

Connecting Paper to Digitization—a Homework Data Processing System with Data Labeling and Visualization

Qingcheng Li, Heng Cao, Ye Lu
College of Computer and Control Engineering
Nankai University
Tianjin, China

liqch@nankai.edu.cn, chaser_cahe@foxmail.com, luye@nankai.edu.cn

ABSTRACT

How to monitor students' learning in daily learning behavior and guide teachers to teach more rationally are very important problems. While the digital teaching resources are popular, we can not ignore the information of traditional paper media. We design a homework processing system to solve the problems. First of all, we begin the operation from the acquisition side, by photo taking and image preprocessing, we make the homework picture with semantic meanings. Then we construct an answer reference system, which shows the answers of different kinds of incorrectness as samples for teachers to refer to when correcting. Semantic pictures are dispatched to teachers handheld devices for intelligent and precise correct activities. We divide correct process into determining the degree of error and giving comments to evaluate and inspire students. Finally, we design a web page for homework data management and correction statistics to help teachers understand the homework completion status of students. The design of homework correction scheme is easy to deploy and can be used in most teaching environments.

CCS CONCEPTS

• **Human-centered computing** → *Ubiquitous and mobile computing systems and tools*;

KEYWORDS

paper homework, data labeling, visualization

ACM Reference Format:

Qingcheng Li, Heng Cao, Ye Lu. 2017. Connecting Paper to Digitization—a Homework Data Processing System with Data Labeling and Visualization. In *MobiQuitous 2017: the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, November 7–10, 2017, Melbourne, VIC, Australia*. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3144457.3152353>

1 INTRODUCTION

The key point in reaching excellent studying is not knowledge or ideas, but the construction of these knowledge and ideas by learners.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobiQuitous 2017, November 7–10, 2017, Melbourne, VIC, Australia

© 2017 Association for Computing Machinery.

ACM ISBN 978-1-4503-5368-7/17/11...\$15.00

<https://doi.org/10.1145/3144457.3152353>

Teachers should answer the students about what to study, how to study and what to study next. A gap between new knowledge and existing knowledge to students must be filled. How to discover the gap is a main consideration in education.

To figure out the solutions, we think formative assessment[3] is one of the important methods to guide teachers in teaching activities and helping students find their own shortcomings. As an important participant in teaching activities to help students carry out knowledge construction activities, teachers should improve their teaching skills besides mastering the students' learning situation. For teachers, teaching activities are not scientific, but the accumulation of experience and feedback, no specific method can be directed to all students, so in the daily teaching activities, adjustments according to the students feedback of teaching plan is very important.

In China, the tide of elementary education reform has been sweeping. Many primary and secondary schools have started using electronic textbooks and using electronic equipment for job feedback. Digital teaching resources provide a convenient channel for data collection and arrangement, and provide data support for formative evaluation. As for traditional media, the data is difficult to be effectively utilized, and it has become an obstacle to digital education. As the paper media is more common in daily operations and has higher convenience, we can not ignore the data on paper.

Our research starts from homework because using homework to improve the academic achievement of students plays a significant role[8] and homework is a key approach to acquire formative assessment of everyday teaching activities. To make up the lack of visibility[14] of learning, we design a homework control website for teachers to do homework arrangement and result analysis to provide a more direct view of students' learning status.

This paper expounds the significance of this homework system from two aspects of learning and teaching. Normalized operation flow:

- **Collection:** we collect homework data from multiple channels: mobile phones, high speed scanners, electronic files. Support multi-format and convenient collection.
- **Tagging:** by designing homework template before publication, customization, and post publication layout decomposition, we generate a file to describe that is in the paper.
- **Correction:** we use historic data to instruct teacher for precise correcting and restore correcting result for visualization and further analysis.
- **Sharing:** as we talked before, we share effective teaching practices and ensure teaching has the greatest possible impact on students' learning.

This paper begins with related work, covering the literature on note-making technology. Then we display the overall five layers structure of our system. This is followed by a detailed introduction of every deployed layer. Next we talk about how to visualize homework correcting data on the web. Finally, we reiterate our key findings and discuss the future work.

2 RELATED WORK

Online marking system [16] has been widely studied, and has been applied to the middle school entrance examination and college entrance examination and other important examinations. By printing the answering paper in advance, a high-speed answer paper image acquisition can be realized by using an optical mark reader [9], and the pre-processing operation for classifying the answers according to blocks pre-defined. By scoring teachers can grade the students' responses through the Internet for both online courses [2] and traditional course. However, this scheme is difficult to deploy, performs higher requirements for instrument and equipment and is not suitable for extensive promotion. We design a scheme to mark the answering region, which means all the homework submitted images will be marked using the XML semantic description. We use daily equipment such as mobile phone [5], digital pen and other high speed photographic apparatus [11] instrument to realize homework pictures collection, and through the intelligent equipment to restore the traditional homework scene.

In the process of reduction of scene, teachers of homework is divided into the marker symbols and evaluation of the students' correct answer, through gesture recognition can be completed on the classification of correcting the results, the sentence similarity of comments recommended.

Record of academic achievement also plays an important role in improving academic achievement [10]. According to the research of learning dashboard [13], visualization of learning results can help teachers improve teaching activities, and also help students understand their own learning shortcomings and conduct targeted learning [7].

3 SYSTEM OVERVIEW

Seeing from the perspective of function module, the homework system consists of five parts respectively, they are: the data processing server, pre-processing module, precise marking module, operation management and visualization analysis module, experience sharing module. Among them, the preprocessing module realizes high-speed automatic collection, semantic processing and uploading function of the homework image. The precise correction module is implemented by handheld intelligent devices, including real-time statistics based on gesture recognition [12], recommendation reference answer score and data upload and download functions. The visualization module provides functions such as organization of curriculum and homework resources, homework correction, result analysis, and so on. The shared module shares the teaching and learning experience with homework as a link. The data processing server provides data preparation and calculation services for the above functional modules.

4 DATA TRANSMISSION

Acquisition and Storage: Students' assignments are taken by means of photography, scanning, etc., and the homework are saved in the form of pictures. Each of the corresponding types of homeworks has a description file corresponding to it. Description files, detailed records of the layout of the homework layout and specific topic information. This descriptive document provides data support for post processing corrections. Both the picture and the description file of the homework are saved in the homework server. The homework management page will read the homework data from the server and provide visual support for the teacher's evaluation of the student's homework. The process realizes the monitoring of students' learning situation and process evaluation.

Correction as Marking: if the acquisition work is marked complete, the system will enter into correcting the results described in the paper, and the correcting homework and save the correcting result description file to be saved. The description file provides data preparation for building the homework semantic network. At the same time, without correcting the homework will be pushed to the marking equipment. To the original image data including marking information, homework description file, the recommended list and available comments recommended classification reference answer. Comments on the list by the server according to the assignment of the knowledge point, answer and correcting of the use of contract arrangement is recommended, while considering the mental characteristics of students in this age group may have the reference to describe marking errors, defects in the answer when students have. The answer for the classification reference reviewers provide reference for answer classification. Reviewers can be viewed in different levels of error for the current sample answer, marking the title of a reasonable error rating, with comments, can be achieved on the depth and precision of correcting students' answers. At the end of the current correction, the result of the correction will also be entered into the grading reference answer.

Transmission of Return: the data at the end of the correction include the corrected picture and the description file of each picture. The description document at this time retains the correction result, and the result includes error grading and commenting. The content is saved to the server, and can also be accessed and managed through the homework management page. Shown in Fig.1.

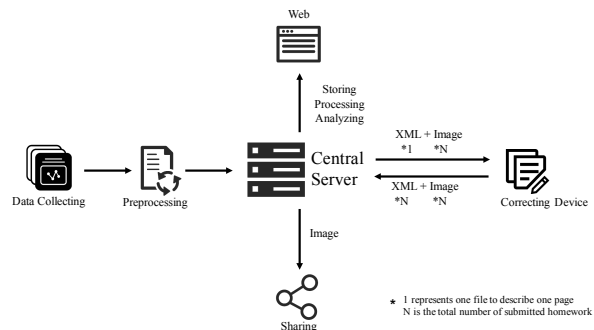


Figure 1: Software Data Flow

5 HOMEWORK CORRECTION ON SMART DEVICES

In this section, we will discuss about how to design a stable, efficient and intelligent correction system using smart mobile devices with touch screens. We aim to provide electronic correction environment and instruction operation, and also can automatically generate statistical visualization information of homework. Specifically, the system uses a hierarchical operating interface design. The hierarchical design can decouple the input information, guarantee the system to be stabilized, and improve the system scalability. According to the different functions, we consider to divide the the system input interface into five layers: base layer, operation storage layer, gesture recognition layer, precise correction reference layer and statistical analysis visualization layer.

5.1 Software design

There are three main parts of the application for precision correction on handheld smart devices which is shown in Fig.2: A storage subsystem, an operating subsystem, and a display subsystem. The storage subsystem implements the storage of the original homework, which includes the original image and its corresponding XML description file. At the same time, correcting the results of statistics and comments on the recommendation list with Android built-in database Sqlite is stored in the local database access from the server in history in. Operation operation subsystem recognizes users, including recognition of gesture recognition classification to judge, the error level screen, save, flip action; operation subsystem data synchronization update to the local database, when marking work is completed, stored in the data will also be updated to the server. The display subsystem is responsible for presenting the results, including corrections for hand gestures, handwriting and graffiti. The subsystem mainly retains the traces of the user’s written actions and displays, to a greater extent, to restore paper-based habits in traditional correction environments. We design a five layers approach of displaying metrics referring to format of multi-layer design of Microsoft Powerpoint[4].

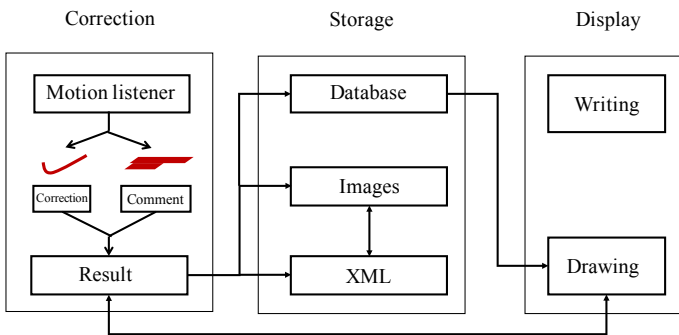


Figure 2: Application Structure Overview

5.2 Base Image Layer

The bottom layer displays the overall picture of the homework to be corrected, and highlights the corresponding homework local map according to the user’s touch position. Specifically, when

a user needs to correct a topic, click the topic first. The bottom layer listens to the click operation and parses the specific coordinates (x, y) of the clicks. After that, the bottom layer determines the location of the coordinates by the XML file in the layout format, and extracts the rectangle region coordinates $(X_{LeftTop}, Y_{LeftTop}), (X_{LeftBottom}, Y_{LeftBottom}), (X_{RightTop}, Y_{RightTop}), (X_{RightBottom}, Y_{RightBottom})$ to check which the topic belongs. Details are also shown in Fig. 3 The XML file is obtained from the server with the homework file, which defines the format and layout information of the homework file. Finally, the bottom layer of the complete picture region extraction operation, the selected rectangle of the subject image to enlarge in top layer now.

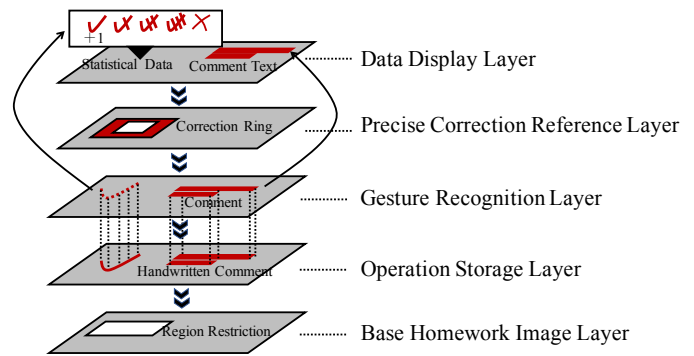


Figure 3: Base Layer with Region Restriction

5.3 Operation Storage Layer

The gesture of homework processing layer in the user level judging and handwritten comment is a specific topic based on the top layer bottom layer operation. So we designed the action save to save all of the user’s actions in the original image layer[1][15].

Specifically, we will save all of the user’s gestures, including the level of checking and correcting handwriting annotation become a combination of layers, according to the specific subject layer zoom and extended to the original image layer set operation.

5.4 Precise Correction Reference Layer

Teachers often evaluate standard answers by correcting their assignments. In order to simplify the operation, precise correction of the standard answer reference layer receives the bottom layer transfer correcting problems in number to find the corresponding database subject and history homework recommended information, comments, and display above the bottom layer, for the teacher to make more standard and reasonable decision marking.

5.5 Gesture Recognition Layer

The homework correcting system integrates the traditional correcting behavior, integrates in the mobile device terminal, and supports the endorsement of the teacher’s gestures. According to the investigation of the usual habit of correcting homework, the system supports the teacher’s evaluation of the five grades of homework, and the graphs of the five grades are shown in Fig.5. For example, the fork is set for all wrong, a tick with one slash is set for almost

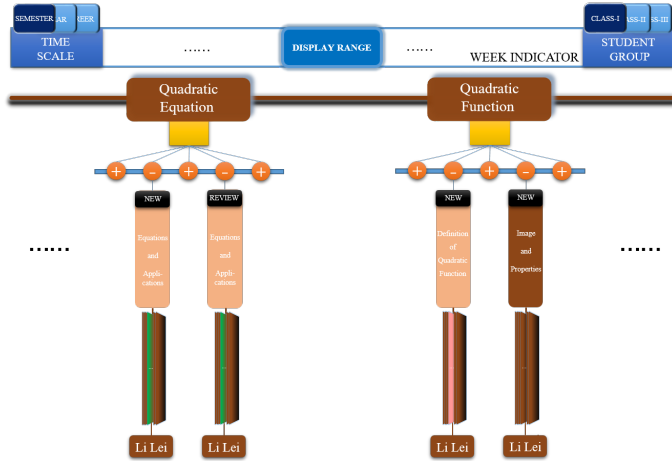


Figure 4: Web Based homework Statistic Data Visualization

corret but with a little defect. The number of slash indicates error level assessed by the commenter. Gesture recognition layer can automatically, efficiently and accurately recognize the gestures of five graphics, and convert the graph into the corresponding scoring level for statistical analysis of final correction result information. At the same time, the layer also supports user handwriting input comments and other marking information to keep the teacher’s usual correcting habits as far as possible. Gesture recognition has used the number of strokes and the Google Android gesture recognition API to recognize input types based on the cos value of the input gesture path.

Specifically, shown in Algorithm 1, the gesture recognition layer initiates a gesture listener that monitors user gestures for every two seconds. Here, we assume that when the teacher corrects, the time of grading the result is less than or equal to two seconds. The specific monitoring algorithm flow is shown in Algorithm 1.

Within one second, if the number of strokes as a gesture, then determine whether it is right. If yes, is full of rank count increased by 1; if it is not right, it belongs to the user input, write graffiti.

If the gesture is two strokes, the difference is that the user input is a fork or half pair. The gesture recognition layer calls the Android gesture recognition API[6] (*android.gesture.GestureLibraries*, *android.gesture.GestureLibrary*, *android.gesture.GestureOverlayView*, *android.gesture.Prediction*). Identify. And the recognition results corresponding to the statistical value increased by 1 processing.

In other cases, the input gesture types are distinguished by counting the strokes, and the corresponding rank statistics are increased by 1.

5.6 Statistical Analysis Visualization Layer

The statistical analysis shows that the layer can make the teacher observe the overall score of the correction group in real time. Specifically, the layer stores the correction result statistics of the object groups. After obtaining the correction grade of the gesture layer recognition, the number of the subjects in each class can be displayed in real time, and the teacher can quickly understand the corresponding knowledge points of the students.

Algorithm 1 Gesture Recognition Process

Input:

The start point of touch event P_{xy} ;

The set of strokes S with time duration T_i and each stroke length L_i ;

Output:

Counter of different recognition result +1;

1: Find corresponding topic area according to P_{xy}

2: **if**

$$\sum_{i=1}^n T_i < THRESHOLD$$

&& each $L_i > THRESHOLD$ **then**

3: Open a gesture recognition listener;

4: **end if**

5: Find size *sum* of the Stroke set.

6: Calculate prediction score of stroke set comparing to the pre-defined user gestures to get *result*;

7: **if** *sum* = 1 **then**

8: Restrict recognition result in *RIGHT* and *NONGESTURE*

9: **else if** *sum* = 2 **then**

10: Restrict recognition result in *ALMOST*, *WRONG* and *NONGESTURE*

11: **else if** *sum* = 3 **then**

12: Restrict recognition result in *SOMEWHAT* and *NONGESTURE*

13: **else if** *sum* = 4 **then**

14: Restrict recognition result in *LITTLE* and *NONGESTURE*

15: **else**

16: Restrict recognition result in *NONGESTURE*

17: **end if**

18: **return** *result*;

In short, through collaboration with these five layers, we design the system can make the teachers without changing the original paper type homework habits under the condition of efficiently correcting in the mobile terminal equipment, and in the correction process, correcting the reference in precision and real-time scores of statistical information case teachers can clearly and accurately understand the current students’ understanding degree, greatly improving the efficiency of the teachers review.

6 HOMEWORK CORRECTION EVALUATION WITH PRECISE CORRECTION RING

In this section, we introduce a auxiliary tools called correction ring to provide teachers correction referenece and recomment comments for specific homework topics. The ring also provides history correction data and current correction status. The rings is shown in Fig. 7. And we show a example of a corrected homework page in Fig. 6.

6.1 correcting ring operation procedure

The correcting ring is an auxiliary tool for correcting the single item subject accurately, shown in Fig. 7. The pre-processed data of precise correction is provided by teaching metadata framework, which is mainly carried out off-line. In correcting ring is not displayed,

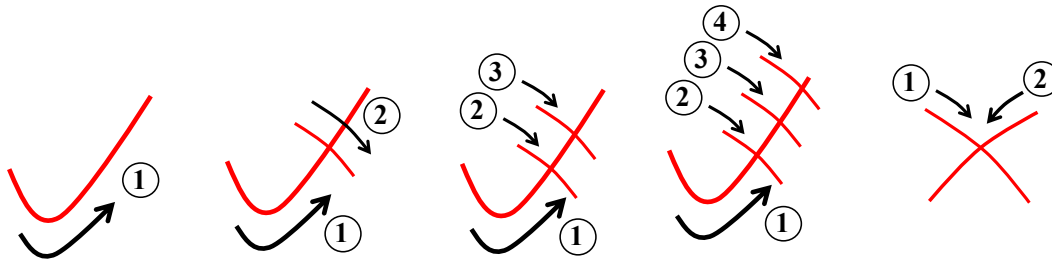


Figure 5: Five Levels of Fault Grade

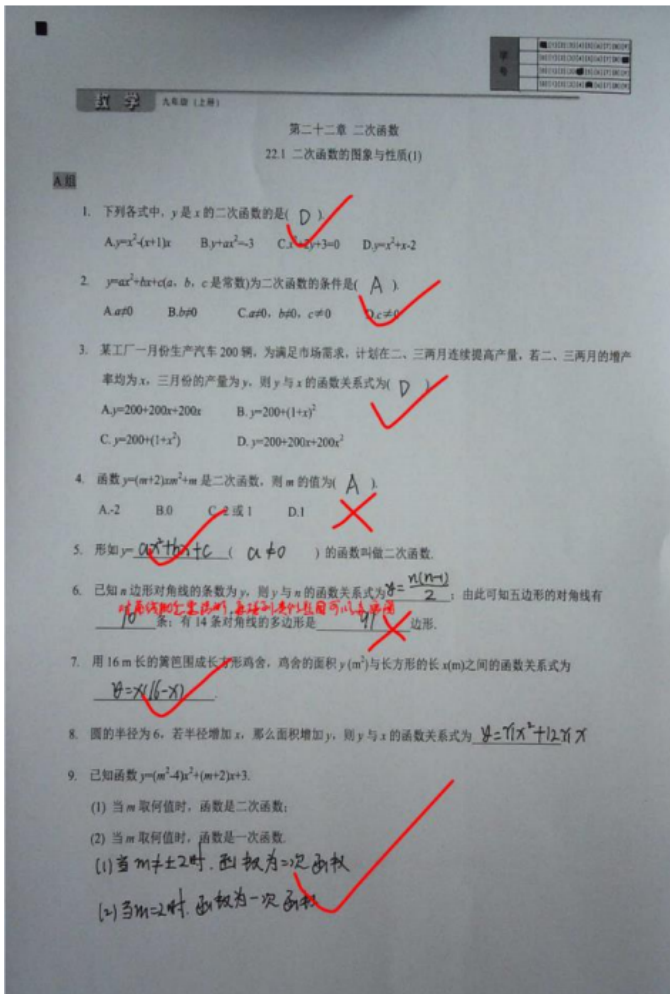


Figure 6: Homework after Correction

recommended by obtaining the evaluation comments correcting result after correcting the choice of recommendation, criticism, and achieve the same effect in the presence of precision correcting ring.

- Enter: double click to specify the exact correction problem area, double-click the pop-up precise correction ring.

- Exit: after precise correction, double click the outer ring area to hide the precise correction ring. After you leave the correct correction ring, the results will be saved to the main page

6.2 Precise Correction References with Answer Classification

The reference answer is not the correct answer, but rather the student's objective reflection of the use of knowledge, and the identification of the cause and extent of the problem by correcting the mark of the marker. Refer to the answer as much as possible to the questions that students may have in answering them.

The reference answer is to provide references for corrections in correcting activities. According to their own experience in correcting this problem causes or difficult to quantify the defect when the reference are given, at the same time, correcting this assignment correcting results will be recorded into the reference answer, and then correcting them

The entry button for the classified answer is located above the correction ring. Click on the button corresponding to the different error levels to view the recommended reference answers under that category.

- Click the button, pop up the classified reference list, and slide the thumbnail view
- Click on different categories of buttons to switch different categories of reference answer lists
- Click the thumbnail and eject the target reference on top of the correction question
- Click on other thumbnails to change your reference answers
- Click the reference answer big picture to exit the larger view mode
- Click the button, drop out the reference answer, and view the list

6.3 Correcting Operation Process

Accurate marking is the correction of errors and defects. Precise correction is mainly divided into two parts, the error level division and the generation of criticism. The error grade is annotated according to the historical answer mark of the answer reference system, and combines the personal experience of the correcting person with the history and the same type to correct the result, and has certain scientific nature.

- Error Level Identification:** To record five kinds of gestures, which correspond to “yes” and four different error grades. Gesture, graphics and classification, answer, sign, sign the same. Other actions are recognized as writing. After correcting the gesture, the 5 types of correction results will be statistically corrected and the current correction statistics “+1” operation.
- Handwritten Comments:** The comments should be distinguished between the five hand operations gestures and writing operations. When the gesture is recognized as writing, the written result will be saved as a picture. Comments can be written in the handwriting correcting ring and also can be written when the correct ring is not appearing. The handwritten characters saved as images will be sent to the server for further processing such as handwritten character recognition and emotion analyzing.

6.4 Selecting Comments

The whole comments consists of two half sentence: Knowledge/non-knowledge defect sentence + emotion encouragement sentence. The first half sentence describe defects on knowledge and answering technique. And the last half is emotion encouragement sentence.

First level comment button - two column vertical arrangement are mutually exclusive, only one button can be selected. The evaluation of knowledge and non-knowledge acquires clicking the vertical buttons or horizontal button, the color of the chosen button will be deepen. Second level comment button - these buttons will shown as a more detailed classification of the first level comment button. We provide five comment sentences in each first level comment button.

The chosen knowledge/non-knowledge half sentence and emotion encouragement half sentence will be composed into a complete comment for labeling the student answer.

When the ring is displayed, by double click outside the ring area, comments will be saved to the main interface on the operation storage layer.

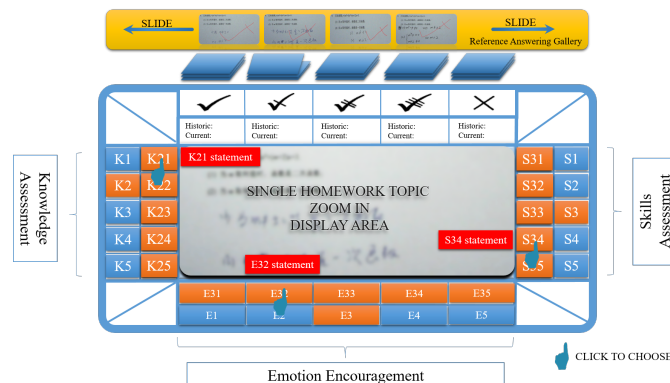


Figure 7: Correction Ring

7 WEB BASED HOMEWORK VISUALIZATION

We display and organize the organization and statistics of the homework resources through the web page. At the top of the page is the timeline, in weeks. Through the left option, choose different time scales, respectively, by semester, school year, career three scales. The middle slider represents the time range corresponding to the resources displayed in the current screen range. For example, when the screen shows 19-21 weeks of teaching content within a semester, the slider covers 19-21 or three times. Teachers can use the mouse wheel to control the number of chapters displayed in the screen. When the content displayed in the screen is too long, the system will merge the similar contents according to certain rules.

At the far right of the timeline is the student group selection button. In our program, a teacher divides classes according to different classes. When a class of labels is selected, all the resources at the bottom are connected to the class.

The contents of the brown box below are the main contents of a chapter. Each chapter organizes the course contents into sections. Teachers can arrange the teaching plan through the plus and minus.. In the boxes under the add and subtract, the black boxes at the top indicate the course. There are new classes, recitations, experimental classes, etc.. The main contents of the section are indicated in the box. When the color of the box is light brown, it represents the end of the course. Dark brown indicates that the subject has not yet started.

The bottom louvered structure is the statistics of student assignments. When you click the Class one button, the window page shows the completion of different topics by color. From red to green represent the severity of the problem. Red indicates a serious problem on the subject. Green means there is no serious problem with the subject. Click the class flag to switch the query to the homework status.

At this point, the shutter is set to single page linkage status. Homework condition of single student will be shown at the same time. Teachers use mouse to slide on the shutter, the system can display learning conditions of single student in different chapters under the completion of the work trend. Teachers can clearly see which chapter of the student’s knowledge is not good, so as to infer what problems the students are having in their studies. It allows accurate positioning of individual student defects. Click on the student’s single page, teachers can thoroughly check the student’s real homework completion and correction.

8 RESULTS AND FUTURE WORK

This paper designs and implements a closed-loop system for homework correction. The system consists of four parts: 1, the original operation collection and pre-processing part; 2, based on mobile devices, 3 homework; homework results visualization of web applications; 4, responsible for the data processing server. The system can make teacher homework correcting and automatic correcting results to the use of gestures in mobile electronic devices, intelligent, efficient and convenient to improve the traditional teacher grades homework habits.

Compared with the traditional ways of homework correcting, the system allows students to submit electronic homework, traditional homework, teachers can conveniently carry the workload out of

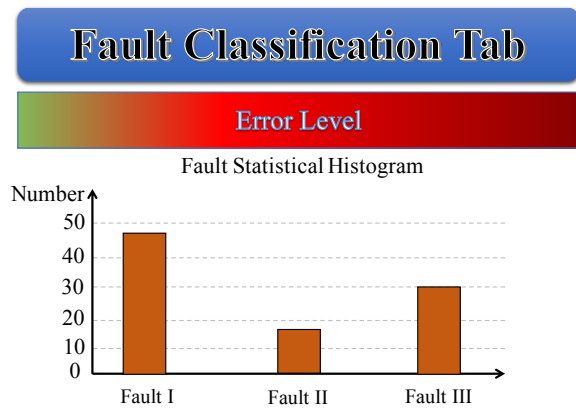


Figure 8: Data Displaying Demo on the Web

office, which greatly improve the quality and efficiency of homework correction procedure. We provides tool support for formative assessment for teachers. We realize our home work correction client system in the Android mobile table, using high speed photographic apparatus for homework data collection. The visualization system is developed based on web using PHP and JS. We verified the feasibility of the system. It is stable and efficient to assist teachers in daily teaching activities and homework.

In the future we will further improve user experience of homework correction , increase the analysis statistics and provide more intelligent answer comment recommendation, and further improve the system's actual application value, promote the development of digital education revolution.

9 ACKNOWLEDGEMENT

This work is financially supported by Ministry of Science and Technology(2016YFC0400709), The National Key Research and Development Program of China(2016YFC0400709), the Natural Science Foundation of Tianjin(No.16JCYBJC15200, 17JCQNJC00300, 15ZXD SGX00020) and the funding named "Research on Visualization Framework and Device Oriented Group Teaching" of Tianjin Read Electronic Co., LTD. The corresponding author is Ye Lu.

REFERENCES

- [1] Aaron Bauer and Kenneth R. Koedinger. 2007. Selection-based note-taking applications. In *Conference on Human Factors in Computing Systems, CHI 2007, San Jose, California, Usa, April 28 - May*. 981–990.
- [2] Chuck Bennett. 2000. Multipart Homework Problems Over the Web. *Computing in Science & Engineering* 2 (2000), 87–95.
- [3] Randy Elliot Bennett. 2011. Formative assessment: a critical review. *Assessment in Education Principles Policy & Practice* 18, 1 (2011), 5–25.
- [4] Robin Cover. 2006. Open Document Format for Office Applications (OpenDocument). (2006).
- [5] Google. 2017. (2017). <http://www.tensorflow.org>
- [6] Google. 2017. (2017). <http://www.android-doc.com/sdk/>
- [7] Hendrik Greller, Wolfgang Drachsler. 2012. Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society* 15, 3 (2012), 42–57.
- [8] John Hattie. 2008. *Visible Learning*.
- [9] Stephan Hussmann, Leona Chan, Celine Fung, and Martin Albrecht. 2003. Low-cost and high-speed optical mark reader based on an intelligent line camera. In *Aerosense*.
- [10] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Sigchi Conference on Human Factors in Computing Systems*.

- 557–566.
- [11] Qingcheng Li, Heng Cao, Haoqiang Yan, and Ye Lu. 2017. High Speed Homework Data Collecting System under K-12 Education Circumstance. In *International Symposium on System and Software Reliability*.
- [12] Sushmita Mitra and Tinku Acharya. 2007. Gesture Recognition: A Survey. *IEEE Transactions on Systems Man & Cybernetics Part C* 37, 3 (2007), 311–324.
- [13] Katrien Verbert, Erik Duval, Joris Klerkx, Sten Govaerts, and Jos Luis Santos. 2013. Learning Analytics Dashboard Applications. *American Behavioral Scientist* 57, 10 (2013), 1500–1509.
- [14] Katrien Verbert, Sten Govaerts, Erik Duval, Jose Luis Santos, Frans Van Assche, Gonzalo Parra, and Joris Klerkx. 2014. Learning dashboards: an overview and future research opportunities. *Personal & Ubiquitous Computing* 18, 6 (2014), 1499–1514.
- [15] M. Ye, P. Viola, S. Raghupathy, H. Sutanto, and C. Li. 2007. Learning to Group Text Lines and Regions in Freeform Handwritten Notes. In *International Conference on Document Analysis and Recognition*. 28–32.
- [16] Jianxiong Zhou, Jiachen Yang, Houbing Song, Syed Hassan Ahmed, Amjad Mehmood, and Haibin Lv. 2016. An online marking system conducive to learning. 31, 5 (2016), 2463–2471.