PACALL: Supporting Language Learning Using SenseCam

Bin Hou, Hiroaki Ogata, Mengmeng Li, Noriko Uosaki
Dept. of Information Science and Intelligent Systems, University of Tokushima, Japan
hou-bin@iss.tokushima-u.ac.jp

Abstract—In our previous works, we developed a system named SCROLL in order to log, organize, recall and evaluate the learning log. However up to now, we just use an active mode to record logs. This means that a learner must take a capture of learned contents consciously and most of learning chances be lost unconsciously. This paper proposes a system named PACALL (Passive Capture for Learning Log) in order to have a passive capture using SenseCam to solve this problem. With the help of SenseCam, learner’s activity can be captured as a series of images. With the help of this system, a learner can find the important images by analyzing sensor data and images processing technology. Finally, the selected images will be uploaded to the current SCROLL system as ubiquitous learning logs. This research suggests that SenseCam can be used to do passive capture of learning experiences and workload of reflection can be reduced by analyzing sensor data of SenseCam.

Keywords- PACALL; SenseCam; life log; learning log; passive capture; ubiquitous learning

I. INTRODUCTION

Learning Log was originally designed for children as a personalized learning resource [1]. It was set by teachers to help their students record their thinking and learning. In this learning log, the logs were usually visually written notes of learning journals. We defined a ubiquitous learning log as a digital record of what a learner has learned in the daily life using ubiquitous technologies. We developed a system SCROLL (System for Capturing and Reminding Of Learning Log) [2] that helps learners collect their learning experiences as ubiquitous learning objects (ULLOs). Also, all of the collected ULLOs are organized, shared in this system, and the learning effect can be enhanced. The model of a learning process is shown in Figure 1 and we call it LORE [2].

![Figure 1. LORE Model in SCROLL](image)

- Log what the learner has learned: When the learner faces a problem in the daily life, s/he may learn some knowledge by her/himself, or ask others for a help in terms of questions. The system records what s/he learned during this process as a UULLO.
- Organize ULL: When the learner tries to add a UULLO, the system compares it with other UULLOs, categorizes it and shows the similar UULLOs if exist. By matching similar objects, the knowledge structure can be regulated and organized.
- Recall ULL: The learner may forget what s/he has learned before. Rehearsal and practice in the same context or others in idle moments can help the learner to recall past UULLOs and to shift them from short-term memory to long-term one. Therefore, the system assigns some quizzes and reminds the learner of her/his past UULLOs.
- Evaluate: It is important to recognize what and how the learner has learned by analyzing the past UULL, so that the learner can improve what and how to learn in future. Therefore, the system refines and adapts the organization of the UULLOs based on the learner’s evaluation and reflection.

Miller and Gildea [3] compared the way that children are taught words from dictionary definitions and a few exemplary sentences with the way vocabulary is normally learned outside the school. They noted that people generally learn words outside school. It suggests that using mobile devices is a good for people to remember the vocabulary since people can use mobile devices anywhere and anytime.

Therefore, in our previous works, we used mobile devices such as smartphone and tablet PC with the aim of registering UULLOs whenever and wherever the learner wants to log. In other words, learner has to take photos and registering UULLOs manually. [13]. It means that learners must record their learning experiences consciously. This is the active mode to record the learning experiences. However under this active mode, learners cannot record all of the learning experiences in the system, for example they maybe forget to take some pictures when they learned some new words, or although they want to take photos but they are busy, as a result most of chance that they can learn will be lost and forgotten.

Passive mode has the ability to solve these problems. In this mode, learner is not required to record learning experiences actively, and all the learning activity will be record by some devices automatically. Therefore, we attempt to introduce the concept of life log into this system to record learning experiences in a passive way. The notion of life log can be tracked back at least 60 year [2]. It means to capture a person’s entire life or large portions of life. It usually uses digital devices to record life log such as wearable cameras or...
video recorders. For example, in the early 1980s Steve Mann captured his life using wearable computer and streaming video and even his everyday life 24 hours a day in order to see what he was looking at [8]. The life log brings us the data of whole life of not only learning but also other activities. However, if there is any way that we can extract the learning part from it, the learning log will be more significant and more sufficient. Besides, our system captures the learning log beyond their consciousness and learners’ burden will also be reduced. Microsoft’s SenseCam [6] is an effective way to capture the life log. It is a wearable camera equipped with a number of sensors. The SenseCam is proposed to record a series of images and capturing a log of sensor data.

In this paper, we propose a system named PACALL (PAssive CApture for Learning Log) to capture the learning log passively using Microsoft’s SenseCam. With the help of analyzing sensor data and image processing technology, it extracts the meaningful images for learning from life log and helps learners upload the learning content easily.

II. RELATED WORKS

A. Learning and cognitive theories

This research is advocated by the following theories:

- Authentic learning and situated cognition: We are learning from what happened in the real world, so-called, from the authentic environment [1]. Especially, we learn a lot of vocabularies from our conversations, TV, and other daily activities. Therefore, it is very important to capture what we learned and recall them to enhance learning.
- Encoding specificity: The same context reminds us of the things that we have experienced there [7]. For example, if s/he is in the upstairs and wants to bring something from the downstairs, s/he may forget in the downstairs. However, if s/he goes back to the upstairs, s/he reminds of the thing that s/he wanted. Therefore, if the user is in the same situation where s/he registered the ULLO, SCROLL automatically asks the user the quiz about the ULLO.
- Picture superiority: People can remember vocabularies with their pictures more than those without their pictures [8]. Therefore, SCROLL shows the picture of ULLO for a quiz.
- Spaced repetition: This is the learning method that uses increasing intervals of time between successive reviews of previously learned knowledge [9]. Based on this theory, SCROLL sends a quiz to the learner.

B. MyLifeBits

MyLifeBits [11] is a Microsoft’s project. The aim of this project is to implement Bush’s Memex model [4] that proposed to store everything that you saw and you heard.

MyLifeBits has a large amount of storage that can store email messages, web pages, books, photos, sounds, videos, etc. It also has a full-text search function to supply users with searching text, audio annotations and hyperlinks.

In addition, the MyLifeBits project team is also using SenseCam to have the passive capture of life log and upload the sensor information along with the photos to the MyLifeBits repository [12].

We have learned a lot from this system. In our previous works, we had made it possible to store the learned material such as photos, sounds, videos and pdf files into the system repository. Besides, we have also implemented recall functions that use quizzes and contextual information to help learners to remember what they have learned. However, all works that we have done are using active logging mode, not passive logging mode. It means that learners must record their learning experiences as learning material by themselves. Comparing to the passive mode, in the active mode we are more likely to lose learning experiences since we are not necessarily able to record what we have learned or sometimes we just forget to record it. Therefore, we planned to introduce passive capture in our project with SenseCam.

C. JAMIOLAS

JAMIOLAS [13] is Japanese mimicry and onomatopoeia learning assistant system. This system uses sensor to get the context information from real world such as temperature, light and sound level and use these data to support learning Japanese mimicry and onomatopoeia words. Because most of these words are just Japanese feelings, this system simulates the feeling of human beings with sensors, and generates proper word to help non-Japanese learners learning mimicry and onomatopoeia words. Each word has relationships with a number of sensors. Sensor data types are attributes of the word. For example, “hiyahiya” means cold in Japanese, so this word has a relationship with temperature sensor, and the temperature data is an attribute of “hiyahiya”. The threshold is set by native Japanese speakers. In the second generation of JAMIOLAS [14], sensor network was introduced into this system, and in the third generation [15] this system also used online sensor data and supported
learning mimicry and onomatopoeia with multimedia materials.

In the PACALL, we also use the sensor data to analyze the situation of an image file. When sensor data is collected by combined sensors, we need to analyze the sensor data, evaluate it as a situation and all the information will be added to each photo as properties.

D. Collaborative Reflection with SensorCam

Fleck and Fitzpatrick [16] used SensorCam to support collaborative reflection. In their research, the students were asked to wear SenseCam when they played arcade games. After that, they did a reflection on their learning experiences. They found that SenseCam images were not only used to support memory aids but also can be used as resources for supporting the collaborative reflective discussion. The research also suggests SenseCam has potential to support reflection and that it is more appropriate in learning situation than videos.

III. SenseCam

SenseCam is a prototype device under the development of Microsoft Research (Hodges et al., 2006). In 2009, SenseCam was licensed to Vicon and is available as a product called Vicon Revue. The price of SenseCam is £299. It is a small digital camera that is combined with a number of sensors to help to capture a series of images of the wearer’s whole daily life at the proper time and it can be worn around the neck (Figure 2). Originally this device is designed for memory aid. The sensor data such as movement, light level and temperature is continuously monitored and any sudden changes in any of these triggers the camera. The data in this sensor will be saved as sensor.csv file and photo series separately.

IV. Research Design

A. Comparison between passive mode and active mode

Since this research is based on our previous works that use active mode to register ULLOs, we have to find out differences between active mode and passive mode in this research. We have compared both on features as Table II shows.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Passive</th>
<th>Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of photos within the same time</td>
<td>Many</td>
<td>Few</td>
</tr>
<tr>
<td>Photo quantity</td>
<td>Poor (SenseCam)</td>
<td>Rich (Camera/Smartphone/Tablet PC)</td>
</tr>
<tr>
<td>Recording time distribution</td>
<td>Continuous</td>
<td>Intermittent</td>
</tr>
<tr>
<td>Content completeness</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Content type</td>
<td>Poor (only image)</td>
<td>Rich (image, audio, video, etc.)</td>
</tr>
<tr>
<td>Reflection</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Workload</td>
<td>Low(record)</td>
<td>Medium(record &amp;review &amp;upload)</td>
</tr>
</tbody>
</table>

In this table, the first aspect of comparison is photo. When we use SenseCam to have the passive way of learning, SenseCam takes photos continuously, while in the active mode when smartphone photos only can be taken at the time we want to. As a result, passive mode can bring more photos than active mode. However, because of the storage problem and some other technology limitation, the photo that token by SenseCam has lower quantity, but it is acceptable to be used as ULLO.

The second aspect is learning content. When we use camera or smartphone, many learning contents are logged in the spare time, for example at lunch time. However when we use SenseCam, because the recording is processed continuously, we can get photos whenever in the classroom or on the street. So content is relatively complete. Of course the content type in active mode is richer than that in passive mode because in this research the only learning content that we can get from SenseCam is photo.

The last aspect is about learner. In the passive mode photos are taken unconsciously, while in the active mode learner must take photos consciously. When learners use SenseCam, they must review the whole learning process, and have a reflection what they have seen and what they have learned and what they haven’t learned but they can learn. This process will help learner to remember learned contents.

By this comparison, we understand that passive mode has so many advantages over active mode for the language learning by photos. Only the quantity of photos is low, but it is acceptable. However the biggest disadvantage is workload. SenseCam takes photo continuously. It means when learners review the photos, the number of photos is large and the workload is high. If this workload is reduced, learners can learn language in passive mode easily. This is the key point to use passive mode in language learning. In this research, we are focusing on reducing workload when reviewing the photos and propose a system that can filter the photos to help learners review and upload ULLOs easily.

B. Photo Classification and Sensor Data

In PACALL, we use SenseCam to have a passive capture of learner’s daily life. However, since this device takes
photos continuously, more than 200 photos will be taken in
one hour, and more than 1500 photos in one day. Therefore,
we propose a method to classify these photos by sensor data.

All photos are divided into 5 levels based on importance
– manual, normal, duplicate, shake and dark.
  • Manual: Manual means the photo is taken by
    pressing manual button consciously. When a learner
    takes a photo manually, it means that this photo must
    be important from his point of view. Manual photos
    are selected by the sensor data with flag CAM and
    the capture reason “M” (manual capture).
  • Normal: Normal means the photo is clear and can be
    used as learning log object. After excluding the
    duplicates, shake and dark, left photos are judged as
    normal.
  • Duplicate: Duplicate means the photos are
    duplicated. Duplicated photos usually have same
    conditions. We use CLR, TMP, ACC, MAG and
    timestamp of photo to detect photos that are taken
    under the same conditions and pick out them as
    duplicated photos.
  • Shake: Shake means the photo is blurred. It usually
    happens when the light level is low and the camera
    shakes. The sensor data CLR help us detect light
    level and ACC help us detect camera shake.
  • Dark: Dark means the photo is taken with
    insufficient light and the photo is dark. It can be
detected by CLR data.

V. IMPLEMENTATION

A. System Architecture

![Figure 3. System Architecture](image)

We use SenseCam - Vicon Revue in this research. When
the SenseCam is connected to the computer, if the software
Vicon Revue Desktop is already installed, all photos will be
imported into computer. The location of SenseCam
repository is in the user’s document folder and the name is
Vicon Revue Data. This system is programmed by Java and
runs in Tomcat as a B/S system on the local machine. When
using this system, Tomcat accesses the repository of
SensorCam photos directly and shows them in web browser.
Figure 3 shows the system architecture. All the photos
captured by SenseCam and sensor data are imported into
repository. When a learner uses this system through browser,
server accesses repository and analyzes the photos by sensor
data, then returns the classified photos to learner. Then he
selects proper photos and uploads them to learning log
system through the server. We have a plan to use image
processing technology to detect the photos which contains
faces or texts.

B. User Interface

In order to help learner have better reflection of their
daily life and find important images. We provide a
classification and filter function in PACALL. As shown in
Figure 4, when learners reviewing their life logs photos,
system will classify all the photos into these 5 categories at
the present: including ALL, MANUAL, NORMAL,
DUPLICATE, SHAKE and DARK:

![Figure 4. Interface of browsing life-log pictures](image)

Once the learner clicks a picture, the system will show a
page to view the large picture and help learner upload the
picture to SCROLL. Currently, this page is very simple, and
there are two buttons – “Upload it” and “Close” and one
picture on it. However, in the future, we plan to expand this
page and show the similar pictures from remote server on it.

If learner decides to upload this picture to the server, s/he
can click the “Upload it” button. Then the picture will be
uploaded to the SCROLL system directly and the page will
jump to the learning log registration page (Figure 5). On the
learning log registration page, learner is required to input the
title of the picture. The title is usually the name of the object
in this picture. Location and other options are also supported
on this page. When an object is registered to the system,
SCROLL system will use “organize”, “recall” and
“evaluate” model to help learner remember uploaded objects
and vocabularies. For example, system will remind learner this vocabulary by quizzes.

Figure 5. Registration of learning log

VI. CONCLUSION

In this paper, we discussed how to support language learning using SenseCam. In order to do it, we used SenseCam to capture life-log passively and developed a system named PACALL to help learner to register learning log objects with vocabulary. We have designed a model of learning process in passive capture mode including capture, reflect, store. The PACALL system has been also developed in order to support reflection and reduce the workload of reviewing photos. We found that the SenseCam that originally designed for memory aid can be also used to capture learning log for passive mode to help learners to learn vocabulary. However, it usually takes too many photos, and many of them are duplicated or dark. Therefore, we must introduce other technology to help learners find out important photos. Currently, we are using sensor data to help us do it.

However, SenseCam is not the only device to support passive mode. All the portable cameras that could take a series of photos automatically can be used in this research. Besides, we can also use the smartphone as programmable camera and develop an application to take photos continuously to support passive mode. In brief, this research proposed a new learning method that supports language learning using portable camera in a passive way.

In the future, we also use images processing technology to detect the contents of photos. Besides, current algorithm and user interface also need improvement. In addition, we plan to conduct an evaluation experiment in the near future.

ACKNOWLEDGMENT

This research work was supported by Japan Science and Technology Agency, PRESTO, and the Grant-in-Aid for Scientific Research No.21650225 from the Ministry of Education, Science, Sports, and Culture in Japan.

REFERENCES