2e). The former can assist users in viewing the overview of the warehouse operation state each month *(T1)*, and the latter can help users to view the changing pattern of goods inventory in the selected month *(T6)*.

Warehouse Overview. It consists of the toggle button and the state table. The toggle button is set to switch the warehouse state, inbound or outbound. The state table has five attributes: date, order quantity, goods quantity, type, and execution time. The horizontal bars in the table represent the corresponding attribute values. Users can click any row to view the warehouse state in a specific month, and the remaining views will update simultaneously.

Turnover Timeline View. This view contains three line charts arranged vertically. From top to bottom, they are the daily inventory of goods, the number of goods outbound, and the number of goods inbound. In each line chart, the blue dot represents the goods selected from the *SKU Profiles View*, and the orange encodes the goods' owner.

5 Evaluation

Inspired by the evaluation method in visual analytics studies (Wang et al 2022b; Li et al. 2022; Fu et al 2018), we demonstrate the usefulness and effectiveness of WarehouseLens with three case studies in Sect. 5.1 and expert interviews in Sect. 5.2.

5.1 Case study

We invite our experts (denoted as E1, E2, and E3) mentioned in Sect. 3.2 to explore WarehouseLens (details of the specific process will be introduced in Sect. 5.2). In this section, we describe how the experts use our system to explore and gain insights into the data described in Sect. 3.1, concluding several cases found by our experts and formulating them into three case studies to fully demonstrate the system.

5.1.1 Case 1: the analysis of warehouse operation state

In this case, we describe how E1 analyzes the operation state of the warehouse and finds interesting phenomena. He is curious about whether the warehouse operation state is different before the promotion period.

In the initial *Warehouse Overview* (Fig. 2a), E1 found that the number of inbound goods in October was the largest. He explained that due to the Double Eleven promotion (Bai et al 2022) in China, all warehouses needed to prepare goods in advance to cope with the huge volume of customer orders. Therefore, he focused on the inbound activities in October particularly. From the calendar overview in Fig. 2b3, E1 observed that the state hexagons in the third row had a significantly longer receipt time than the rest rows. Thus, E1 chose the state indicator *T(Receipt)* to update trend triangles and noticed that the receipt time reached a small peak of 3 h and 14 min on Saturday (i.e., 17th October). E1 further checked the detailed information on October 17 in the hexagon analysis view (Fig. 2b2). From the radar chart, E1 found that the average receipt time in October is 1 h and 55 min, significantly less than the time on the 17th. He then inspected the stacked bullet chart and noticed that the number, weight, and volume of inbound goods on that day exceed the monthly mean. Further, E1 wanted to check if the special case of October 17 continued on October 18. In other words, he needed to determine if the receipt time on the 18th was still much longer than the average for the month. He double-clicked the state hexagons representing the 17th and 18th and add these two days to the tri-bar comparison view in Fig. 2b1. Compared with the 17th, the receipt time on the 18th was significantly shorter, which was close to the monthly mean, indicating that the operation state on the 18th was normal. After consulting warehouse managers, E1 speculated that the longer receipt time on the 17th was due to the complexity and large volume of inbound goods rather than other unexpected circumstances. In other words, warehouse managers need not make production adjustments. E1 continued to view the *Event Circle View* (Fig. 2c), hoping to understand the periodicity pattern of warehouse operation. He easily found that the frequency of packing events was much higher from 8 o'clock to 14 o'clock. E1 believed that in this case, warehouse managers must consider arranging more production resources, such as workers and equipment, within this fixed period.

Overall, after analyzing the trend and periodicity pattern of goods turnover events in the warehouse, E1 can have a good understanding of the overall operation state of the warehouse, which can provide suggestions for warehouse managers to develop a reasonable management strategy.

5.1.2 *Case 2: the discovery of system anomalies*

Various factors can lead to warehouse anomalies, such as insufficient production or human resources, equipment failure, etc. Warehouse managers need to detect anomalies in time and take quick countermeasures to ensure the orderly operation of the warehouse. E2 aims to check if there is an anomaly in the goods turnover event and what is the cause of the anomaly.

E2 first chose the state indicator *T(Pick)* to initialize the *State Calendar View* in Fig. 9c and found that the picking time on October 18 was much longer, which caught the attention of E2. Moreover, from the tri-bar comparison view in Fig. 9a, E2 observed that compared with the picking time on the 17th of only 38 min, the time on the 18th was 1 h and 21 min, almost twice the monthly mean in November of 43 min. E2 wanted to what caused the long picking time and then explored the detailed information on the 18th in the hexagonal analysis view (Fig. 9b). From the bottom stacked bullet chart, she found that the values of the four goods attributes were in the normal range, not exceeding the monthly mean. Thus, E2 presumed that goods attributes were not the reason for the abnormal picking time. According to E2's management experience, she pointed out that the picking time was only related to goods attributes or the human resources of the warehouse. After confirming with the human resources department, it verified that as of the 18th, the warehouse company had not renewed its contract with the labor company, resulting in a shortage of workers and extended picking time. At last, E2 concluded that warehouse managers must promptly identify the causes of such anomalies and take timely countermeasures to avoid a similar situation.

In summary, E2 believes that WarehouseLens can assist her in detecting, analyzing, and interpreting warehouse anomalies to some extent, from which she can gain insights to handle anomalies.

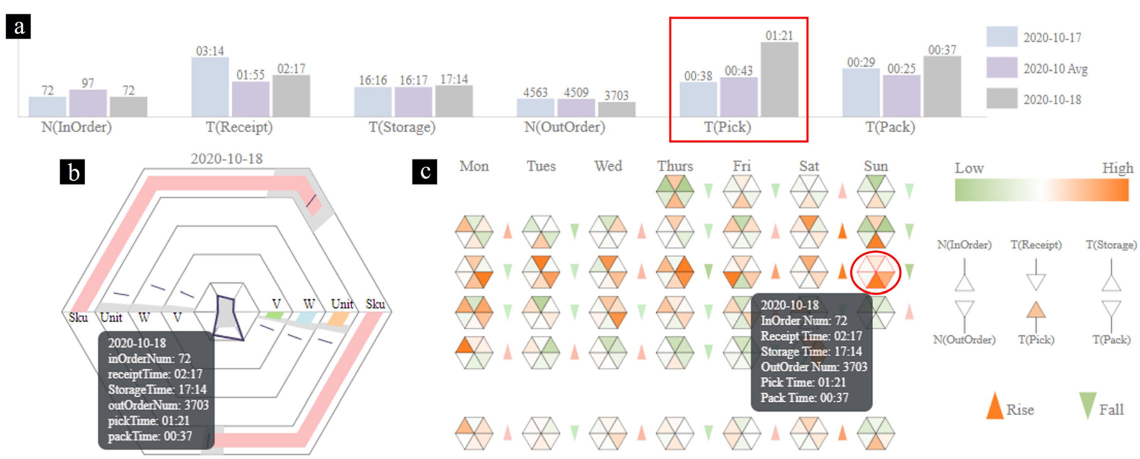

Fig. 9 The anomaly discovery in October. **a** Comparison of operation state on different dates. The picking time on October 18 is almost twice the monthly mean. **b** Inbound and outbound details of goods on October 18. **c** The changing trend of the operation state indicator *T(Pick)* in October

5.1.3 *Case 3: the exploration of goods outbound mode*

Analyzing the outbound mode of goods can provide a reference for storage location adjustment, thereby improving the efficiency of picking events. In this case, we describe how E3 discovers key goods and explores the turnover patterns of these goods.

From Fig. 10a1, E3 noticed that the order execution time in November exceeded 12 h, resulting in a heavy warehouse operation load, which is an unusual situation. To understand the turnover patterns of goods in a usual situation, E3 chose December for analysis (Fig. 10a2), with approximately 9 h for order execution. He selected *OutNum* (Fig. 10b1) and *Pick Time* (Fig. 10b2) as the sorting attribute in the *SKU Profiles View*, respectively, and checked the distribution pattern of the whole goods. As depicted in Fig. 10b1, E3 found that the outbound number was not significantly related to other goods attributes according to the mixed colors of lines. While in Fig. 10b2, he noticed that the goods with longer picking time also spent much more time packing. To find goods worthy of analysis, E3 brushed the longer picking time range on the right part of the parallel rectangle with all the chosen goods displayed in the general list (Fig. 10c). Then, he further added goods 24242 and 59630 belonging to different owners to the detailed list for the outbound mode analysis.

In Fig. 11a1, E3 observed that three outbound modes of goods 24242 in the first row were all normal because the combinations' picking time were close to the individual time of goods in the combinations. He continued to check goods 59630, and its first outbound mode caught his attention. The picking time of this combination was significantly longer than the individual picking time of the two goods that make up the combination. Moreover, by looking at the nested rectangle view (Fig. 11a2, the outbound frequency of goods 59630 was 301 times, approximately half of its owner with 648 times. E3 speculated that goods 59630 was the key goods worthy of further exploration of its turnover patterns. E3 further checked the three outbound modes of goods 59630 in Fig. 11b1: mode A with goods 59630 and 47927, mode B with goods 59630 and 57131, and mode C with goods 59630, 57131, and 47927. Seeing the picking time statistics of goods 59630, 47927, and 57131, respectively, in Fig. 11b2, E3 noticed that when the goods 59630 and

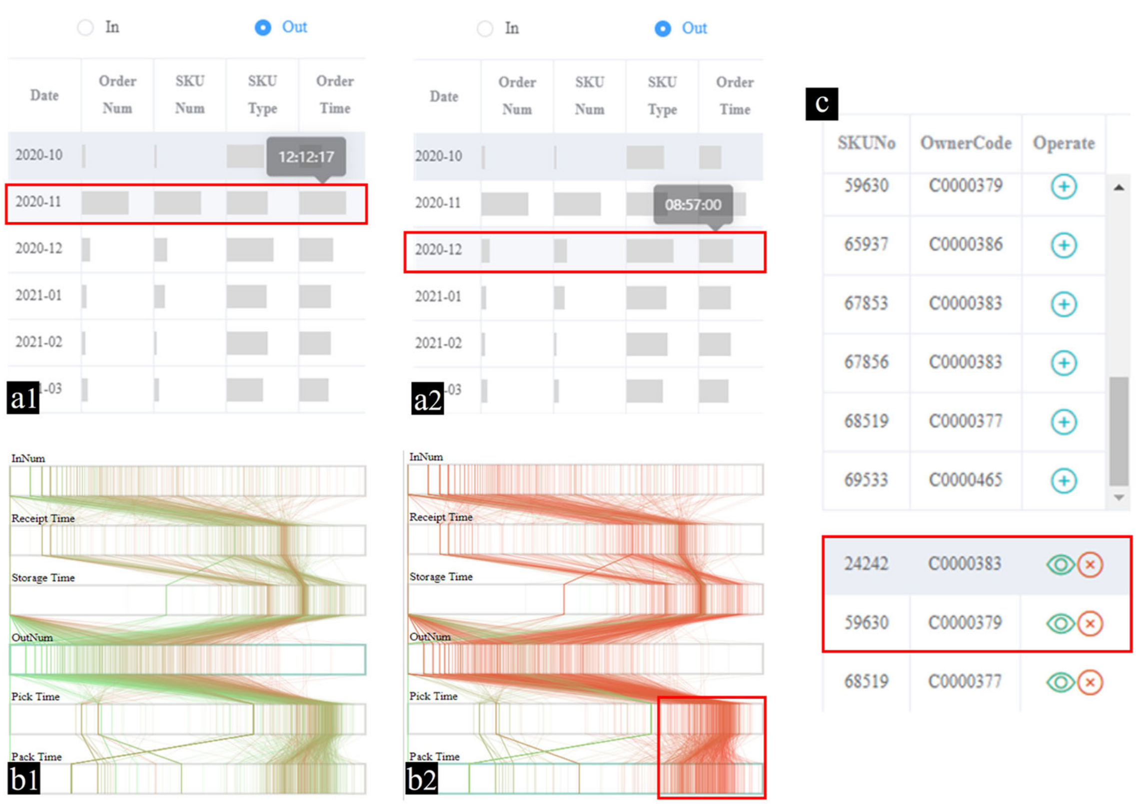

Fig. 10 Filter key goods through the *SKU Profiles View*. Outbound information in November (**a1**) and December (**a2**). Select *OutNum* (**b1**) and *Pick Time* (**b2**) as the sorting attribute, respectively, in the *SKU Profiles View*. **c** Goods with long picking time selected from (**b2**)

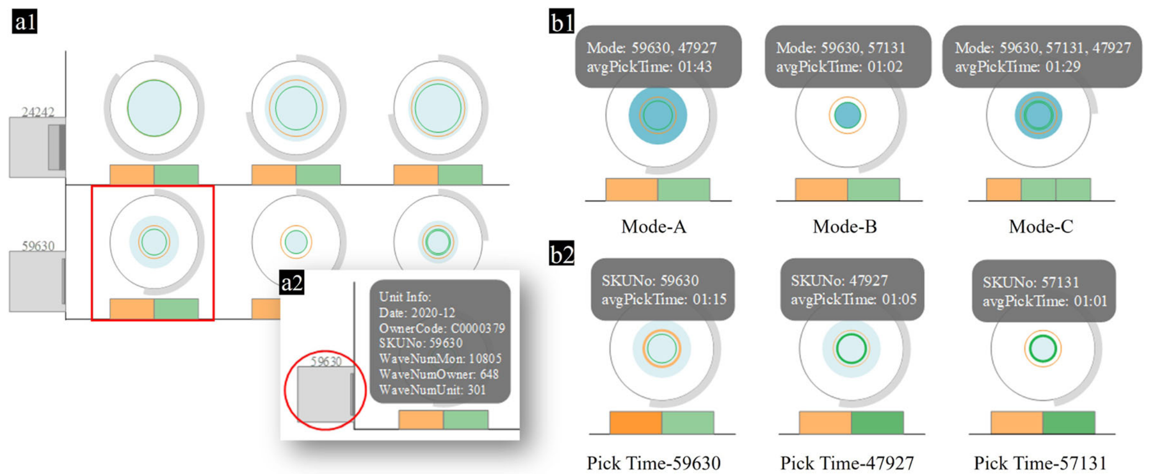

Fig. 11 **a1** The frequent outbound modes of goods 24242 and 59630. **a2** The outbound frequency of goods 50630 with 301, which is approximately half of its owner with 648. **b1** The picking time of three outbound modes for goods 50630. **b2** The individual picking time of goods 59630, 47927, and 57131

47927 were outbound simultaneously, like modes A and C, the picking time of the combinations were higher. While for mode B, the combination's picking time was normal. E3 speculated that was because the goods 59630 and 47927 were far apart, while they were frequently outbound in the same batch, resulting in a longer picking time for the combination. To verify this speculation, E3 inspected the outbound mode of goods 59630 in the next month and still obtained the same rule. Thus, E3 believed that warehouse managers should adjust the storage location of goods 59630 and 47927, shorting their combinations' picking time and thereby improving outbound efficiency.

In summary, E3 appreciates the system's ability to gain deep insights into the frequent outbound mode of key goods, providing guidance for optimizing warehouse location management.

5.2 Expert interview

We conduct in-depth interviews with our collaborating domain experts and collect their feedback and comments.

5.2.1 Participants

We interview 10 experts (3 females and 7 males, age_{mean}=29.5, age_{sd}=2.94) from our collaborating company. Three of them are experts (E1-E3) involved in our design process in Sect. 3.2. The other seven experts (E4-E10) are new, who use WarehouseLens for the first time. E4 is the product manager with six-year working experience. Both E5 and E6 are managers of different warehouses. The rest (E7-E10) are senior engineers who work on developing big data management products for warehouses with 4 to 7 years of working experience. All experts have little experience with visualization techniques.

Table 1 The tasks designed for experts to guide the exploration

	Task
Task 1	Observe the changes in operation state indicators and compare the operation state on different dates
Task 2	Check the temporal pattern of six goods turnover events and identify key time points
Task 3	Explore the distribution pattern of the goods attributes and select the goods of interest for further analysis of turnover patterns
Task 4	Compare high-frequency outbound modes of key goods and judge whether the picking time is normal

Table 2 User questionnaire. Q1-Q5 are designed to evaluate the usefulness of WarehouseLens and Q6-Q10 focus on the system's effectiveness in analyzing warehouse operation state and goods turnover patterns

Q1	It is easy to learn the system
Q2	It is easy to use the system
Q3	It is easy to understand the visual design of the system
Q4	I am willing to use the system in digital warehouse analysis scenarios
Q5	I will recommend the system to other warehouse managers
Q6	It is easy to discover the changing trends in warehouse state indicators
Q7	It is easy to summarize the periodicity pattern of goods turnover events
Q8	It is easy to observe warehouse anomalies
Q9	It is easy to filter key goods of interest
Q10	It is easy to analyze abnormal outbound modes of different goods

5.2.2 Procedure

The interview is conducted using a remote conferencing software, where participants use their own computers to access WarehouseLens deployed on a cloud server. Each interview consists of four sections and lasts for approximately 60 min. The data we use for the interview are identical to those in Sect. 3.1. First, we spend 5 min briefly introducing the project background. Second, we use a comprehensive example to demonstrate the visual encodings and interactions of the system (20 min). Third, experts are asked to freely use the system for 25 min. We design the tasks (see Table 1) to guide their open-ended exploration. Experts are encouraged to think aloud and speak out about whatever they are thinking and doing during their exploration. Finally, we ask the experts to finish a post-study questionnaire to evaluate our system on visual coding, interactions, and system function (10 min). In our questionnaire, we adopt the bipolar survey design (Li et al 2021) with negative statements at the left end of the five-point Likert-scale (Likert 1932) (1–5 represents "strongly disagree" to "strongly agree" for each statement). All questions are listed in Table 2. We also take notes about their feedback and record their comments during the interview.

5.2.3 Results

In the following, we report our results from the expert interview, including participants' task performance, questionnaire response, and qualitative feedback to the system. Figure 12 shows participants' ratings on their impression of WarehouseLens on the post-study questionnaire. Overall, both the usefulness and the effectiveness of our system are appreciated by participants.

Task performance. All participants could successfully complete the exploration of our system. No participant gave up in the middle of the tasks. Below, we summarize users' performance from the four tasks.

- For *Task 1*, all participants would first click on the funnel legend to update the trend triangle in the *State Calendar View*. The three most frequently clicked state indicators were *T(Pick)*, *T(Pack)*, and *N(OutOrder)*. E5 explained that the picking event was the most time-consuming and labor-intensive in warehouse operations, which was the focus of warehouse management optimization. If two or more consecutive trend triangles showed an upward or downward trend, on average, participants would inspect 3.1 ($\delta = 0.66$, use δ for standard deviation in the full text) state hexagons during this period. At the same time, they tended to add these days to the tri-bar comparison view to perform inter-comparisons to better understand the operation state. Eight participants tried to explain the operation state by examining the goods attributes day by day in the hexagon analysis view; further half of the eight participants preferred to carry out a more in-depth examination in the *SKU Profiles View*.
- When participants selected the state indicator *T(Pick)* in *Task 1*, next, they would check the periodic pattern of the picking event in the *Circle Event View* to complete *Task 2*. Seven participants preferred to first inspect the circular heat map in the circumferential direction (i.e., the single ring) and then in the radial direction. As E4 said, "*Users' purchasing behavior tend to be cyclical, such as Friday and Monday near the weekend*." In the radial direction, nine of the participants tended to view the outermost circle first. "*I can obtain an overview of the frequency distribution of the six turnover events over a 24-hour period through the outermost circle*." We observed that all experts were willing to look at the text description above the circular heat map to locate the key circular grid quickly.
- Further, for *Task 3*, the three most selected sorting attributes by participants in the *SKU Profiles View* were *OutNum*, *Pick Time*, and *Pack Time*. E5 and E6 mentioned that goods with high outbound frequency and longer picking time were the key objects of warehouse management. Participants usually

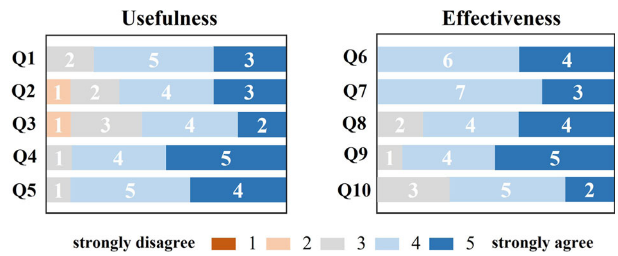

Fig. 12 Results of our questionnaire. 1–5 with 1 as the most negative and 5 as the most positive. The white number on the ribbon indicates the number of participants

added 15 to 20 goods to the general list by brushing the lines in the parallel rectangle. Then, check the outbound modes of these goods one by one in the *Picking Mode View*.

- Regarding *Task 4*, we noted that in order to check the outbound mode of each type of goods, participants tended to conduct at least two queries about its picking time. E8 explained that he needed to compare the outbound modes of the same goods for at least two consecutive months. In addition, eight of the participants would like to check the frequent outbound mode of all goods in an outbound combination by clicking on the rectangle below the mode unit. As E10 stated that he wanted to determine if the picking time of the combination would affect the individual picking time of the goods.

Overall, most participants preferred to analyze the warehouse operation at the overall level through the *State Calendar View* and the *Circle Event View* first. At the same time, the *SKU Profiles View* and the *Picking Mode View* may also be used to assist in explaining certain phenomena in the warehouse operation. The most used view was the *State Calendar View*; and the *SKU Profiles View* and the *Picking Mode View* were closely related and used in parallel. The least used view was the *Turnover Timeline View*, which was just an auxiliary view.

Questionnaire response. Fig. 12 shows experts' ratings on their impression of WarehouseLens on the post-study questionnaire. Overall, WarehouseLens was perceived easy to learn ($Q1$), easy to use ($Q2$), and easy to understand ($Q3$), with mean scores of 4.1, 3.8, and 3.7, respectively. However, one participant rated 2 points on $Q2$ and $Q3$. He stated that the *Picking Mode View* was a little complex and takes some effort to master, which affected his acceptance to WarehouseLens. Participants also felt that the system was helpful ($Q4$, $Q5$), with mean scores of 4.4 and 4.3. Participants had positive opinions on the *State Calendar View* and considered it useful ($Q6$), with a mean score of 4.4 ($\delta = 0.48$). They also preferred the *Circle Event View* and thought it could be mastered quickly to summarize the periodicity pattern of turnover events ($Q7$), with a mean score of 4.3 ($\delta = 1.45$). Two participants rated $Q8$ lower because it took them long to get the desired results. For the parallel rectangles in the *SKU Profiles View*, they overall believed that the improvement over the traditional view was effective, which helped to discover the distribution pattern of goods ($Q9$), with a mean score of 4.4 ($\delta = 0.66$). For the analysis of the outbound mode, the *Picking Mode View* was perceived useful in general ($Q10$), but several mentioned that they could not get started quickly.

Qualitative feedback. Our interview mainly focused on collecting participants' feedback on all the aspects of WarehouseLens in order to assess our system qualitatively.

- *System usefulness.* All participants believed that they could use the system easily after a brief tutorial. WarehouseLens enabled them to quickly analyze the goods turnover data by showing information from different perspectives with a user-friendly interface. E2 commented that "*In the past, we have not tried to analyze turnover events to assess the warehouse operation state. It's impressive that we can explore our data in a new way that is more effective.*" They also liked the visualizations adopted, which help them find new insights. As E1 said: "*I am impressed by the State Calendar View. This view provides an overview of the operation state in the warehouse for me and allows me to view detailed information further.*" We found that participants had different opinions on the *SKU Profiles View* and the *Picking Mode View*. Both E5 and E6 stated that the two views were easy to understand but took some effort to master. E7 commented: "*There are many interactions in the two views, which increase learning costs for users.*" While four participants appreciated the clear workflow between the two views. As E3 said: "*I can find some interesting distribution patterns of goods and select target goods in the SKU Profiles View. Then, further, analyze the corresponding outbound modes of these goods in the Picking Mode View.*" Also, E4 said that "*Although it takes some time to analyze the outbound mode of goods skillfully, insights* *gained from the analysis really instruct me to discover some abnormal outbound modes for some key goods.*"

- *System effectiveness.* All participants considered WarehouseLens as a comprehensive visual analytics system that could successfully fulfill all the design tasks. The seven new experts (E4-E7) could understand most of the visual designs after the demonstration phase and finish the task-driven exploration within the allotted time. Specifically, they regarded the *State Calendar View* and the *Event Circle View* as useful for discovering trend changes and periodic patterns. E8 commented that "*The trend triangle vividly reflects the changes of different state indicators in the current month, which is helpful for me to compare the operation state on different dates.*" E4 agreed, "*I like the labels and legends in the Event Circle View, and the text description above the circular heat map helps me quickly locate key periods* [...] *I observe that the picking event often occurs in the morning while the packing event lasts throughout the whole day.*" Also, E6 said that "*I tend to find abnormal warehouse operation state through the calendar overview in the State Calendar View and get the explanation for the anomaly from the hexagon analysis view directly or further inspect the SKU Profiles View.*" As for the parallel rectangles designed in the *SKU Profiles View*, we observed that participants liked this improvement and tended to choose different sorting attributes to observe the distribution pattern of goods. E9 commented that "*It is easy to filter the outbound goods with large quantities and long picking time in the SKU Profiles View and check whether the outbound modes of these goods are normal in the Picking Mode View.*" However, E10 felt that lines in the parallel rectangles may cause a little visual confusion, as E10 said: "*During the promotion period, there is a large amount of inbound and outbound goods. Thus there may be a little messy due to much line crossing in the SKU Profiles View.*" As for the outbound mode analysis of goods, most participants approved the design of the mode unit. E5 stated that "*Through the mode unit, I can know the picking time of one outbound combination and the individual picking time of goods in the combination at the same time. And this comparison is intuitive enough.*" While E7 commented that "*In some cases, the radius of the rings and solid circle are too small to see clearly, which need to adjust the scale of data mapping.*"
- *Improvement.* Participants also put forward some valuable suggestions for improving the performance of the prototype system. First, E2 suggested that it would be better to support more days of warehouse operation state comparison instead of only two days in the *State Calendar View*. E3 focused on the analysis of goods turnover patterns and figured out that, "*The system can only allow the selection of three key goods in the SKU Profiles View. It should support the analysis of the outbound mode of more goods at the same time.*" Additionally, E4 suggested adding a search function to the *SKU Profiles View* and explained that "*In some cases, it is necessary to check the outbound mode of certain goods for two consecutive months. If the user can manually enter the code of target goods, the search efficiency can be improved a lot.*" As for the *Picking Mode View*, both E5 and E6 hoped to show more outbound combinations of each goods instead of only three combinations. As E6 suggested that "*Alternatively, the system can consider allowing users to manually adjust the number of combinations they want to check.*" E8 commented that "*The Turnover Timeline View is only an auxiliary view and it is a little space-wasting, which can be improved to reflect richer information.*" In addition, E9 suggested to swap the position of the *Turnover Timeline View* and the *Picking Mode View*. As E9 mentioned that "*Users usually need to look back and forth between these two views. Therefore, it will be more convenient to observe if the two views are placed next to each other.*"

6 Discussion

In this section, we first summarize lessons learned from our design study and then, discuss the generalizations and limitations of WarehouseLens.

Lessons learned. Collaborations with domain experts during the whole process provide us with valuable experiences in developing WarehouseLens. For requirement analysis, experts are mostly not familiar with visual analytics and cannot clearly describe their requirements. Thus, iterative interviews are necessary to summarize domain requirements and analytical tasks. For visual design, designing an overview of the warehouse operation state need to consider the cognitive preferences of experts. As such, we adopt a calendar layout to present the daily operation state of the warehouse, which allows users easily to compare operational performance differences on different dates. Moreover, considering the characteristics of the outbound mode data in Sect. 3.1, we design a customed mode circle to better compare the respective picking time of different goods in the outbound combinations.

Generalizability. Although WarehouseLens is presented to analyze goods turnover events in the digital warehouse, it can be applied to other scenarios. First, our prototype system can be improved to visualize and analyze data from multiple warehouses, not limited to the ones we are currently working with. The main effort required is to pre-process the raw data into the same format. For large-scale warehouses, the system can be applied with slight changes. For example, there may be millions of inbound and outbound goods each month and we may set thresholds to filter more worthy goods for analysis in the *SKU Profiles View*. Depending on the job requirements of the warehouse manager, they tend to view different warehouses one by one. Therefore, the new system will still present the operation state and goods turnover pattern of the individual warehouse. But there will be a drop-down box for warehouse managers to select the warehouse they want to inspect. Second, the temporal visualization method based on a calendar diagram can be used to monitor multiple phases of the assembly line in the factory, discover potential anomalies and analyze the causes of anomalies. Third, the improved circular heat map is also applicable to the discovery of periodic patterns in other time-series data, e.g., web clickstream data.

Limitation. The case studies and expert interviews demonstrate the usefulness and effectiveness of WarehouseLens. However, there are still limitations to our work.

(1) Anomaly analysis. The system only supports the verification of the relationship between goods attributes and warehouse anomalies, without production resource data combined for thoughtful analysis, such as worker and equipment data, etc.

(2) Goods outbound modes analysis. Our work only supports finding the abnormal outbound mode of goods, which lacks location data of goods to carry out an in-depth analysis of the location of the goods and is unable to guide storage location adjustment effectively.

(3)Data scope. The goods data used in our work are mostly toiletries, while there are various types of goods in the actual scene, such as clothing, food, and electrical appliances. Therefore, a complete solution is to consider category data of goods in the warehouse anomaly analysis. For example, the order execution time with bulky goods like electrical appliances and furniture is usually longer, which should be considered a normal operation state. Hence, introducing category data of goods can assist warehouse managers in better analyzing the cause of the anomaly.

(4) Evaluation. We have presented three case studies and conducted interviews with five domain experts, which illustrate the ability of WarehouseLens in supporting optimal warehouse management. However, we believe that it would be better to recruit more professional experts with practical experience in warehouse management to validate the effectiveness of the system workflow.

7 Conclusion

In this paper, we present an interactive system, WarehouseLens, for the visual analysis of goods turnover data. The system integrates all processes of goods turnover to describe the workflow of warehouse operation. Novel views and multiple interactions support the understanding of the warehouse operation state (e.g., the trend and periodicity pattern) and the exploration of goods turnover patterns (e.g., the outbound mode and inventory change). Our case studies and expert interviews reveal the usefulness and effectiveness of the system. In future work, we plan to improve the system's usefulness by considering the analysis of production resource data, goods location and category data, further realizing the all-around supervision of warehouse operation state and the in-depth exploration of goods turnover patterns.

Acknowledgements This research is partially supported by the School-City Cooperation Special Fund Project (2020CDSN-02) and School-City Strategic Cooperation Project (2021CDSN-13). We would like to thank the industry sponsor Sichuan-wulianyida Technology Co., Ltd. for providing with the data.

References

Bai J, Li Z, Lin Y (2022) The application analysis of game theory on double eleven. In: 2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022), Atlantis Press, pp 377–382, https://doi.org/10.2991/aebmr.k.220307.060

Bräuer P, Mazarakis A (2020) Visualization of turnover rate in a warehouse using augmented reality: a demo with the microsoft hololens. In: Alt F, Schneegass S, Hornecker E (eds) Mensch und Computer 2020 - Tagungsband, Magdebug, Germany, September 6-9, 2020. ACM, pp 519–522, https://doi.org/10.1145/3404983.3410422

Cogo E, Žunić E, Beširević A, et al (2020) Position based visualization of real world warehouse data in a smart warehouse management system. In: 2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH), IEEE, pp 1–6, https://doi.org/10.1109/INFOTEH48170.2020.9066323

De Koster R, Le-Duc T, Roodbergen KJ (2007) Design and control of warehouse order picking: a literature review. Eur. J. Oper. Res. 182(2):481–501. https://doi.org/10.1016/j.ejor.2006.07.009

Deng Z, Weng D, Liu S, Tian Y, Xu M, Wu Y (2023) A survey of urban visual analytics: advances and future directions. Comput Vis Media 9(1):3–39. https://doi.org/10.1007/s41095-022-0275-7

Du F, Plaisant C, Spring N, et al (2016) Eventaction: Visual analytics for temporal event sequence recommendation. In: 2016 IEEE Conference on Visual Analytics Science and Technology (VAST), pp 61–70, https://doi.org/10.1109/VAST.2016.7883512

Fang W, Zheng S, Liu Z (2019) A scalable and long-term wearable augmented reality system for order picking. In: 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp 4–7, https://doi.org/10.1109/ISMAR-Adjunct.2019.00016

Fu S, Dong H, Cui W et al (2018) How do ancestral traits shape family trees over generations? IEEE Trans Vis Comput Graph 24(1):205–214. https://doi.org/10.1109/TVCG.2017.2744080

Gu J, Goetschalckx M, McGinnis LF (2007) Research on warehouse operation: a comprehensive review. Eur J Oper Res 177(1):1–21. https://doi.org/10.1016/j.ejor.2006.02.025

Guo X, Yu Y, Koster RBD (2016) Impact of required storage space on storage policy performance in a unit-load warehouse. Int J Prod Res 54(8):2405–2418. https://doi.org/10.1080/00207543.2015.1083624

Guo Y, Guo S, Jin Z et al (2022) Survey on visual analysis of event sequence data. IEEE Trans Vis Comput Graph 28(12):5091–5112. https://doi.org/10.1109/TVCG.2021.3100413

Guo S, Jin Z, Chen Q, et al (2019) Visual anomaly detection in event sequence data. In: 2019 IEEE International Conference on Big Data (IEEE BigData), Los Angeles, CA, USA, December 9-12, 2019. IEEE, pp 1125–1130, https://doi.org/10.1109/BigData47090.2019.9005687

Han J, Pei J, Yin Y (2000) Mining frequent patterns without candidate generation. ACM SIGMOD Rec 29(2):1–12. https://doi.org/10.1145/342009.335372

Hou Y, Wang C, Wang J et al (2022) Visual evaluation for autonomous driving. IEEE Trans Vis Comput Graph 28(1):1030–1039. https://doi.org/10.1109/TVCG.2021.3114777

Jaghbeer Y, Hanson R, Johansson MI (2020) Automated order picking systems and the links between design and performance: a systematic literature review. Int J Prod Res 58(15):4489–4505. https://doi.org/10.1080/00207543.2020.1788734

Jin Z, Cui S, Guo S et al (2020) Carepre: an intelligent clinical decision assistance system. ACM Trans Comput Heal 1(1):1–20. https://doi.org/10.1145/3344258

Jo J, Huh J, Park J et al (2014) Livegantt: interactively visualizing a large manufacturing schedule. IEEE Trans Vis Comput Graph 20(12):2329–2338. https://doi.org/10.1109/TVCG.2014.2346454

Krstajic M, Bertini E, Keim DA (2011) Cloudlines: compact display of event episodes in multiple time-series. IEEE Trans Vis Comput Graph 17(12):2432–2439. https://doi.org/10.1109/TVCG.2011.179

Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29(1):1–27

Lee CKM, Lv Y, Ng KKH et al (2018) Design and application of internet of things-based warehouse management system for smart logistics. Int J Prod Res 56(8):2753–2768. https://doi.org/10.1080/00207543.2017.1394592

Li C, Cao M, Wen X, Zhu H, Liu S, Zhang X, Zhu M (2022) MDIVis: visual analytics of multiple destination images on tourism user generated content. Vis Inform 6(3):1–10. https://doi.org/10.1016/j.visinf.2022.06.001

Likert R (1932) A technique for the measurement of attitudes. Arch Psychol 22(140):1–55

Lin Y, Wong K, Wang Y et al (2021) Taxthemis: interactive mining and exploration of suspicious tax evasion groups. IEEE Trans Vis Comput Graph 27(2):849–859. https://doi.org/10.1109/TVCG.2020.3030370

Li H, Xu M, Wang Y, et al (2021) A visual analytics approach to facilitate the proctoring of online exams. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp 1–17, https://doi.org/10.1145/3411764.3445294

Mei H, Eisner J (2017) The neural hawkes process: A neurally self-modulating multivariate point process. In: Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA, pp 6754–6764

Mu X, Xu K, Chen Q, et al (2019) Moocad: Visual analysis of anomalous learning activities in massive open online courses. In: EuroVis (Short Papers), pp 91–95, https://doi.org/10.2312/evs.20191176

Nguyen PH, Henkin R, Chen S et al (2020) VASABI: hierarchical user profiles for interactive visual user behaviour analytics. IEEE Trans Vis Comput Graph 26(1):77–86. https://doi.org/10.1109/TVCG.2019.2934609

Pan JCH, Shih PH, Wu MH et al (2015) A storage assignment heuristic method based on genetic algorithm for a pick-and-pass warehousing system. Comput Ind Eng 81:1–13. https://doi.org/10.1016/j.cie.2014.12.010

Perer A, Gotz D (2013) Data-driven exploration of care plans for patients. In: Mackay WE, Brewster SA, Bødker S (eds) 2013 ACM SIGCHI Conference on Human Factors in Computing Systems, CHI '13, Paris, France, April 27 - May 2, 2013, Extended Abstracts. ACM, pp 439–444, https://doi.org/10.1145/2468356.2468434

Perer A, Wang F (2014) Frequence: Interactive mining and visualization of temporal frequent event sequences. In: Proceedings of the 19th international conference on Intelligent User Interfaces, pp 153–162, https://doi.org/10.1145/2557500.2557508

Pinto ARF, Nagano MS (2019) An approach for the solution to order batching and sequencing in picking systems. Prod Eng Res Dev 13:325–341. https://doi.org/10.1007/s11740-019-00904-4

Plaisant C, Mushlin R, Snyder A, et al (1998) Lifelines: using visualization to enhance navigation and analysis of patient records. In: AMIA 1998, American Medical Informatics Association Annual Symposium, Lake Buena Vista, FL, USA, November 7-11, 1998. AMIA

Ramtin F, Pazour JA (2015) Product allocation problem for an as/rs with multiple in-the-aisle pick positions. IIE Trans 47(12):1379–1396. https://doi.org/10.1080/0740817X.2015.1027458
Sedlmair M, Meyer MD, Munzner T (2012) Design study methodology: Reflections from the trenches and the stacks. IEEE Trans Vis Comput Graph 18(12):2431–2440. https://doi.org/10.1109/TVCG.2012.213
Sun D, Huang R, Chen Y et al (2020) Planningvis: a visual analytics approach to production planning in smart factories. IEEE Trans Vis Comput Graph 26(1):579–589. https://doi.org/10.1109/TVCG.2019.2934275
Tang J, Zhou Y, Tang T et al (2022) A visualization approach for monitoring order processing in e-commerce warehouse. IEEE Trans Vis Comput Graph 28(1):857–867. https://doi.org/10.1109/TVCG.2021.3114878
Tappia E, Roy D, Melacini M et al (2019) Integrated storage-order picking systems: Technology, performance models, and design insights. Eur J Oper Res 274(3):947–965. https://doi.org/10.1016/j.ejor.2018.10.048
Tarigonda A, Hymes B, Nikonovich-Kahn A (2018) E-commerce flow management in fulfillment centers through data visualization. In: International Conference on HCI in Business, Government, and Organizations, Springer, pp 767–778, https://doi.org/10.1007/978-3-319-91716-0_60
Viégas F, Wattenberg M, Hebert J, et al (2013) Google+ripples: A native visualization of information flow. In: Proceedings of the 22nd International Conference on World Wide Web. Association for Computing Machinery, p 1389–1398, https://doi.org/10.1145/2488388.2488504
Vrotsou K, Johansson J, Cooper M (2009) Activitree: interactive visual exploration of sequences in event-based data using graph similarity. IEEE Trans Vis Comput Graph 15(6):945–952. https://doi.org/10.1109/TVCG.2009.117
Wang Q, Mazor T, Harbig T et al (2022) Threadstates: state-based visual analysis of disease progression. IEEE Trans Vis Comput Graph 28(1):238–247. https://doi.org/10.1109/TVCG.2021.3114840
Wang Y, Peng T, Lu H et al (2022) Seek for success: a visualization approach for understanding the dynamics of academic careers. IEEE Trans Vis Comput Graph 28(1):475–485. https://doi.org/10.1109/TVCG.2021.3114790
Wongsuphasawat K, Gotz D (2012) Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization. IEEE Trans Vis Comput Graph 18(12):2659–2668. https://doi.org/10.1109/TVCG.2012.225
Wongsuphasawat K, Shneiderman B (2009) Finding comparable temporal categorical records: A similarity measure with an interactive visualization. In: 4th IEEE Symposium on Visual Analytics Science and Technology, IEEE VAST 2009, Atlantic City, NJ, USA, October 11-16, 2009, part of VisWeek 2009. IEEE Computer Society, pp 27–34, https://doi.org/10.1109/VAST.2009.5332595
Xu P, Mei H, Ren L et al (2017) Vidx: visual diagnostics of assembly line performance in smart factories. IEEE Trans Visl Comput Graph 23(1):291–300. https://doi.org/10.1109/TVCG.2016.2598664
Xu H, Farajtabar M, Zha H (2016) Learning granger causality for hawkes processes. In: Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016, JMLR Workshop and Conference Proceedings, vol 48. JMLR.org, pp 1717–1726
Yang P, Miao L, Xue Z et al (2015) Variable neighborhood search heuristic for storage location assignment and storage/retrieval scheduling under shared storage in multi-shuttle automated storage/retrieval systems. Transp Res Part E Logist Transp Rev 79:164–177. https://doi.org/10.1016/j.tre.2015.04.009
Zhang W, Wong JK, Wang X, et al (2022) Cohortva: a visual analytic system for interactive exploration of cohorts based on historical data. In: IEEE Transactions on Visualization and Computer Graphics pp 1–11. https://doi.org/10.1109/TVCG.2022.3209483
Zhao J, Drucker SM, Fisher D, et al (2012) Timeslice: interactive faceted browsing of timeline data. In: Tortora G, Levialdi S, Tucci M (eds) International Working Conference on Advanced Visual Interfaces, AVI 2012, Capri Island, Naples, Italy, May 22-25, 2012, Proceedings. ACM, pp 433–436, https://doi.org/10.1145/2254556.2254639
Zhao J, Liu Z, Dontcheva M, et al (2015) Matrixwave: Visual comparison of event sequence data. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp 259–268, https://doi.org/10.1145/2702123.2702419
Zhen L, Li H (2022) A literature review of smart warehouse operations management. Front Eng Manag. https://doi.org/10.1007/s42524-021-0178-9
Zhong R, Lan S, Xu C et al (2015) Visualization of rfid-enabled shopfloor logistics big data in cloud manufacturing. Int J Adv Manuf Technol. https://doi.org/10.1007/s00170-015-7702-1