香港中文大學

計算機科學及工程學系

Department of Computer Science and Engineering,

The Chinese University of Hong Kong

Research on Tablet

Tablet Apps for Supporting Research Activity

Li Wai Lun (1009615163)

Mo Lim Kwan Pakey (1009618374)

Supervised By

Prof. LYU Rung Tsong Michael

Abstract: To use the tablet devices for supporting Research Activities, we may need to install a number of apps to perform various tasks, such as paper/reference library searching, PDF reader & annotation tool, file sharing, and group communication client. In this project, we would like to develop an application to help in supporting research activities in CSE Department. The final product would include the following three features: paper or reference management, research activity log and communication. A server is required to be used as paper/reference repository, research activity synchronization and instant messaging.

Table of Contents

1. Introduction

1.1 Motivation

Tablet devices gain greater and greater importance around the world. The portable design attracts users for reading information and watching videos. Although there are still limitations of tablet devices and it is not possible to replace the traditional personal computer, the tablets can be further utilized in many aspects.

Figure 1.1 Samsung Galaxy Tab

Figure 1.2 Using tablet to reading e-book

Seeing the advantages of the portability and the ease of reading books, the tablets should be promoted to perform academic objectives. We would like to introduce it to research activities. However, the design of tablet aims at the simplicity. As mentioned previously, the tablet is different from personal computer. Too complex tasks should not be implemented. So what can be done without a substantial keyboard? Yes, the task should be performed without large amount of input data. We noticed that the large screen of tablet eases the difficulties of reading books. Papers are important components of research activity, and that is what can be easily done without a substantial keyboard.

Purely reading paper seems an easy job. It seems far easier than e-books, because e-books may not available on the internet while huge resources of paper are easily “Googled” (since there is “Google scholar” http://scholar.google.com) on the web. Additionally, PDF reader is common installed in tablets.

Figure 1.3 Google Scholar

Nevertheless, what we hope to do is not only giving a PDF reader to read paper. We would like to build up on the existing resources, and develop a tailor-made application for users to view paper easily. For example, are there ways to help us to view the paper easier? Paper seems like ordinary printings with long long pages of words, but they are indeed very different. References are especially for academic papers. Would the application improve the ease of viewing the reference? And though it is mainly used for viewing paper, we believe that more stuff can be completed with the assistance of the tablet. Extra featured are expected to achieve.

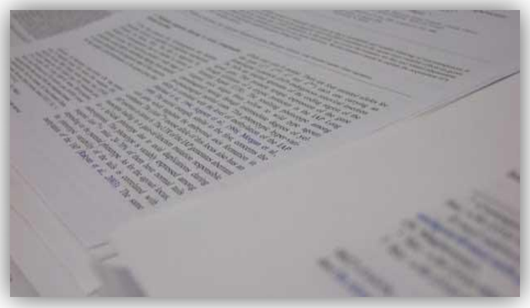


Figure 1.4 Paper in hard copy

We hope that the product constructed can relieve much difficulties of viewing paper and help in research activity. We envisage this product can truly reach the objectives with the rapid popularization of the tablets,

1.2 Objective

Our final project is hoped to develop a tablet application to fulfill the following four objectives.

1.2.1. Make reading paper simply

Reading paper is quite a simple task. However, it is sometimes quite annoying – especially for reading papers. Let's think of the situation of flipping a hardcopy of a paper. Browsing the pages one by one, looking at the dense lines of words and several pieces of images and diagrams, everything runs smoothly. However, some not special but common nagging digits start to annoy us quickly and frequently. The digits within the [ ] bracket show its prevalence in the whole passage, if you would like to know what is that means, you should flip to the last page to have a note – usually you know nothing because the line tells you a string of unknown article names and strange names of researcher – to have a look. Repeat the procedure for 10 times or above – see? Yes, it is annoying!

The above case seldom appears in daily life indeed, since we are not common to refer to reference in daily reading. Nonetheless, the reference weighs heavily in the scholars' writing. Tones of unknown phrases and findings may be shown and these are all valuable to be referred.

Of course, the problem is easily solved using the PC by opening the file twice. The problem can also be addressed on the hardcopy by tearing the pages out. But we would like to make everything simpler, to view the reference or annotation effortlessly by just one click. If the tool is powerful enough, we expect the tool can directly show us the referred articles too. All these helps make things simpler.

1.2.2. Improve efficiency

Obviously, the tablet devices aim to help saving much more time. We can treasure our time by doing vivid tasks anytime and anywhere – and what the users focusing is not just “what they can do”, but “how they can complete their tasks simply and efficiently”. For the word “simple“, we would like to define the word as doing much more things with minimal work done on it. We try to achieve this “simple” which has been analyzed in the last paragraph.

So what is “efficient” mean? To make it more materialized, we make it related to time consuming. Time is definitely our major concern, and finish the tasks in a simple way is one way to increase the efficiency. Besides, efficiency requires strong supports behind simplicity. Strong, powerful and well-designed database and libraries can greatly devote to the enhancement of the efficiency, which let you locate what you need within the minimal time spend.

In fact, efficiently cannot be easily measured. Comparing to the effectiveness, we may quantify some targets to see whether it can be completed fully or partially, or in how much percentage it completed, while efficiency is heavily depends on many different tiny stuffs. We hope to achieve this by having research on the methodology of reading papers and design our system carefully.

1.2.3. Exchange ideas easily

Research papers, to a certain extent, are the concepts of spreading the themes and theories – or we name it “ideas” – around the world. Research activity involves a great scale of exchanging ideas in principle, in order to invoke much more new inventions. Sharing papers are what we can commonly do nowadays, yet would there be rooms to improve it? Can the communication be further enhanced? If yes, then how?

Building upon reading research papers, we may add features to further boost the circulation of ideas. We should have experienced such kind of situation that we would have some questions or challenges to raise, or we would have our own ideas or different point of views concerning certain point described about the topic. These should all be considered as academic sharing of ideas and they may stimulate us to rethink deeply. Sparks invoked may contribute to new ideas too.

1.2.4. Minimize workload of managing resources

http://www
Resolution 4000x3000 px - Free hi-res JPG file download - www.psdgraphics.com

Figure 1.5 Internet contains lots of information

Information is flooding, especially on the Internet. Millions and millions of information floods when we look for a certain topic. Perhaps it helps us to include as much as information as can, while we often spend quite a lot time for filtering and locating what we actually need. Duplication of information always occurs, which prevents us from extracting the pieces we want.

To avoid this, good management may address the problem. If the system is clever and mature enough to manage paper neatly and avoid the repeats, workload of researchers can be minimized. Much more time can be spent to concentrate on the research topic. All these help are more than just saving time of the researchers and scholars – these also save the resources of our computer to achieve green computing in some sense.

Figure 1.6 Green Computung

1.3 Project Overview

To achieve our objectives, we should implement our project steps by steps. This report will discuss each part clearly, and here will give a whole picture of our project.

1.3.1 Research on relevant topics

Chapter 2.2 gives some analysis of our survey which has been conducted before starting our project with coding. Following are the area we would like to study:

A) ***Format of research paper : what can be shown to user?***

Understanding the skeleton of papers makes the filter of information run smoothly. Some parts are essential to identify the uniqueness of the paper and building index of the paper library, just like the title and authors of a paper. But would there be more information can be easily recognized by the server, so as to improve the accuracy of storing paper or showing useful information to the users?

B) ***Format of PDF files : what PDF does?***

It is not surprising that most papers are saved as PDF format. Printing as PDF prevents it from editing or mix-up due to careless mistakes and make it portable to various computer working environment and operating system. Of course, we are not trying to analyze "why PDF is used", but "what PDF does". Format of a paper is readable for human beings, while PDF format is known by computer. Studying more and getting familiarized to this format would lead us to better development of the application and the buildup of server.

Figure 1.7 A PDF file

C) ***Occurrence of the phrases : how to make the paper readable for computer?***

Besides understanding the paper format, the phrases used by the papers are important for teaching the system too. Keep in mind that papers are designed to be read by human. Languages of people, unlike the languages of computers, can vary within a certain range. Different wordings (even with "syntax errors") may tell the same ideas. Analysis of the occurrence of some key phrases improves the accuracy for computer to locate the data.

D) ***More features can be utilized : how to achieve more?***

Not just referencing is significant; we should investigate more characteristics to enrich the application and accomplish our aims. For example, are there any communication features have been implemented in the tablet that can be made good use of to facilitate the exchange of ideas? Or are there any open source software can be imported to help the development runs smoothly? Just have a research to find out the answers.

E) ***Tools help to build the app: any API available?***

Building an application on the tablet devices is not easy. Android platform, to a certain extent, is quite a new platform to us. It is quite hard for us to start to learn to write a PDF viewer from knowing nothing. We would like to find out if there is any APIs can help us to build our application smoothly. After all, learn to build a PDF viewer is not our ultimate goal. We would like to save time so as to run more towards our goals.

1.3.2 Design of the whole system

Chapter 3 concludes the overall design of our system. Our system includes server side and client side development, that both sides heavy workload to handle.

For the server side, it has been developed using PHP web server language. The server side aims to do every parsing and analysis on the PDF files and manage the database. PHP is chosen because PHP has strong libraries for reading the PDF files. This make the analysis on the parsing paper runs smoothly.

For the client side, our application has been implemented on the Android platform. Originally it is expected to be implemented on the iOS. Yet, due to the limitation of hardware resource that we do not have iOS development platform, we give up the iOS and switch to Android platform. The Android application is built by Java language and the development is comparatively portable to the application runs on iOS. Reiterate that our focus for once – we would like to make research activity simpler on tablet devices but not on MACs only.

Figure 1.8 Android platform

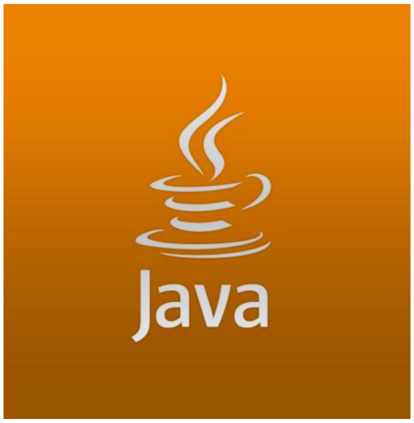


Figure 1.9 Android platform uses Java to develop

To ensure the progress of the project, the project is roughly divided into two parts. One student focuses on client side and the other student focuses on the server side.

There three main modules to achieve our goals:

1) Paper management
2) Communication
3) Research Log

Further explanations on design and implementation are included in Chapter 3 and 4.

1.3.2 Division of works into two phases

To make the tasks clearly, we divided the works into two phases. We have tried to conduct a comprehensive survey and implement some key features of the applications during the first phase. In the second phase, we would aim at increase the accuracy of the application and make fine tunes to improve the efficiency.

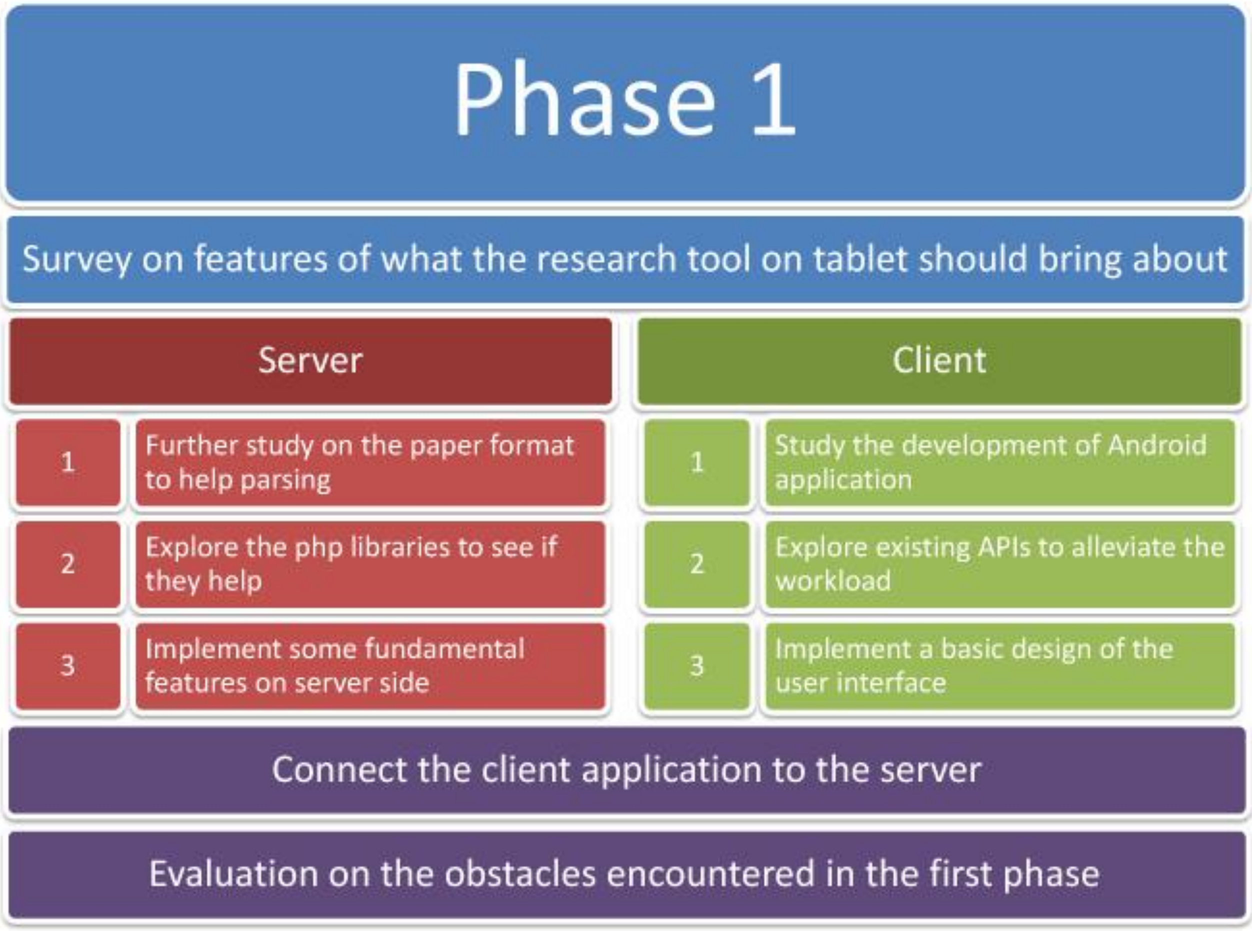

Phase 1
Survey on features of what the research tool on tablet should bring about
Server
Client
1
Further study on the paper format to help parsing
1
Study the development of Android application
2
Explore the php libraries to see if they help
2
Explore existing APIs to alleviate the workload
3
Implement some fundamental features on server side
3
Implement a basic design of the user interface
Connect the client application to the server
Evaluation on the obstacles encountered in the first phase

Figure 1.10 Works in phase one

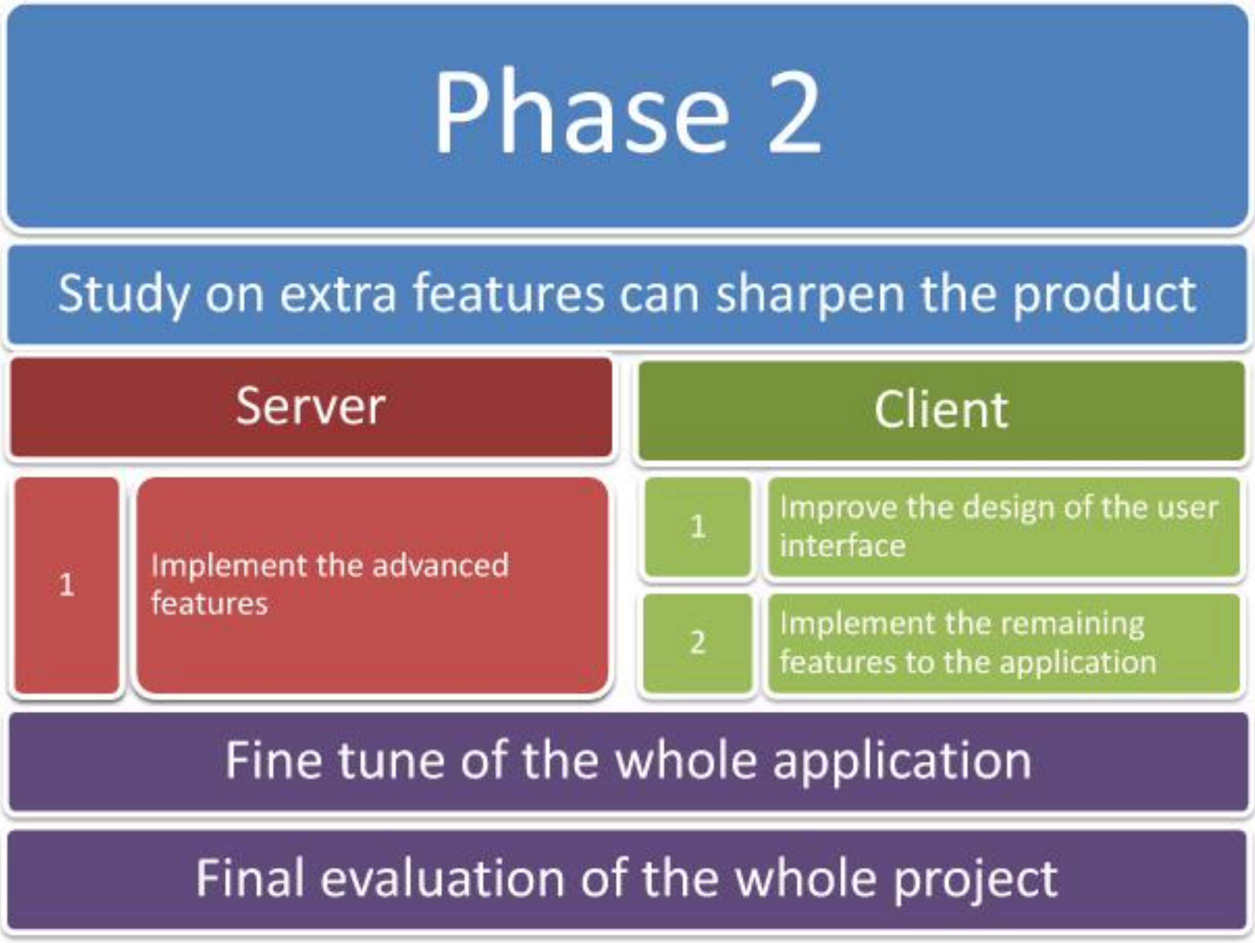

Phase 2
Study on extra features can sharpen the product
Server
1
Implement the advanced features
Client
1
Improve the design of the user interface
2
Implement the remaining features to the application
Fine tune of the whole application
Final evaluation of the whole project

Figure 1.11 Works in phase two

2. Background

2.1. Raising population of tablet device

Figure 2.1 iPad

When it comes to tablet devices, they are seldom used in academic aims. Chasing back to the time when Steve Jobs presented his “truly magical product” – iPad[1], he defined it with many functions: “browsing the Web, sending e-mail, sharing photos, watching videos, listening to music, playing games and reading e-Books”. Since this most popular tablet in the world released, we almost found most of users treat it just as a smartphone with a larger screen. Surely most of features of the tablets, not matter iPad or Android tablets, can be found on the iPhones or Android smartphones. The main differences between the tablets and smartphones are the calling and SMS functions and the screen size.

Figure 2.2 Screen Size difference between smartphones and tablets

However, we thought that these portable devices with a larger screen size should not be treated like an alternatives of smartphone. They are truly different, and we should fully utilize the differences. Once we noticed what Steve Jobs presented, one point should be focused – the reading e-Books function of the iPad. As one of experienced smartphone users, this point woke us up. Among the abovementioned features, we are comforted to perform most of them. We may browse the web for dealing with some urgent stuff or finding emergent information, yet we seldom use the phone to read books. Of course, reading books is possible on phone devices, while the small size screen is not preferable compared with the tablet. And this is what we looking for – the absolute difference between tablet devices and the small touchscreen gadget. On recognizing the difference, we decided to develop a tool to make good use of it.

2.2. Lack of academic application for tablets

We would like to introduce the tablet devices to academic aspects. Take a look at the Android market. Click on the category tab of "Books and References" – the one seems related to academics research – and check if there is anything helps in research or paper referencing. However, we would get a list of e-books or e-book readers. Even we click onto the tab of "education", we also get similar results. How about we type "research" and some relevant terms? Most of the results show application related to business or stocks. One result of "research paper" catches our eyes – well, the app "Writing Your Research Paper" seems just an E-book, not an app help me to writing the

Figure 2.3 Android Market is lack of academic application

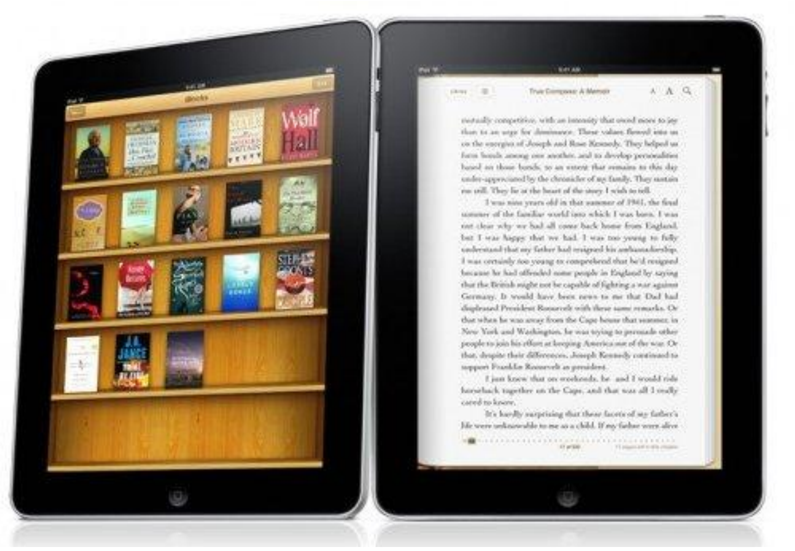
paper.

The case may not such extreme indeed. If we search “research Android App” on Google, meaningful results can still be found. However, there are still few meaningful results. The two applications we found is just like a database and libraries for us to search for materials [2, 3]. There is lack of strong and useful applications for supporting activity.

2.3. Reading paper takes the advantage of tablets

As mentioned in the motivation part, complex task should be avoided violating the original concept of the tablet design. This tells reasonably that nowadays the application available only provide the function of searching papers or research resources for users.

Taking of advantages that tablet can read books and keeping the origin of tablet, we would like to develop an application related to reading research paper. Reading paper seems quite a simple task. It seems far simple than e-books. Books may not be available on Internet, while there are great resources of paper on the web.



Yet when we read papers, no matter the hardcopy or softcopy on the tablet, we find it annoying to clip the pages over and over, to grab the reference and chase where the ideas come from. The clipping maybe not the great problem, but the referencing can surely be done better and better. Somehow, just clipping to the last page we just know more about a list of strange titles and authors. What we want is the paper! And that is what we looking for – a paper reader with automatically referenced and linked to others.

Figure 2.4 The large screen of tablet is good for reading books and papers

2.4. More features beyond reading

If a platform is designed for reading research paper, then could it be further extend to a research tool? Quite more than reading paper and required little data input. We would like to improve and sharpen the product in order to realize a helpful research tools that truly cater the needs of researchers. This goal prompts us to think more, that will be further clarified in the research part and the design part.

3. Research on relevant topics

Research tools are not rare in personal computer. What we are now seeking for is not just software of a research tool, but an application specifically catering the aim of reading paper and using the tablet devices. Yet, Scholars' writings, though have been quite restricted in the writing formats, vary much because they are written by human. To make the paper more manageable, we should first start to see the difference and tell the computer what parts to look at. There are four main topics we have conducted relevant surveys.

3.1. Format of research paper : what can be shown to user?

To start with, we should clearly define what the research paper is. Somehow our product may not powerful enough to recognize all PDF files without a standard. Quoted from Wikipedia: "Research paper, (also called scholarly paper), which is published in academic journals and contains original research results or reviews existing results" [4]. Under such definition, we try to restrict the paper format fulfill the rules of the publication of the academic journals.

Firstly, generalizing on the ideas of the some reference page of the University of California [5] and the Rice University [6] from USA, the research paper have to include the title, author(s) (and maybe the contact information), the abstract and the references. Of course there are introduction and contents, but abstract is enough for achieving our goals.

Second, we conduct a rough survey on a pool of research papers. The pool of papers contains 169 PDF files and 149 of them fulfill the definition of a scholarly paper. We analyze the position of title and author, the appearance of abstract and keywords, and how the reference show. Nearly 95% of the papers have an abstract, and 90% of them start with the word “Abstract”. The keywords appear in 50% of the papers. Almost 95% of the references indexed by the [ ], while few of them using other format or with no index (just show a list of references). Some of research papers consist of unstandardized format, such as author goes before title, or in the form of report. 15% of them start with a cover page which is not a standard.

To conclude, over 80% of the scholar papers in the research pool follow the format that includes title, authors, abstract, introduction, content and references. Though it may not be the real statistics of all the papers available on the web, it is worth to have a try in parsing papers in the format investigated.

3.2. *Format of PDF files : what PDF does?*

According to Wikipedia: ”Portable Document Format (PDF) is an open standard for document exchange. This file format, created by Adobe Systems in 1993, is used for representing documents in a manner independent of application software, hardware, and operating systems.”[7]

Structure of PDF files

The PDF files consists of eight primitive objects including Boolean values, numbers, strings, names, arrays, dictionaries, streams, and the null objects[8]. PDF contains everything as objects, to standardize different data printed by various files originally. Take text as one of the examples. Adobe creates an element for the text and makes it as stream [7]. The primitive type “streams” is used for storing a large amount of data. The element records not only the text wordings, but also the font size, the font family, etc. PDF also records the layers of each object.

Metadata

There is one more significant feature of PDF files. There is metadata of PDF files. The metadata records much key fields of the title, author and update dates. The metadata seems to be powerful, which could help us a lot in completing or goals. PDFmeat [9] is a powerful open-source tool to acquire the metadata of PDF Files and we have done a research on the metadata using the PDFmeat. Though there are few errors, over 70% of papers are correctly located.

PDFmeat

PDFmeat stands for PDF Metadata Acquisition Tool [9]. The tool provides great support for analyzing the metadata in PDF header and managing the PDF files. It aims at convert the metadata to BibTex. To run the program, PDFmeat provides a user-friendly GUI which embedded in the Firefox browser. On click it will generate BibTex data to catch the metadata of PDF file. However, the GUI is not feasible for realize our aim. We would like to have the CLI program to run and catch data. We have done many searching and refer to the official page for many times but we still can access the PDFmeat successfully. Unless we can achieve it successfully, we are impossible to use the metadata.

PHP Library - PDFlib

There were many share libraries available for PHP, and some of them related to PDF files. Nonetheless, most of them are aimed to create PDF files rather than parsing / reading the PDF files. Filtering those useless, we focus on the PDFlib. [10] The PDFlib get various value like font size, position in X and Y, page numbers etc. [11] All these features shall be taken at note for implementation.

Figure 2.5 Icon of PDFmeat

Figure 2.6 Icon of PDFlib

Figure 2.7 Icon of Xpdf

Xpdf – pdftotext

Though PDFlib can read the PDF files, most of the functions build-in are expected to be used in creating or editing PDF files. Like the functions in PDFlib, these is no any function can directly extract the text of the existing PDF. Parsing PDF text may be not relied on the PHP library. Nonetheless, there are other open-source programs help to extract the text in the PDF, like PDFBox [12] and Xpdf [13]. We would first implement Xpdf for experiment.

3.3. Occurrence of the phrases : how to make the paper readable for computer?

To study how to parsing the paper, we analyze the phrase occurred in the paper which may make the paper more understandable to the computer.

Title

95% of the papers do not have the “title” word to indicate the title of the paper. Usually the font size of the title is the largest word appeared in the whole file. Few of the titles appear after the author name. Titles may not the first line of the paper, because there may be headers of the page or some further information.

Author

There are a huge number of papers which are written or collaborated by more than two authors. Besides the name of authors, more than a half of papers we surveyed contain a line of information and contact information about the authors. For those do not show author information at the top of the paper, the information may be added at the end of the paper (after the appearance of references).

Abstract

More than 80% of abstract is written following the presence of the "Abstract" word. Some of them are written directly after the name / the contact information of the authors. When the abstract ends, introduction or keywords will be presented.

References

Nearly all the papers have the "References" or "Bibliography", while there is a small variation between paper and papers. Statistic shows 90% of references have shown in the index form [ ]. The rest shows without [ ] but numbers only, and some of them without indexing. Those lists of references without indexing will not be considered in this project because of the meaninglessness.

Keywords

Presence of keyword is not a necessary part of a standard scholarly paper. Yet, after consideration, we would like to check if there is any presence of the keywords column. If there is, the keywords will be recorded in our database to assist the searching of paper. The keywords usually appear with the presence of different noticing words like "keywords", "index terms" or "general terms".

3.4. More features can be utilized : how to achieve more?

We have done some searching in further exploration on the functionality of the application. We tried to start with our objectives set, and some sparkles invoked.

To improve the efficiency?

As undergraduate students, we have never got a chance to perform a formal and comprehensive research on any topic. However, we believe that every time we do research and investigation, we need plenty of information and bibliography to support what we are saying. Quite annoying scenario arouses – when we try to look for some new information, we read and find some unknown words and terms. Then we try to make it clear, and we search for other pages of information and more and more terms will come out. Just like some kind of recursion. Finally, after we open for hundreds pages of reference, we almost forget what we asking for!

Thus, we find it will be valuable to have a research log for each user. User information are recorded in our database and users have to login to use the application. Every time they read a paper, there are their own logs to save the record. Even though they go though from one paper to another for hundreds of times, they can still find what they originally look for and can more handily filter what they need.

To exchange ideas?

When it comes to communication, it is not surprised to think of the idea of instant chatting. Chatting or messaging helps for better discussion on a topic, while some views are worth arguing. It seems reasonably, but impractical indeed. Firstly, due to absence of substantial keyword, the application is designed for as less input as possible. Discussing on instance messages requires a certain amount of input data. Moreover, the main advantage of a tablet is the portability. Serious discussion seems not meaningful on the railway journey, or if it does, phone calling would be more practical and efficiency.

We want some features are not redundant but useful. Does no instant messaging mean there is no need towards exchange of ideas? Definitely not. Ideas and views are aroused during the reading of other’s writing, and we may give some response to the writer, or share what we think to a specific point. Yes! The comments.

Adding comments onto the paper, when someone has his own view, his can put it down with few sentences. The comments will be saved when anyone reopen the file. Every users can comment and view what others thinking by just one click. Exchange of ideas can hence be achieved, handily.

To minimize workload of managing resources?

Papers are our main characters of the show. We cannot implement an application without papers. We are actually not worried about functionality and the implementations, because once the algorithm and functions work successfully, it seldom have bugs. (Of course further analyze and maintenance is needed to improve the accuracy.)

However, even if the application is precisely implemented with perfect algorithms, the renewal of paper should be carried out everlastingly.

Could the paper library enriches and renews itself automatically? We raise this question and have a draft answer. We may embed a web browser into our application. When the user cannot find what he wants in our existing database, he can open the web browser to look for new information. If he would like to download a new paper, the library would save it add do checking and parsing. If there is duplication, the new file will be abandoned. Otherwise, the paper will be parsed for every data requires and a new entry will be added. With this feature, the maintenance and management by human being can be greatly reduced.

We noticed that there is a http://scholar.google.com/ which is a strong library of scholarly papers. We put this back and will look if we can take this advantage in the Project Phase 2.

3.5. *Tools help to build the app: Any API available?*

To alleviate our workload, we choose to use the API of PDF viewer. We have found several in the Internet, such as 'APV - Android PDF Viewer', 'vudroid', ebookdroid' and 'apdfviewer'. We have tested their performance by reading different types of PDF files.

	APV	vudroid	ebookdroid	apdfviewer
Only text	2	3	1	4
Only image	2	4	3	1
Both text and image	1	3	2	4
Small file size	1	3	2	4
Very large file size	2	4	1	3

(1 is the best, 4 is the worse)

We have tested the four APIs by five different types of PDF files, one is only containing texts, one is only containing images, one is containing both texts and images, one is of small file size and the final one is of very large file size.

APV - Android PDF Viewer

'APV' has the best overall performance, and does best when reading PDF containing both texts and images with small file size.

vudroid

'vudroid' does worst when reading PDF with only images and of very large file size. Its overall performance is the worst among the four APIs.

ebookdroid

'ebookdroid' has very good performance when reading PDF with only texts and of very large file size. However it does not perform good when reading images.

apdfviewer

'apdfviewer' does the best when reading images, however, it has the worst performance when reading texts.

Conclusion

Although 'ebookdroid' does better than 'APV' when reading PDF with texts only and of very large file size, a paper contains both texts and images, and its file size would not be very large, we believe that 'APV' is the best API to read papers, and we choose to use 'APV' in our project.

4. Overall Design

4.1. System Architecture

Our system consists of server side and client side for better communication and management, and for better collaboration with the database.

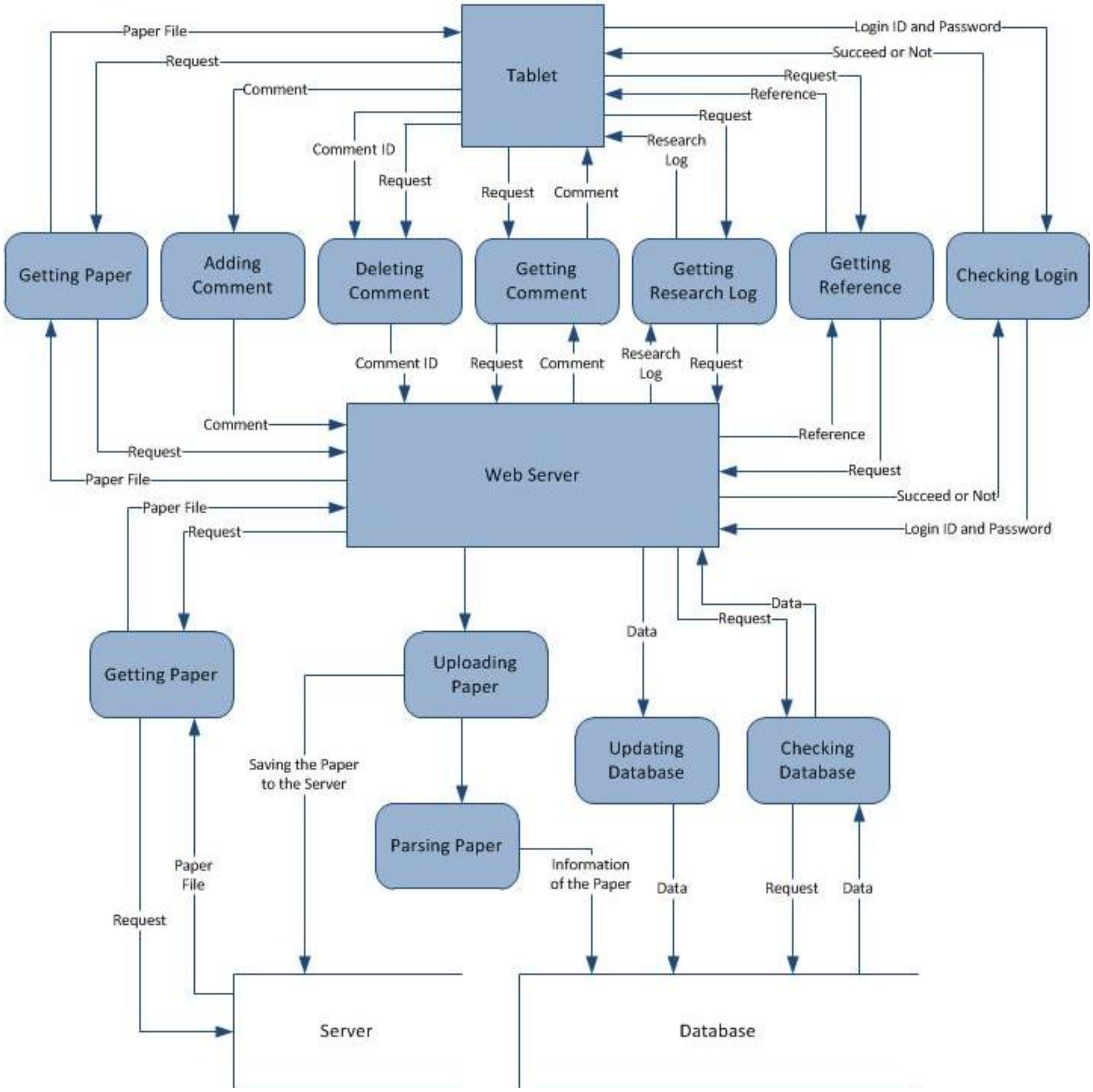


Figure 4.1 Architecture diagram on server base

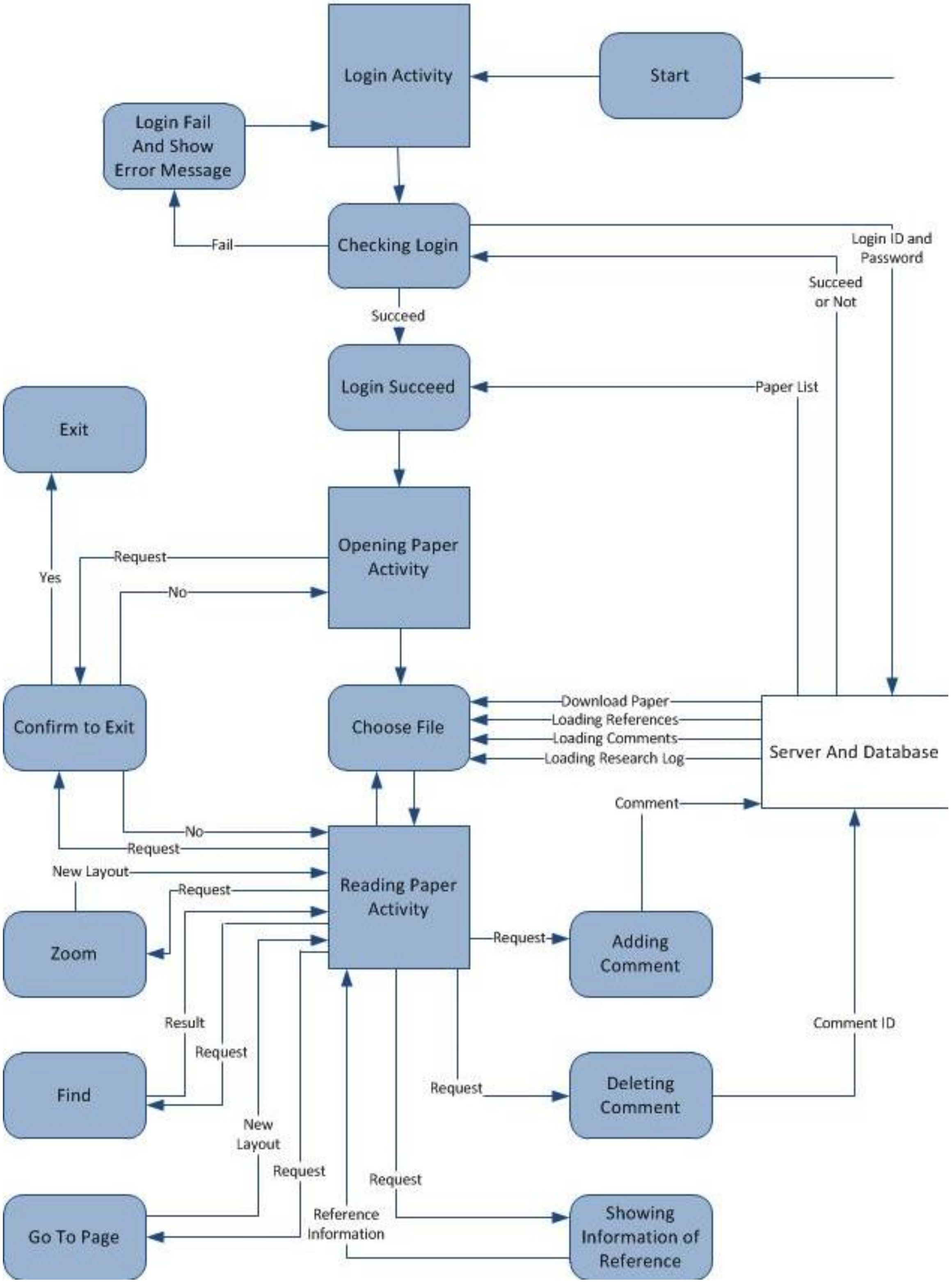


Figure 4.2 Architecture diagram on tablet base

4.2. Module Description

To make the tasks simpler, we have divided our works into three different modules.

4.2.1. Paper Management

Paper management plays a great role in our project. On one hand, well paper management helps researchers and scholars to locate their target resources and information easily. If the library is strong enough, the researchers can save much more time and improve the research efficiency. Just say, how to guarantee each paper is unique? Even when we type some keywords onto anyone of search engine, a list of repetitive results is not surprising. Yet most of the search engine is not clever enough to ease out those duplicated results. Extracting the information cost much time when we try to perform search. Therefore, we would try to alleviate the problem of duplication – try to store only one file despite different file name of same paper, or save the same paper only if the version is different.

On the other, the available paper resources online are vivid. Of course it is impossible to grab all papers from the internet once the application released, yet it is our ultimate goal to keep as much file as possible in our library. The tool may not powerful and intelligent enough to download all files online automatically, but we hope to achieve the task with the help of the users. The application on tablet embeds a web browser and let the users to search on web. When a paper is downloaded, the app will raise a signal and send the link of the paper to the server. After comparing the paper newly downloaded is unique and has not been saved on the server, a new file will be included and the source path is recorded to avoid duplication in the future. The library can hence expanded without heavy maintenance and involving a great amount of manpower.

Moreover, assisting the users to view the paper easily is one of our main objectives. To achieve this, other than show the references in just one click, we would also like to introduce some linkage between papers. While the user takes a look at the reference of some specific points, a hyperlink is provided to link the reference paper and the user can therefore have further investigation on the related topics. If the idea can truly be realized, it can surely minimize much more time for the user to do searching and collecting.

4.2.2. Communication

The concept of exchange of ideas has been discussed previously. We expect no any instant messaging because we hope to avoid inputting great amount of data without a substantial keyword, and we believe the user would relied on other more effective and efficient communication method rather than using our application.

We will let the use to post his comment onto the paper. The comment can be added onto anywhere of the paper. There will be a button or a flag for identification of the existence of the comment. Every other user can view the comment and can give response to the point too. The users who have echoed their ideas before would receive a notice that someone posted response. These stimulate the discussion by a small scale of input data. Since all the comment will be saved, the discussion can be shared to many users and the circulation of ideas can be achieved.

4.2.3. Research Log

Researchers face a huge obstacle that has been noticed for many times before: the huge amount of information flooding into our mind. When we go through for a referenced idea, we flip over the pages for the first term. To understand the meaning of a theory or a term, we shall then refer to one another. We flip the pages over and over, and for many times after we almost forgot what we are reading at first.

The research log shows everything that each user has done. Every single paper that his has read is record and sorted by the opening time. This allows the user to chase back what his is originally looking for. Bookmarking will be considered as one of the features. If the user finds some papers that are truly inspiring, the user can put a star on it. Then his can reopen it by referring to the log.

We would like to make the research as strong as possible. Even the time the user posts a comment on a specific paper would be record. And if the comment posted has been replied by someone else, the user would be told by the research log too.

4.3. Functional Specification

This section will show general functions that the system can perform.

4.3.1. File Exploring

User can click on the file explorer to open the research paper that he needs. According to the interface design, after the user choose the file and the paper will be shown on the right hand side. A searching function will be available for user to search keywords, title or authors.

4.3.2. Paper Parsing

Once the paper is added into the server, the paper shall be parsed in detail and all data shall be recorded in the database for better management. All fields will be parsed once to check if it is duplicated with existing papers. If the paper is new, all fields of data will be stored in the database for future uses.

4.3.3. Referencing

The newly added paper will be parsed to check if the reference index available. If yes, a list of references will be parsed into the database. The reference index throughout the whole passage will be highlighted and linked and user can refer to that point by a click. No matter the paper have indexed reference, a checking and linking procedure will be performed on the references in the list. If a piece of reference points to another paper that exists in our database, a linkage will be built for a faster referencing.

4.3.4. Library Renewal

A basic paper managing panel should be provided for uploading paper onto the server. The panel is monitored by administration and the uploading process is manual. We would like to introduce a browser to be embedded within the application. Even if the introduction is fail, we still hope the application can be called automatically when a PDF file downloaded. The application will first check if it is a scholar paper. The paper will be parsed and uploaded to the server. All these are expected to be implemented with manual control.

4.3.5. Posting Comments

User can post comment onto the paper. Different actions can be done, such as viewing, deleting and replying. The important stage is how can those have commented or responded to the specific point can be informed to extend the discussion. This part will be collaborated with the user research log functionality.

4.3.6. User Login

The application requires user to login. User would have his or her own profile for other users to refer to. Only little basic information is needed, like the full name and simple description of the user.

4.3.7. User Log

Each user has his user log for recording the history of browsing each paper so as to make the process of referencing run easily. Bookmarks are allowed. The log also contains the history of posting comments and notifies the user if there is anyone responded to the comments he posted.

4.4. Development Platform

Server side

The server is built by php5.3.6 with Ubuntu. Due to the research related to the PDF, there is PDFlib in PHP to read PDF. PDFlib supports only the version newer than PHP 5.0.Xpdf and PDFBox are also need to be installed in the server for parsing the text in the PDF.

As for the pdfmeat, since we still cannot implement the CLI successfully, we would like to implement the project without it in phase1. If the self-implemented result is not satisfied, the pdfmeat will be introduced in phase2.

Client side

Since the Android platform is finally selected by the end of September, the latest version of Android at that time was chosen. Our project will be implemented in the Android 3.1 (the version which Samsung Galaxy tab 10.1 implemented [14]).

5. Detail Design

This chapter describes the design of the whole system in details. The shows all tiny features we would like to implement. In this part, we would state clearly whether the features will be implemented in phase 1 or 2 of the project.

5.1. Module 1 : Paper Management

This module performs four main functions.

5.1.1. File Exploring

The first major concern is on opening files.

Show list of paper (Phase 1)

The user interface is composed by two parts. The sliding door on the left hand side is a file explorer. The list shows all the papers available on the server. In the first phase we will show the file name only. In the second phase we may consider to show the paper with title or author.

Additional feature (Phase 2)

On the first click the abstract of the paper will appear. The paper will be shown on the second click.

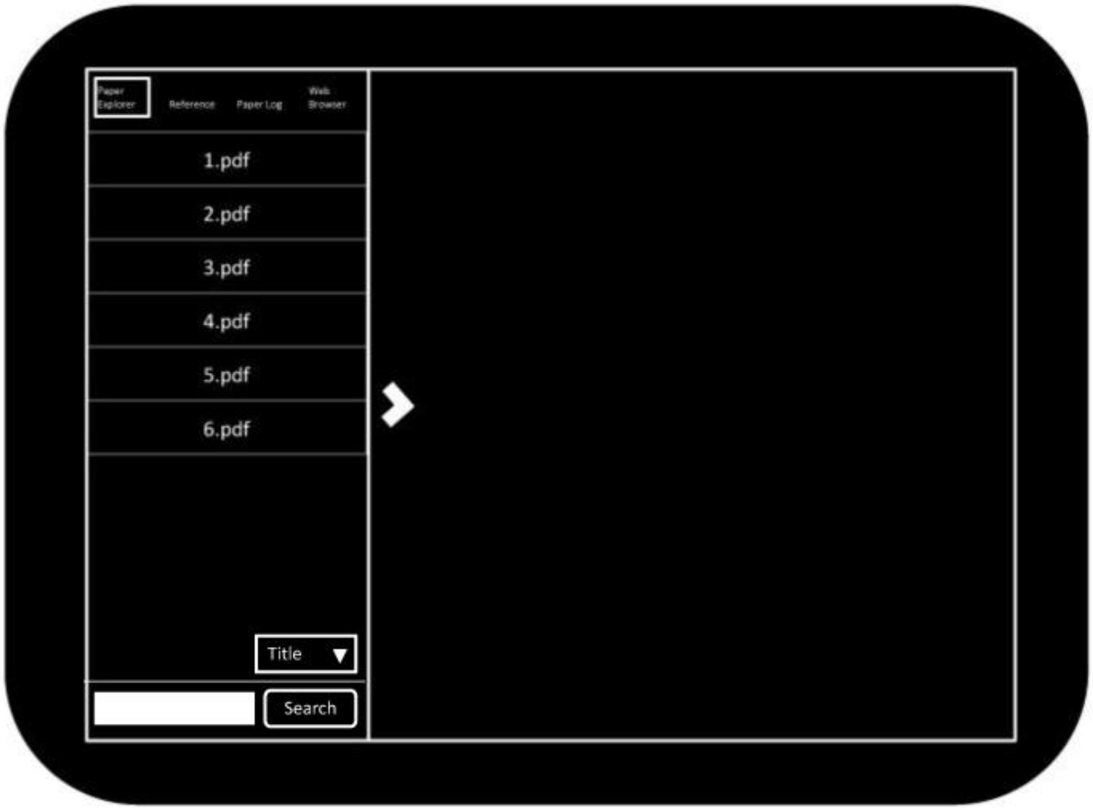
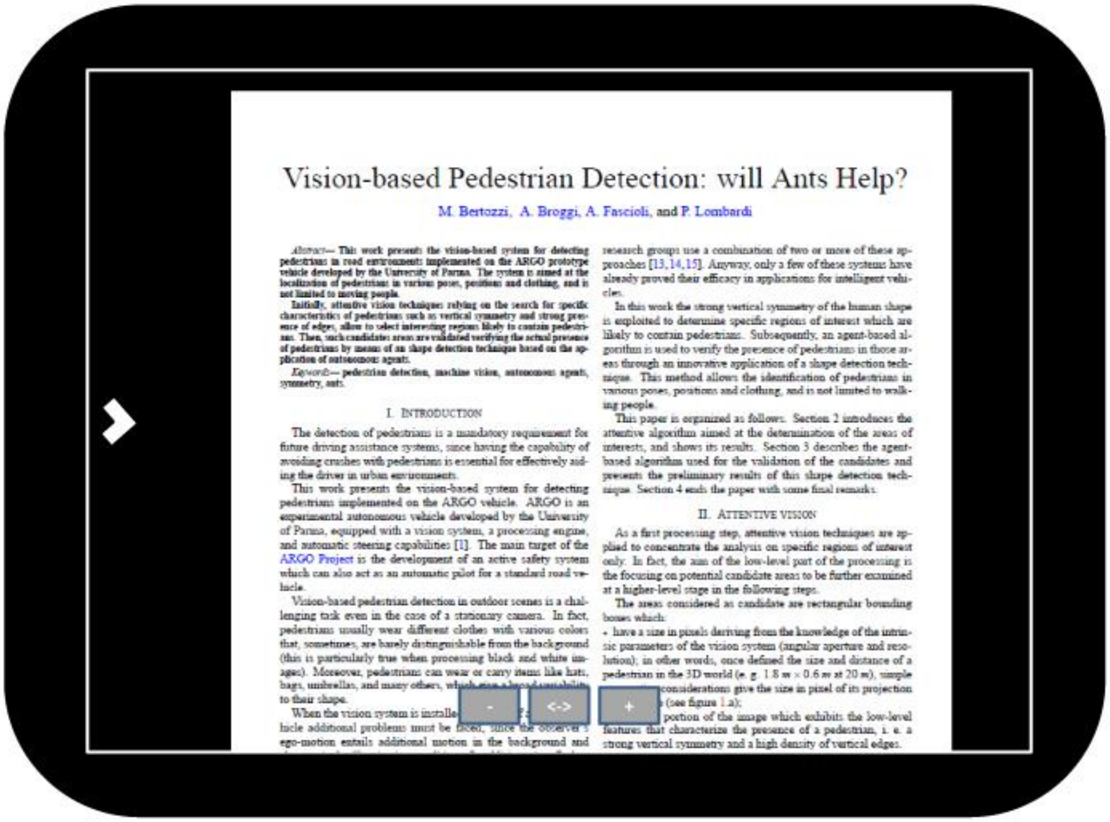


Figure 5.1 Basic UI design on the application

Show paper (Phase 1)

On clicking the file listed on the explorer, the paper will be shown on the right hand side. The PDF file should be allowed to zoom in or zoom out. It allows the user to perform searching on the content of the file. Directly going to a certain page is also allowed.



Vision-based Pedestrian Detection: will Ants Help?

M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi

Abstract— **This work presents the vision-based system for detecting pedestrians in road environments implemented on the ARGO prototype vehicle developed by the University of Parma. The system is aimed at the localization of pedestrians in various poses, positions and clothing, and is not limited to moving people.**

Initially, attentive vision techniques relying on the search for specific characteristics of pedestrians such as vertical symmetry and strong presence of edges, allow to select interesting regions likely to contain pedestrians. Then, such candidates areas are validated verifying the actual presence of pedestrians by means of an shape detection technique based on the application of autonomous agents.

Keywords— **pedestrian detection, machine vision, autonomous agents, symmetry, ants.**

I. INTRODUCTION

The detection of pedestrians is a mandatory requirement for future driving assistance systems, since having the capability of avoiding crashes with pedestrians is essential for effectively aiding the driver in urban environments.

This work presents the vision-based system for detecting pedestrians implemented on the ARGO vehicle. ARGO is an experimental autonomous vehicle developed by the University of Parma, equipped with a vision system, a processing engine, and automatic steering capabilities [1]. The main target of the ARGO Project is the development of an active safety system which can also act as an automatic pilot for a standard road vehicle.

Vision-based pedestrian detection in outdoor scenes is a challenging task even in the case of a stationary camera. In fact, pedestrians usually wear different clothes with various colors that, sometimes, are barely distinguishable from the background (this is particularly true when processing black and white images). Moreover, pedestrians can wear or carry items like hats, bags, umbrellas, and many others, which give a broad variability to their shape.

When the vision system is installe hicle additional problems must be faced, since the observer's ego-motion entails additional motion in the background and

research groups use a combination of two or more of these approaches [13,14,15]. Anyway, only a few of these systems have already proved their efficacy in applications for intelligent vehicles.

In this work the strong vertical symmetry of the human shape is exploited to determine specific regions of interest which are likely to contain pedestrians. Subsequently, an agent-based algorithm is used to verify the presence of pedestrians in those areas through an innovative application of a shape detection technique. This method allows the identification of pedestrians in various poses, positions and clothing, and is not limited to walking people.

This paper is organized as follows. Section 2 introduces the attentive algorithm aimed at the determination of the areas of interests, and shows its results. Section 3 describes the agent-based algorithm used for the validation of the candidates and presents the preliminary results of this shape detection technique. Section 4 ends the paper with some final remarks.

II. ATTENTIVE VISION

As a first processing step, attentive vision techniques are applied to concentrate the analysis on specific regions of interest only. In fact, the aim of the low-level part of the processing is the focusing on potential candidate areas to be further examined at a higher-level stage in the following steps.

The areas considered as candidate are rectangular bounding boxes which:

- have a size in pixels deriving from the knowledge of the intrinsic parameters of the vision system (angular aperture and resolution); in other words, once defined the size and distance of a pedestrian in the 3D world (e. g. 1.8 m × 0.6 m at 20 m), simple considerations give the size in pixel of its projection (see figure 1.a);

portion of the image which exhibits the low-level features that characterize the presence of a pedestrian, i. e. a strong vertical symmetry and a high density of vertical edges.

Figure 5.2 UI design of showing a paper

Sort of paper (Phase 2)

For locating the file accurately, there should be a number of sorting methods provided. For example, it can be sorted by title, authors or creating time.

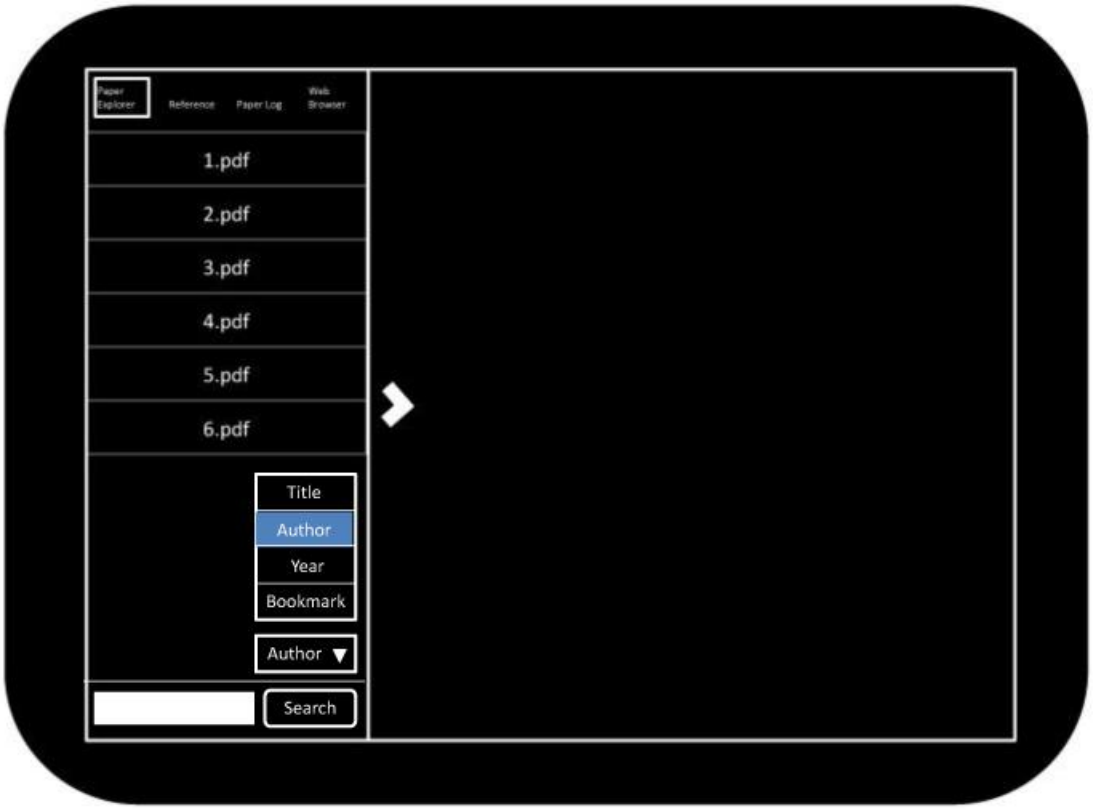


Figure 5.3 UI design of sorting papers

Search for paper (Phase 2)

The searching function should also be available. This makes the process of opening the correct paper run faster and more efficient. The searching will echo the keywords in title, authors, key terms and abstracts of the paper. When the list of searching is shown, abstract will follow the title and authors to provide sufficient information to users.

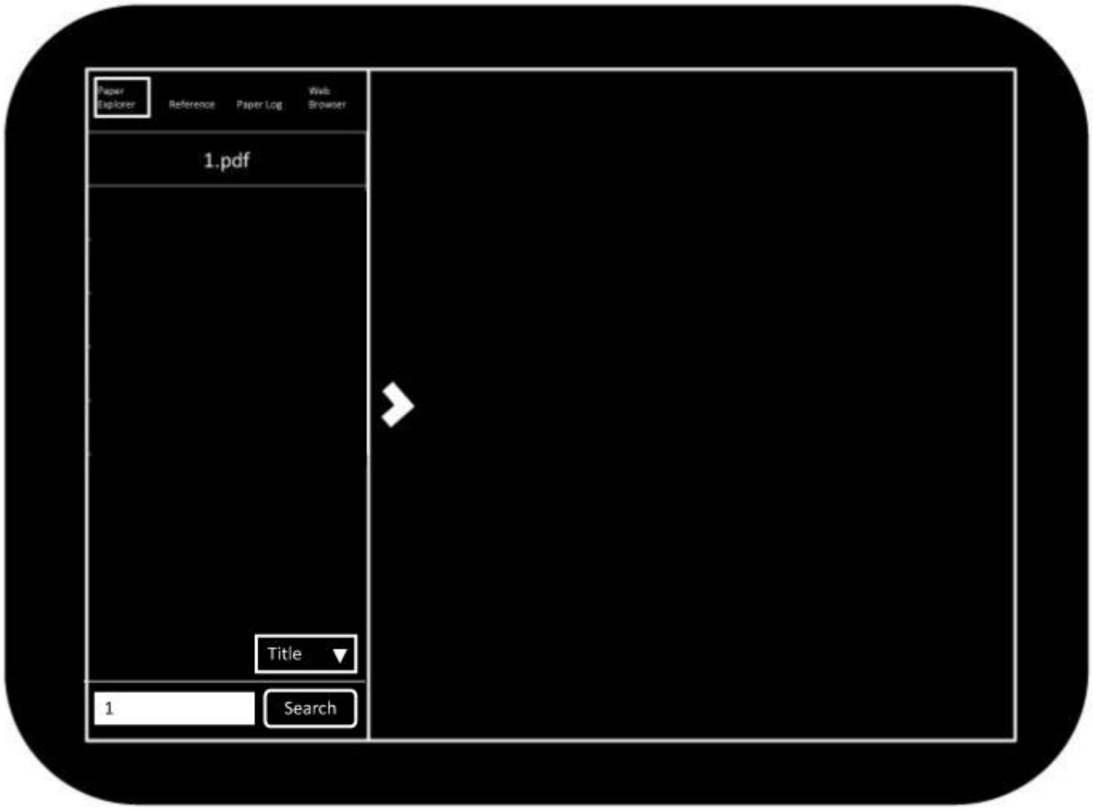


Figure 5.4 UI design of searching papers

5.1.2. Paper Parsing

Paper requires parsing and saving data in database.

<u>Get information needed (Phase 1)</u>

Based on the result of our studies, we need several fields of paper: title, authors, abstract, keywords (if any) and references. The title, authors name, abstract and key terms are all help to identify every single paper. The title and author play important roles in display the stream of papers while the abstract and keywords help in searching efficiently. References are parsed for achieving better referencing.

<u>Check duplication (Phase 2)</u>

During the parsing, we would like to compare the original file name and title with the others in the database. If we found the title is duplicated, we will further compare the authors and abstract. If the file is truly duplicated, it will be abandoned. However, we would try to see if the version is different or few modifications are applied. In this case, we would like to store two or more versions for reference.

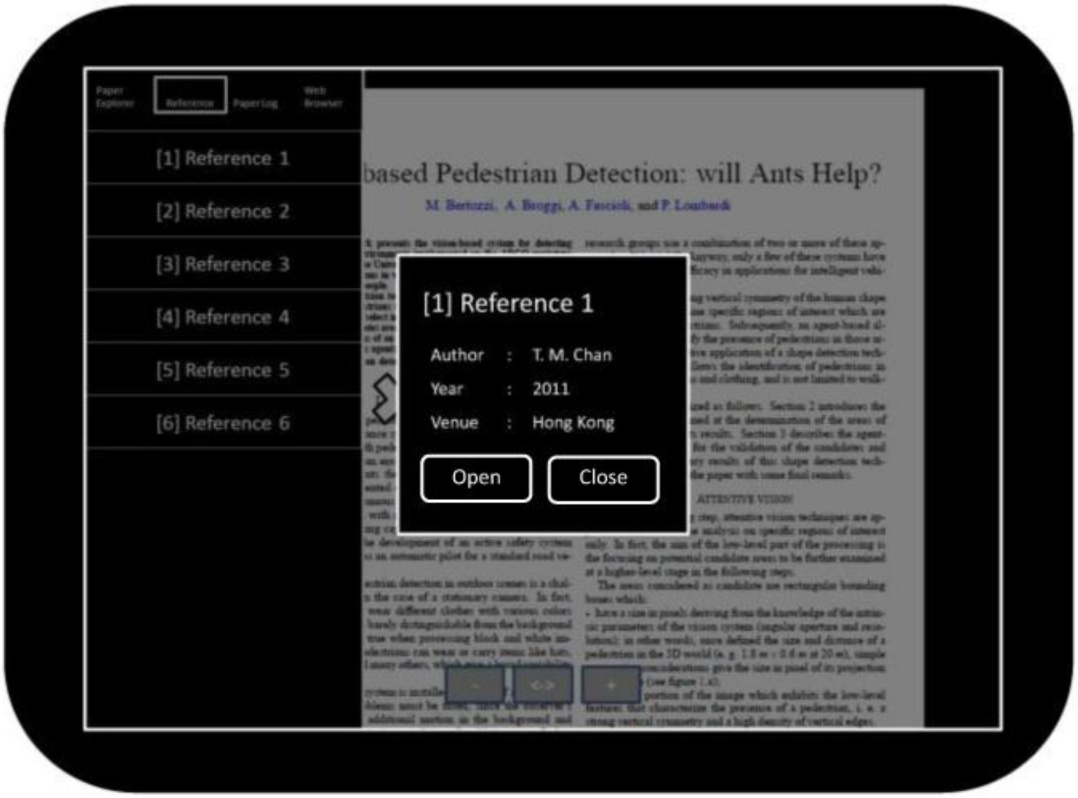
<u>Check the linkage (Phase 2)</u>

When a file is being parsed, it should be compared with the reference list of the other papers. If the references of other paper refer to this paper, they should be recorded and shown to the users.

5.1.3. Referencing

Show the reference (Phase 1)

The indexes appear throughout the paper will be highlighted. When one of the indexes is clicked, a box will be prompted and show the piece of reference.



based Pedestrian Detection: will Ants Help?

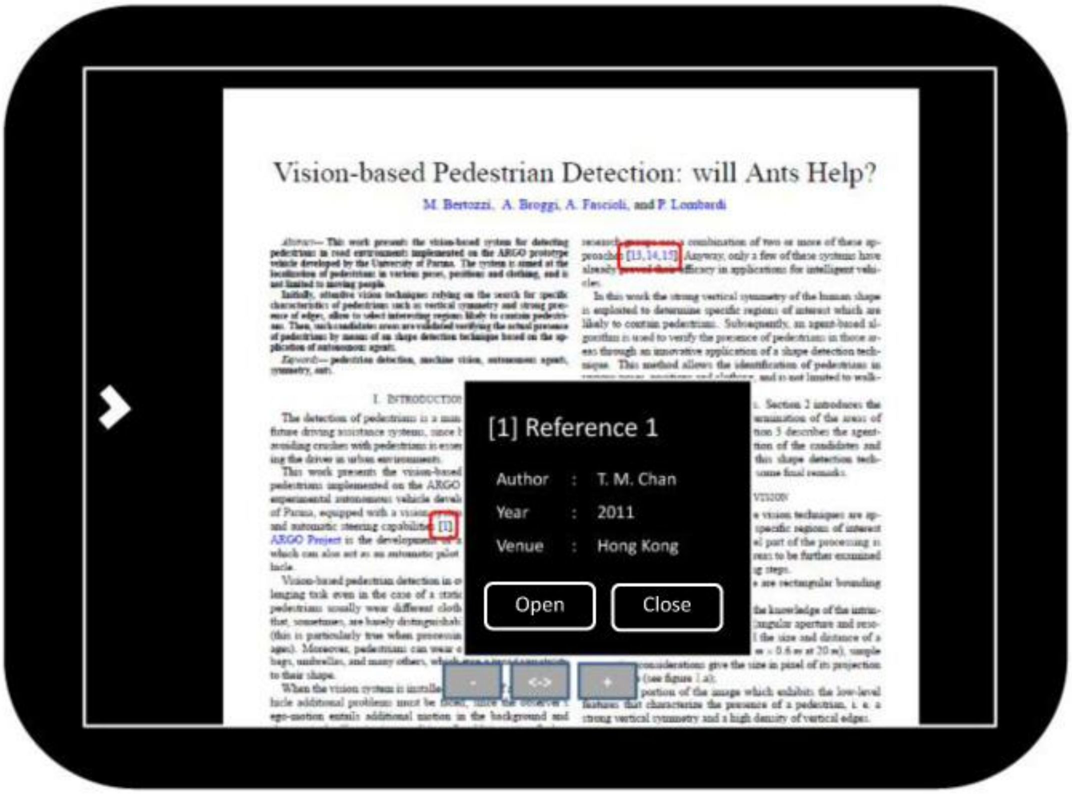
M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi

[1] Reference 1

Author : T. M. Chan
Year : 2011
Venue : Hong Kong

Figure 5.5 UI design of showing reference list

When one of the button of reference is clicked, a box will be prompted and show the piece of reference.



[1] Reference 1

Author : T. M. Chan

Year : 2011

Venue : Hong Kong

Vision-based Pedestrian Detection: will Ants Help?
M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi
[13,14,15]
[1]
Open
Close
-
<->
+

Figure 5.6 UI design of showing reference list

Build linkage to other paper (Phase 2)

If linkage is found during parsing, the result would be recorded in the database. One the linking reference is prompted out, a hyperlink is built and the user can check what it refers to on clicking.

5.1.4. library renewal

Manage files manually (Phase 1)

A panel is designed for the administrator to perform the paper management handily and efficiently. The panel allows the admin to upload files, to control the parsing procedure manually and to delete files. The panel helps the developer for further enhancement by showing all fields and entries to ease viewing the parsing performance.

Embed a browser (Phase 2)

A browser is suggested to be embedded within the application, so as to make the further search more handily. When downloading PDF is invoked during the searching in the browser, the application can automatically chase the file downloaded.

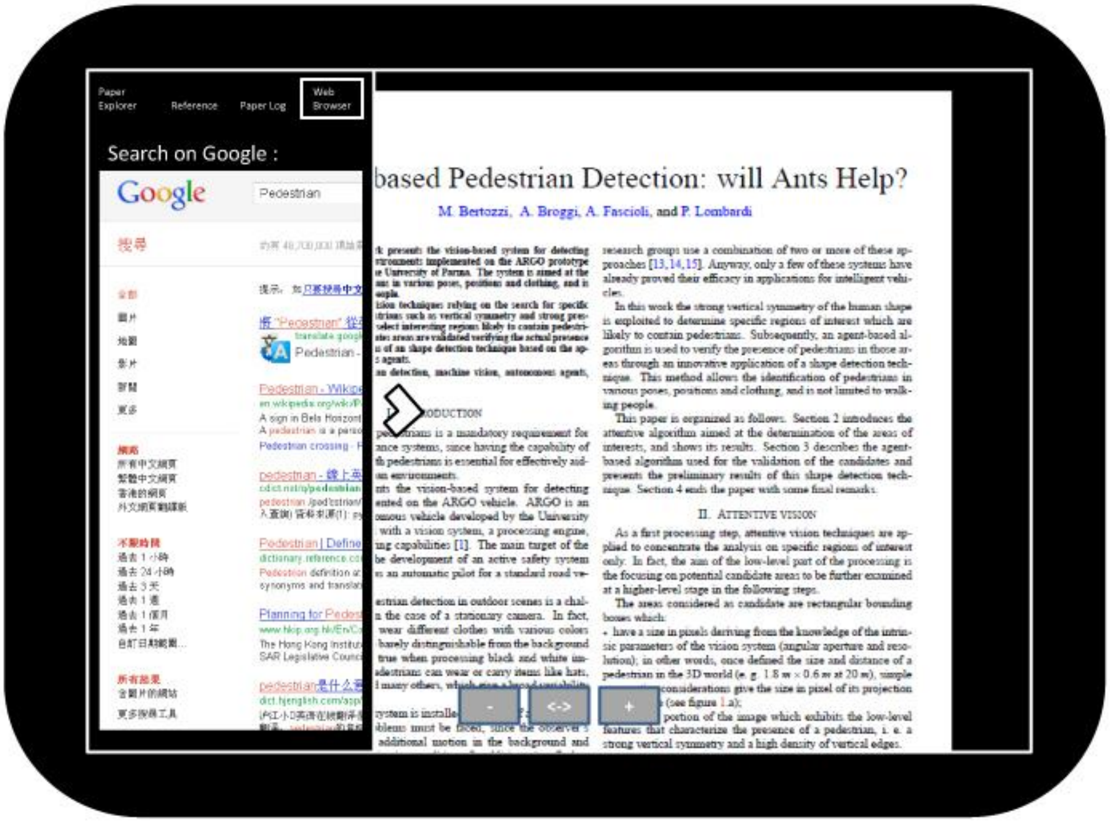

Paper Explorer
Reference
Paper Log
Web Browser
Search on Google :
Google
Pedestrian
搜尋
based Pedestrian Detection: will Ants Help?
M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi
-
<->
+

Figure 5.7 UI design of showing web browser

Update automatically (Phase 2)

When application noticed there is PDF downloading relevant to certain papers, it would first check if the PDF is a scholarly paper. The paper will be grabbed to the server and under ordinary parsing, and hence the paper library can be further extended.

5.2. Module 2 : Communication

5.2.1 posting comments

Add comments (Phase 1)

Every user can add their comment at anywhere of the PDF file. When adding the comments, the user ID and the time will be recorded.

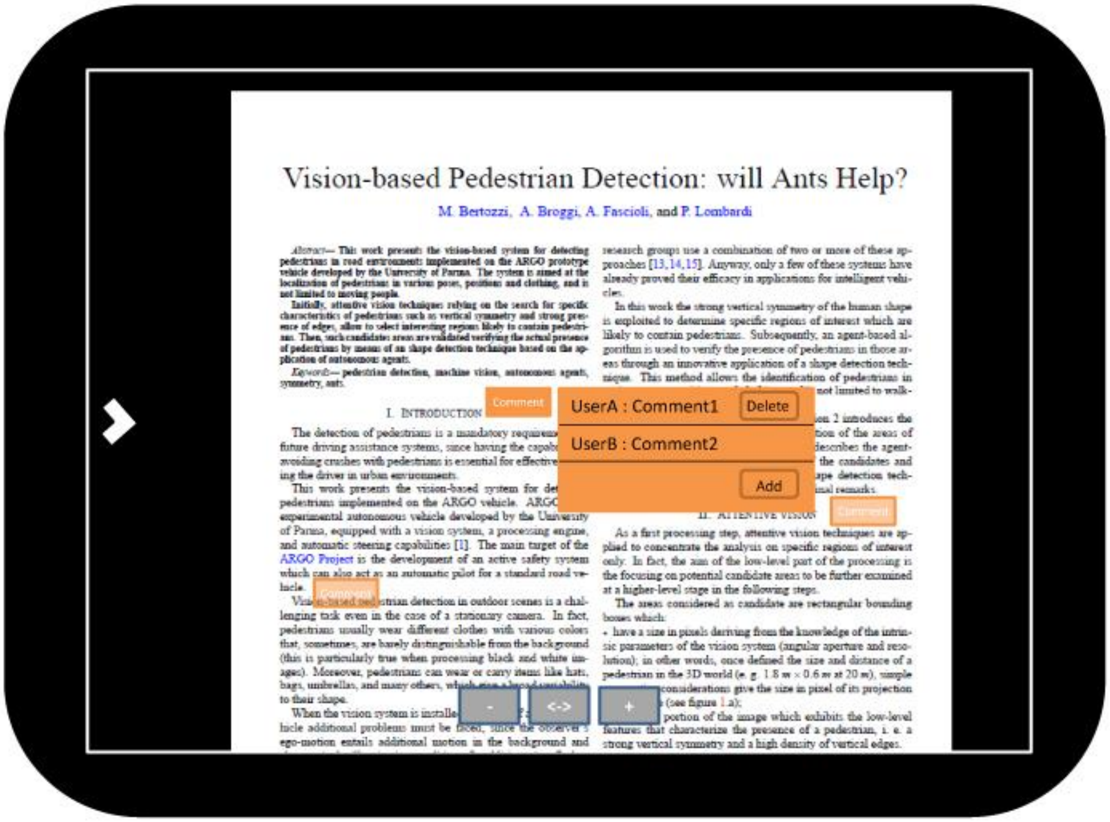

Vision-based Pedestrian Detection: will Ants Help?
M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi
Comment
UserA : Comment1
Delete
UserB : Comment2
Add
Comment
Comment
-
<->
+

Figure 5.8 UI design of showing comments

Show comments (Phase 1)

When a user added a comment, a flag / a button will be shown on the position he added. When the flag is clicked, a dialogue box will show and all the content of the discussion can be viewed. The discussion will be stored in the database and whenever any user opens the same paper, the comments will stay there and every one can view.

Reply comments (Phase 1)

On the dialogue box showing the details of a discussion, a reply button is available for others to keep the discussion.

Delete comments (Phase 1)

For better management and avoiding misunderstanding in discussion, editing comment is not allowed. However, deleting comment is allowed. If the head topic is deleted, all the replies will be deleted. Yet if the reply is deleted, the topic will still keep.

Notify user of new replies (Phase 2)

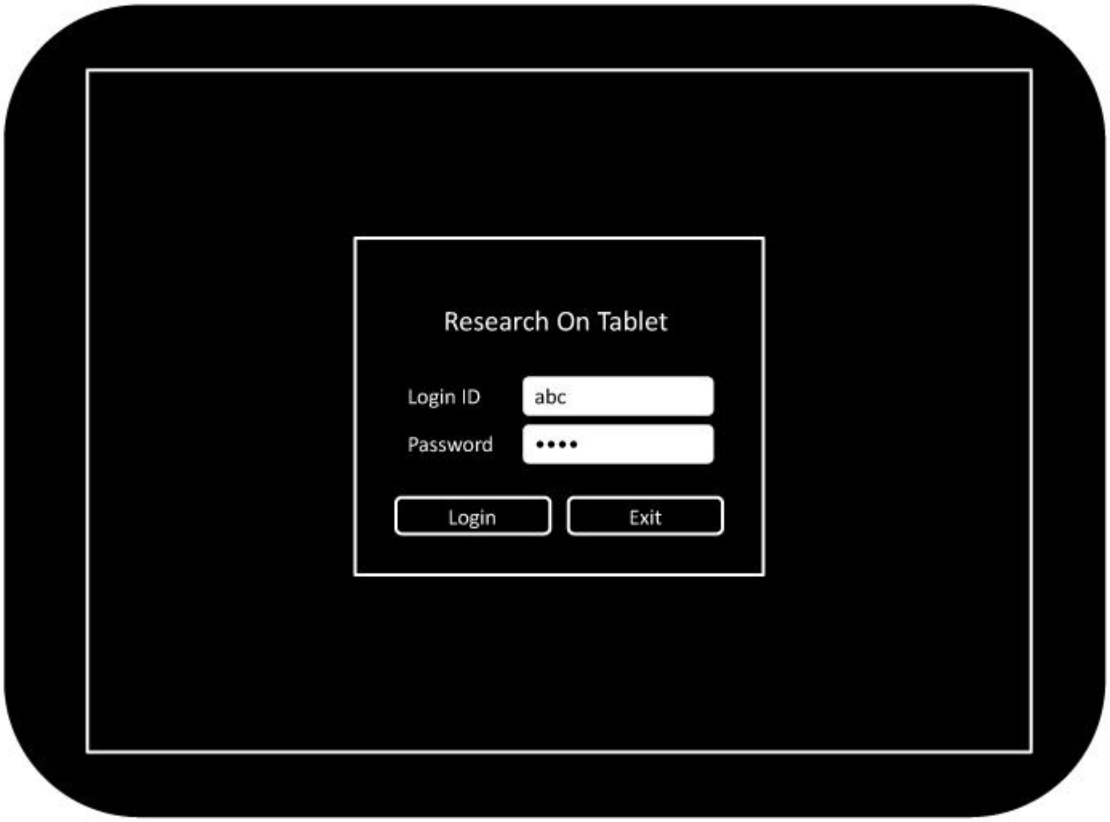
The users who have participated in the discussion should all be notified when a new reply responded to the discussion is posted. This could further promote the exchange of ideas. This requires collaboration with the development of research log.

5.3. Module 3 : Research Log

5.3.1. User accounts

<u>Login to system (Phase 1)</u>

Before the user use the application, he should login to the page. Unauthorized login is not allowed to use the application. This ensures every user can perform all the features smoothly.



Research On Tablet

Login ID: abc

Password: ••••

Figure 5.9 UI design of login

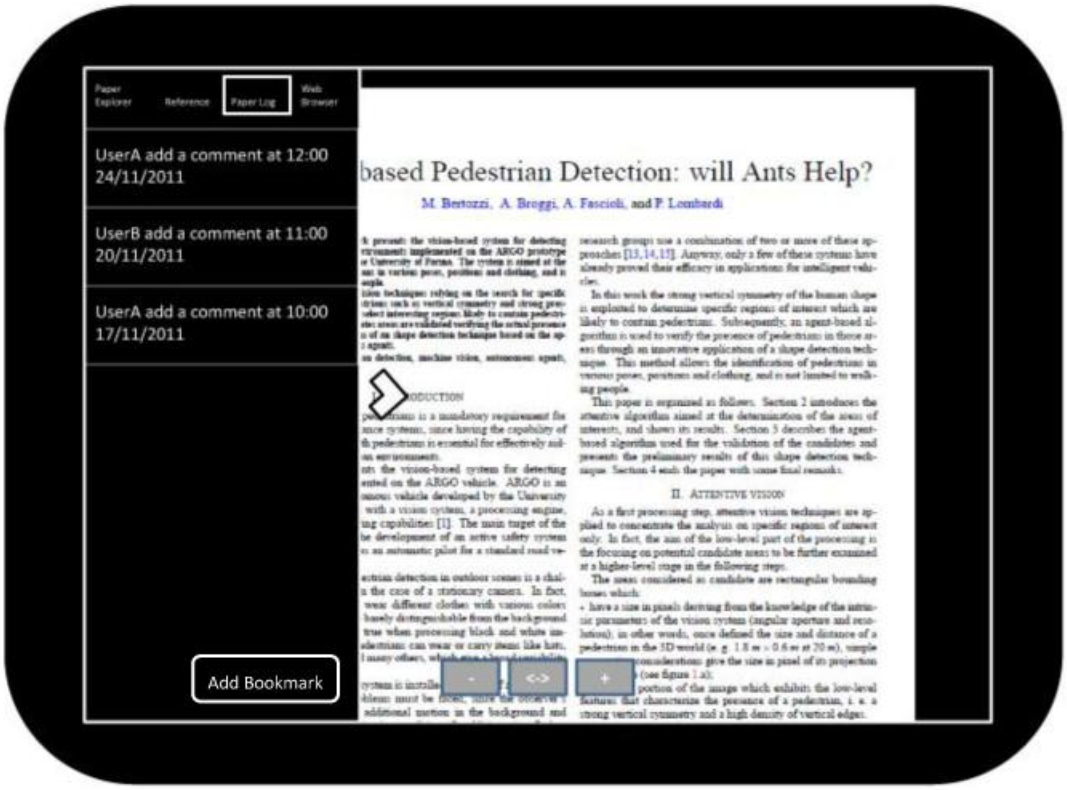
Edit user profile (Phase 2)

There is a simple and basic profile for each user, just to make things clear and help the users to know each other. The profile may include the name, description, and contact information.

5.3.2. Research log

View history of reading (Phase 2)

The reading log will refresh every time the user open a paper to read. The entries of the log will be sorted by the time happened in descending order. For better resources management, about 100 entries will be recorded for each user.



UserA add a comment at 12:00 24/11/2011

UserB add a comment at 11:00 20/11/2011

UserA add a comment at 10:00 17/11/2011

based Pedestrian Detection: will Ants Help?

M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi

Figure 5.10 UI design of research log

Bookmark paper (Phase 2)

If a paper is found interesting or valuable for researching, the user can bookmark the papers for future reference. They can found their bookmarks in the bookmark tab of the research log.

View commenting activities (Phase 2)

The log will show the all the commenting activities that the user performed, including both adding comments and replying. The commenting time, title of paper will be shown.

Show notification of new reply (Phase 2)

While the discussion the user participated has been renewed, the user should be notified by referring the log.

5.4. Database design

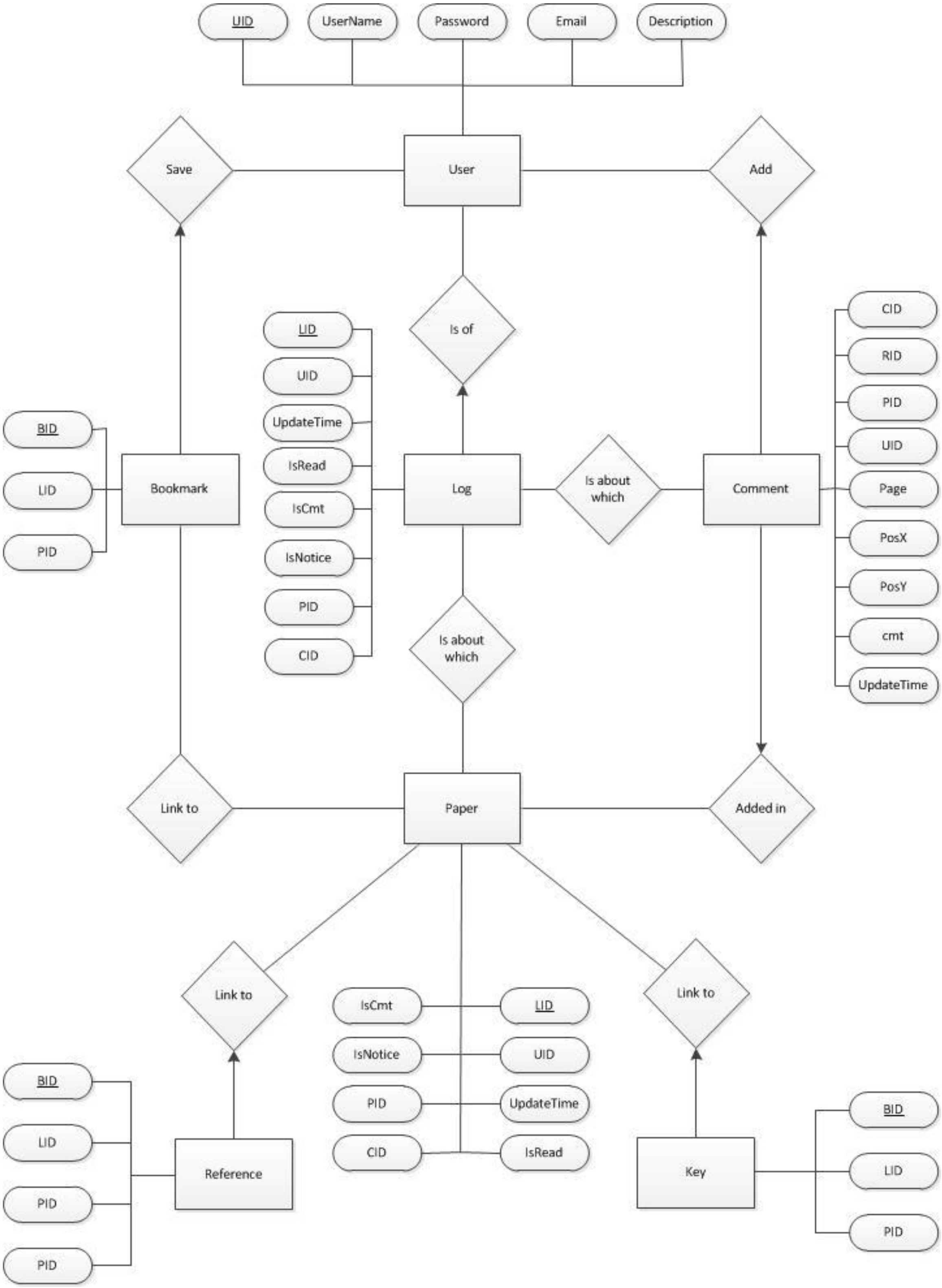


Figure 5.11 ER diagram

6. Implementation

In the chapter, we are going to describe the implementation of features we implemented in project phase 1 only. For every module, we will describe the implementation method for each function in both client side and server side respectively.

6.1. Module 1 : Paper Management

6.1.1. File Exploring

Show list of paper

A) SERVER SIDE –

The server calls functions in PHP to open the directory "paper/" which store all the files. It returns all the files except the current directory path and parent directory path.

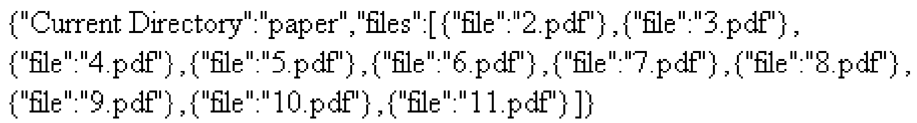
```
- Pseudo code -

IF (open the directory successfully)
{       WHILE (there is file)
        {       GET every file within the directory
                PUT the file into the json array     }
        CLOSE the directory         }
RETURN the json array
```

Note that the file name is changed into paper ID during the uploading of paper. Original file name with be kept in the database, for the uses of checking duplication in the future.

The expected result will contain a list of file names and will be encoded as a JSON array.

- Value Return - Example -



```
{"Current Directory":"paper","files":[{"file":"2.pdf"},{"file":"3.pdf"},
{"file":"4.pdf"},{"file":"5.pdf"},{"file":"6.pdf"},{"file":"7.pdf"},{"file":"8.pdf"},
{"file":"9.pdf"},{"file":"10.pdf"},{"file":"11.pdf"}]}
```

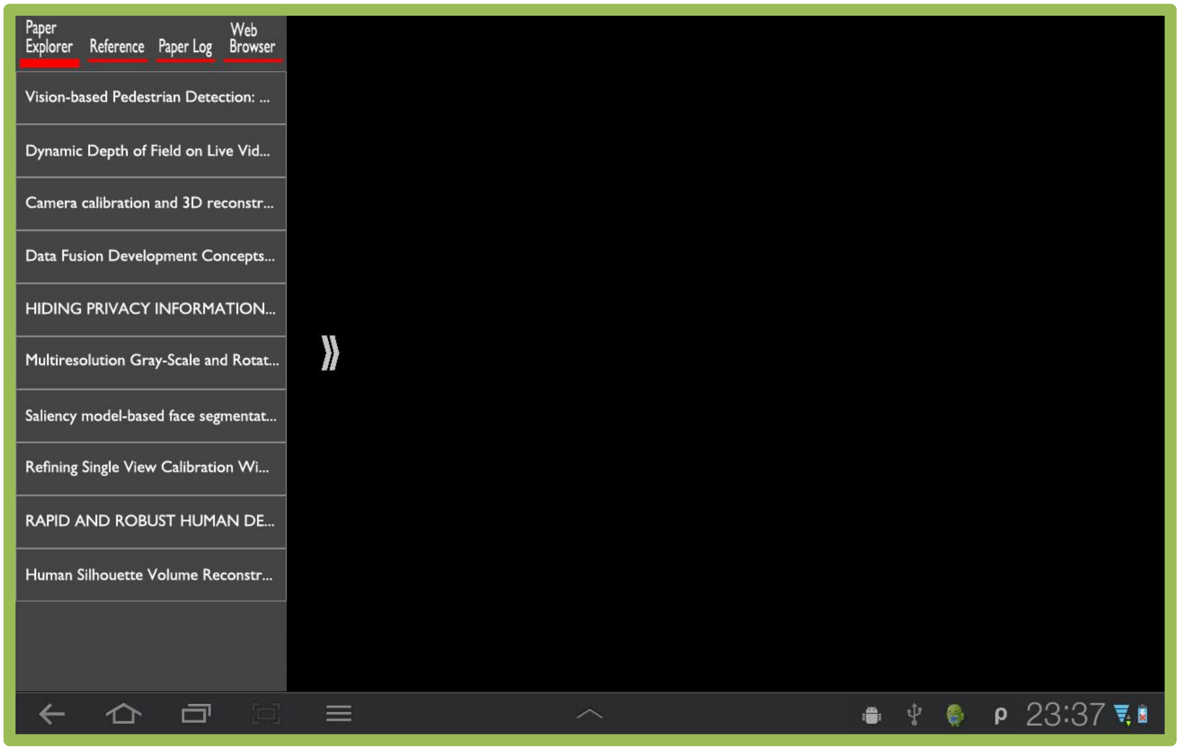
B) CLIENT SIDE –

After login validates, the client device sends request to the server and gets the list of file names. It will show the list in the file explorer in transparent panel on the left hand side one by one.

- Pseudo code -

```
IF (login validate)
{       SEND request to web server for the list of papers
     GET the result from web server
        SHOW the list of papers in the transparent panel      }
```

Although the file name is changed into paper ID, the list will show the title of the papers obtained from the database through the web server.



Vision-based Pedestrian Detection: ...

Dynamic Depth of Field on Live Vid...

Camera calibration and 3D reconstr...

Data Fusion Development Concepts...

HIDING PRIVACY INFORMATION...

Multiresolution Gray-Scale and Rotat...

Saliency model-based face segmentat...

Refining Single View Calibration Wi...

RAPID AND ROBUST HUMAN DE...

Human Silhouette Volume Reconstr...

Figure 6.1 Screen capture of the paper list

Show paper

A) SERVER SIDE –

The server saves the PDF files by changing the name to paper ID. This can avoid crashing due to the presence of special characters in file name.

B) CLIENT SIDE –

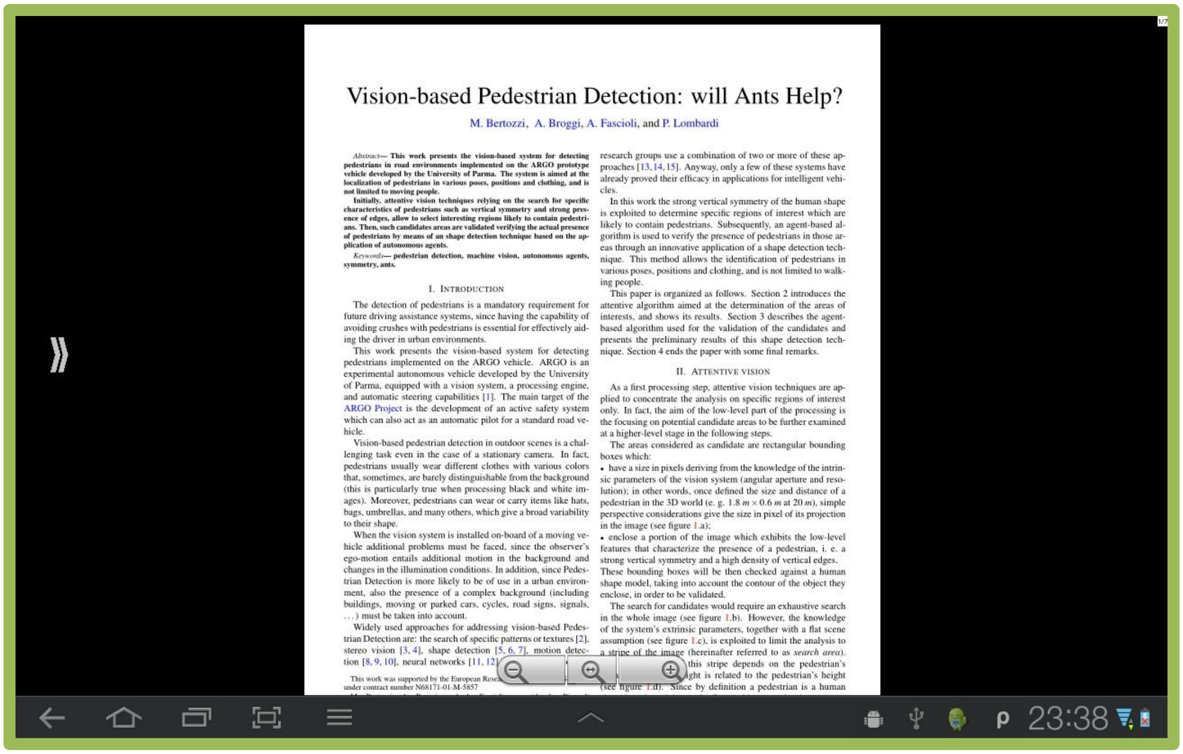
When the user selects a paper, the client device will download the paper through web server and save it in a temp folder. Then the client device will show the paper with some functions, including zooming, finding texts and going to pages. Besides, it will also show the buttons for reference and previous comments in the paper after downloading the paper, these two features will be talked about later.

- Pseudo code -

```
IF (select paper)
{       SEND request to web server to download the paper
    SEND request to web server for the references and previous
        comments
    GET the result from web server

    SHOW paper

      SHOW buttons for reference and comments     }
```



Vision-based Pedestrian Detection: will Ants Help?

M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi

Abstract— **This work presents the vision-based system for detecting pedestrians in road environments implemented on the ARGO prototype vehicle developed by the University of Parma. The system is aimed at the localization of pedestrians in various poses, positions and clothing, and is not limited to moving people.**

Initially, attentive vision techniques relying on the search for specific characteristics of pedestrians such as vertical symmetry and strong presence of edges, allow to select interesting regions likely to contain pedestrians. Then, such candidates areas are validated verifying the actual presence of pedestrians by means of an shape detection technique based on the application of autonomous agents.

Keywords— **pedestrian detection, machine vision, autonomous agents, symmetry, ants.**

I. INTRODUCTION

The detection of pedestrians is a mandatory requirement for future driving assistance systems, since having the capability of avoiding crushes with pedestrians is essential for effectively aiding the driver in urban environments.

This work presents the vision-based system for detecting pedestrians implemented on the ARGO vehicle. ARGO is an experimental autonomous vehicle developed by the University of Parma, equipped with a vision system, a processing engine, and automatic steering capabilities [1]. The main target of the ARGO Project is the development of an active safety system which can also act as an automatic pilot for a standard road vehicle.

Vision-based pedestrian detection in outdoor scenes is a challenging task even in the case of a stationary camera. In fact, pedestrians usually wear different clothes with various colors that, sometimes, are barely distinguishable from the background (this is particularly true when processing black and white images). Moreover, pedestrians can wear or carry items like hats, bags, umbrellas, and many others, which give a broad variability to their shape.

When the vision system is installed on-board of a moving vehicle additional problems must be faced, since the observer's ego-motion entails additional motion in the background and changes in the illumination conditions. In addition, since Pedestrian Detection is more likely to be of use in a urban environment, also the presence of a complex background (including buildings, moving or parked cars, cycles, road signs, signals, . . .) must be taken into account.

Widely used approaches for addressing vision-based Pedestrian Detection are: the search of specific patterns or textures [2], stereo vision [3, 4], shape detection [5, 6, 7], motion detection [8, 9, 10], neural networks [11, 12]

This work was supported by the European Rese
under contract number N68171-01-M-5857

research groups use a combination of two or more of these approaches [13, 14, 15]. Anyway, only a few of these systems have already proved their efficacy in applications for intelligent vehicles.

In this work the strong vertical symmetry of the human shape is exploited to determine specific regions of interest which are likely to contain pedestrians. Subsequently, an agent-based algorithm is used to verify the presence of pedestrians in those areas through an innovative application of a shape detection technique. This method allows the identification of pedestrians in various poses, positions and clothing, and is not limited to walking people.

This paper is organized as follows. Section 2 introduces the attentive algorithm aimed at the determination of the areas of interests, and shows its results. Section 3 describes the agent-based algorithm used for the validation of the candidates and presents the preliminary results of this shape detection technique. Section 4 ends the paper with some final remarks.

II. ATTENTIVE VISION

As a first processing step, attentive vision techniques are applied to concentrate the analysis on specific regions of interest only. In fact, the aim of the low-level part of the processing is the focusing on potential candidate areas to be further examined at a higher-level stage in the following steps.

The areas considered as candidate are rectangular bounding boxes which:

- have a size in pixels deriving from the knowledge of the intrinsic parameters of the vision system (angular aperture and resolution); in other words, once defined the size and distance of a pedestrian in the 3D world (e. g. 1.8 $m \times 0.6$ m at 20 m), simple perspective considerations give the size in pixel of its projection in the image (see figure 1.a);
- enclose a portion of the image which exhibits the low-level features that characterize the presence of a pedestrian, i. e. a strong vertical symmetry and a high density of vertical edges.

These bounding boxes will be then checked against a human shape model, taking into account the contour of the object they enclose, in order to be validated.

The search for candidates would require an exhaustive search in the whole image (see figure 1.b). However, the knowledge of the system's extrinsic parameters, together with a flat scene assumption (see figure 1.c), is exploited to limit the analysis to a stripe of the image (hereinafter referred to as *search area*).

this stripe depends on the pedestrian's

ight is related to the pedestrian's height (see figure 1.d). Since by definition a pedestrian is a human

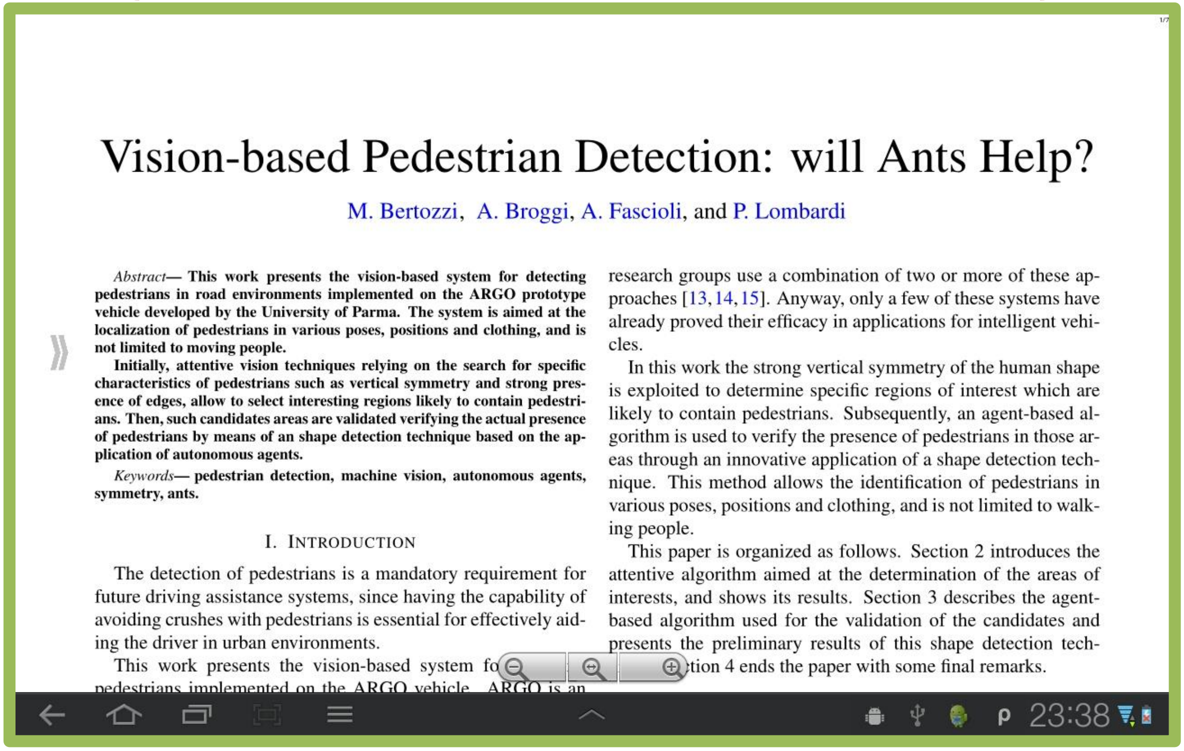
Figure 6.2 Screen capture of showing paper

The three functions, including zooming, finding texts and going to pages, are the built-in functions of the API of PDF viewer we have used. We have studied how they work and are going to make some changes to them, so as to make it more suitable to our design.

For the zooming functions, the user can zoom in, zoom out or zoom to fit the width of paper. These zooming functions are implemented by pressing three buttons in the bottom of the screen, or double tapping the screen.

- Pseudo code -

```
IF (press button of zooming in or double tap)
{   CALCULATE the new paper size
    SHOW the new view          }
IF (press button of zooming out)
{   CALCULATE the new paper size
    SHOW the new view          }
IF (press button of zooming to fit the width)
{   CALCULATE the new paper size
    SHOW the new view          }
```



Vision-based Pedestrian Detection: will Ants Help?

M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi

Abstract— This work presents the vision-based system for detecting pedestrians in road environments implemented on the ARGO prototype vehicle developed by the University of Parma. The system is aimed at the localization of pedestrians in various poses, positions and clothing, and is not limited to moving people.

Initially, attentive vision techniques relying on the search for specific characteristics of pedestrians such as vertical symmetry and strong presence of edges, allow to select interesting regions likely to contain pedestrians. Then, such candidates areas are validated verifying the actual presence of pedestrians by means of an shape detection technique based on the application of autonomous agents.

Keywords— pedestrian detection, machine vision, autonomous agents, symmetry, ants.

I. Introduction

The detection of pedestrians is a mandatory requirement for future driving assistance systems, since having the capability of avoiding crushes with pedestrians is essential for effectively aiding the driver in urban environments.

This work presents the vision-based system fo pedestrians implemented on the ARGO vehicle. ARGO is an research groups use a combination of two or more of these approaches [13, 14, 15]. Anyway, only a few of these systems have already proved their efficacy in applications for intelligent vehicles.

In this work the strong vertical symmetry of the human shape is exploited to determine specific regions of interest which are likely to contain pedestrians. Subsequently, an agent-based algorithm is used to verify the presence of pedestrians in those areas through an innovative application of a shape detection technique. This method allows the identification of pedestrians in various poses, positions and clothing, and is not limited to walking people.

This paper is organized as follows. Section 2 introduces the attentive algorithm aimed at the determination of the areas of interests, and shows its results. Section 3 describes the agent-based algorithm used for the validation of the candidates and presents the preliminary results of this shape detection tech- tion 4 ends the paper with some final remarks.

Figure 6.3 Screen capture of zooming

For the function of finding texts, the user can find the texts in the paper after pressing the button of finding texts. The texts found by the user will then be bounced three red lines. Besides, the user can choose to find the previous and next result. If this texts finding is finished, the user can hide all the result, and the view is returned to the normal view.

- Pseudo code -

```
IF (choose finding text)
{  ASK the user to input the text for finding
   FIND the text in the paper
   SHOW the results in the paper
   WHILE (showing paper with results of finding)
   {        IF (choose previous result of finding)
            {        MOVE the view to the previous result     }
            IF (choose next result of finding)
            {        MOVE the view to the next result }
            IF (choose hiding the results of finding)
            {        HIDE all the results of finding
                     BREAK        }        }        }
```

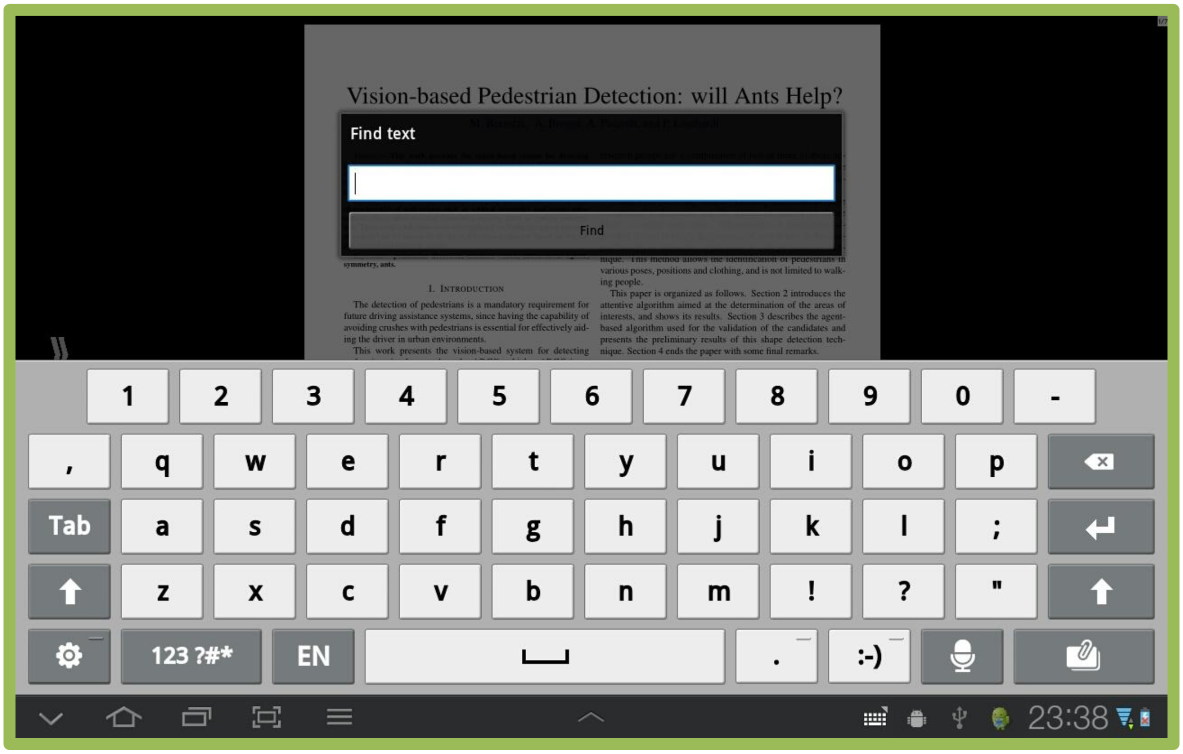
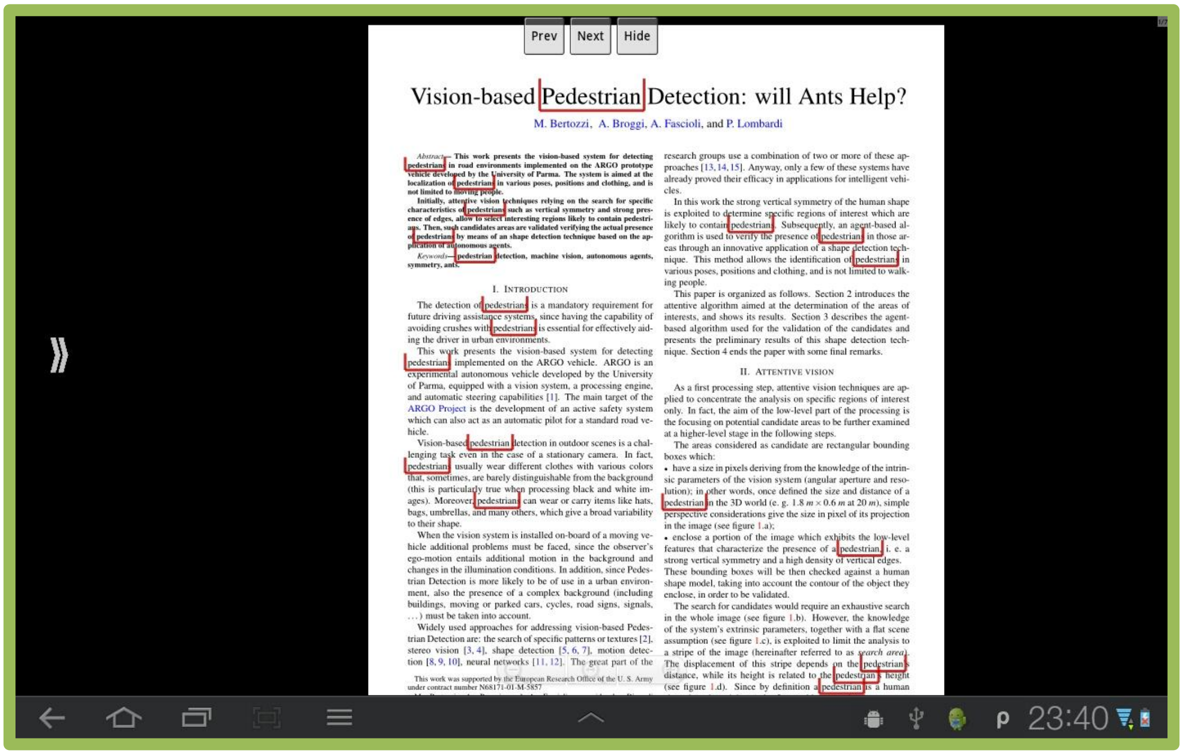


Figure 6.4 Screen capture of finding texts



Vision-based Pedestrian Detection: will Ants Help?

M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi

Abstract— **This work presents the vision-based system for detecting pedestrians in road environments implemented on the ARGO prototype vehicle developed by the University of Parma. The system is aimed at the localization of pedestrians in various poses, positions and clothing, and is not limited to moving people.**

Initially, attentive vision techniques relying on the search for specific characteristics of pedestrians such as vertical symmetry and strong presence of edges, allow to select interesting regions likely to contain pedestrians. Then, such candidates areas are validated verifying the actual presence of pedestrians by means of an shape detection technique based on the application of autonomous agents.

Keywords— **pedestrian detection, machine vision, autonomous agents, symmetry, ants.**

I. Introduction

The detection of pedestrians is a mandatory requirement for future driving assistance systems, since having the capability of avoiding crushes with pedestrians is essential for effectively aiding the driver in urban environments.

This work presents the vision-based system for detecting pedestrians implemented on the ARGO vehicle. ARGO is an experimental autonomous vehicle developed by the University of Parma, equipped with a vision system, a processing engine, and automatic steering capabilities [1]. The main target of the ARGO Project is the development of an active safety system which can also act as an automatic pilot for a standard road vehicle.

Vision-based pedestrian detection in outdoor scenes is a challenging task even in the case of a stationary camera. In fact, pedestrians usually wear different clothes with various colors that, sometimes, are barely distinguishable from the background (this is particularly true when processing black and white images). Moreover, pedestrians can wear or carry items like hats, bags, umbrellas, and many others, which give a broad variability to their shape.

When the vision system is installed on-board of a moving vehicle additional problems must be faced, since the observer's ego-motion entails additional motion in the background and changes in the illumination conditions. In addition, since Pedestrian Detection is more likely to be of use in a urban environment, also the presence of a complex background (including buildings, moving or parked cars, cycles, road signs, signals, . . .) must be taken into account.

Widely used approaches for addressing vision-based Pedestrian Detection are: the search of specific patterns or textures [2], stereo vision [3, 4], shape detection [5, 6, 7], motion detection [8, 9, 10], neural networks [11, 12]. The great part of the research groups use a combination of two or more of these approaches [13, 14, 15]. Anyway, only a few of these systems have already proved their efficacy in applications for intelligent vehicles.

In this work the strong vertical symmetry of the human shape is exploited to determine specific regions of interest which are likely to contain pedestrians. Subsequently, an agent-based algorithm is used to verify the presence of pedestrians in those areas through an innovative application of a shape detection technique. This method allows the identification of pedestrians in various poses, positions and clothing, and is not limited to walking people.

This paper is organized as follows. Section 2 introduces the attentive algorithm aimed at the determination of the areas of interests, and shows its results. Section 3 describes the agent-based algorithm used for the validation of the candidates and presents the preliminary results of this shape detection technique. Section 4 ends the paper with some final remarks.

II. Attentive vision

As a first processing step, attentive vision techniques are applied to concentrate the analysis on specific regions of interest only. In fact, the aim of the low-level part of the processing is the focusing on potential candidate areas to be further examined at a higher-level stage in the following steps.

The areas considered as candidate are rectangular bounding boxes which:

- have a size in pixels deriving from the knowledge of the intrinsic parameters of the vision system (angular aperture and resolution); in other words, once defined the size and distance of a pedestrian in the 3D world (e. g. 1.8 $m \times 0.6\, m$ at 20 m), simple perspective considerations give the size in pixel of its projection in the image (see figure 1.a);
- enclose a portion of the image which exhibits the low-level features that characterize the presence of a pedestrian, i. e. a strong vertical symmetry and a high density of vertical edges.

These bounding boxes will be then checked against a human shape model, taking into account the contour of the object they enclose, in order to be validated.

The search for candidates would require an exhaustive search in the whole image (see figure 1.b). However, the knowledge of the system's extrinsic parameters, together with a flat scene assumption (see figure 1.c), is exploited to limit the analysis to a stripe of the image (hereinafter referred to as *search area*). The displacement of this stripe depends on the pedestrian's distance, while its height is related to the pedestrian's height (see figure 1.d). Since by definition a pedestrian is a human

This work was supported by the European Research Office of the U. S. Army under contract number N68171-01-M-5857

Figure 6.5 Screen capture of result of finding texts

The last function is going to a certain page. The user can go to the first page, last page or a certain page entered.

```
-   Pseudo code   -

IF (choose going to certain page)
{  ASK the user to input the page number
   SHOW the paper at that page      }
IF (choose going to first page)
{  SHOW the paper at the first page }
IF (choose going to the last page)
{  SHOW the paper at the last page }
```

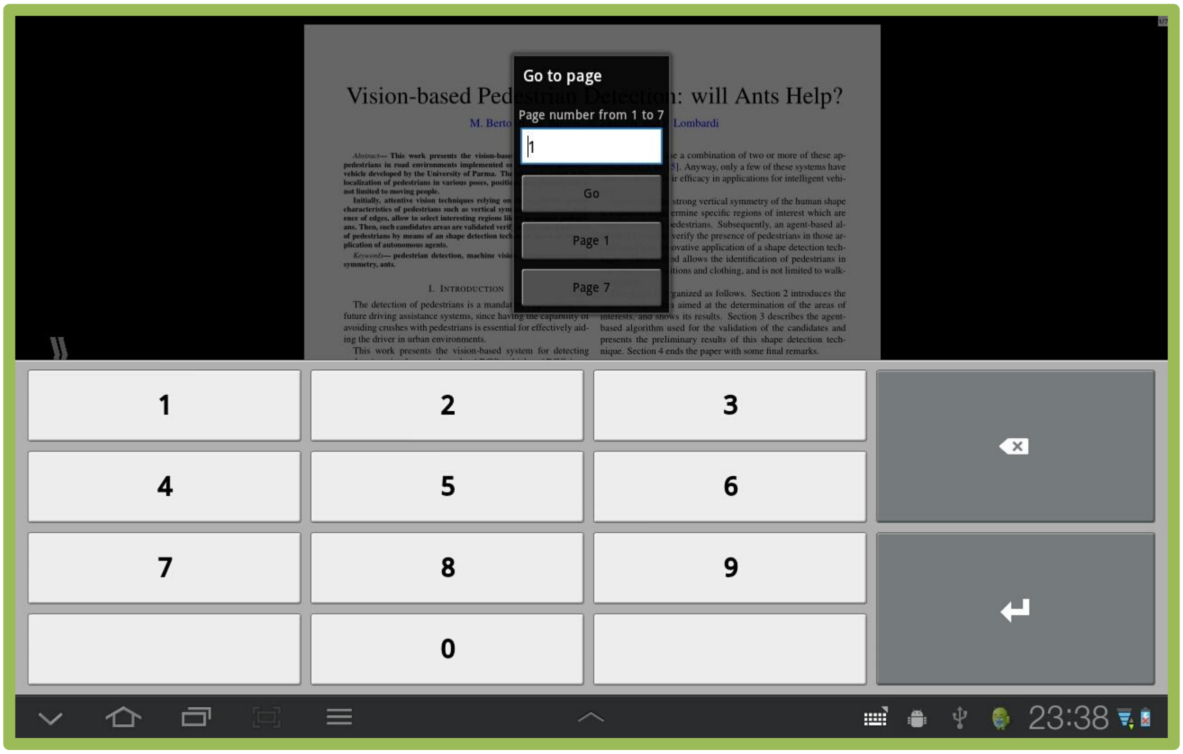


Figure 6.6 Screen capture of result of going to page

6.1.2. Paper Parsing

Get information needed

A) SERVER SIDE –

To begin with, once the paper is uploaded into the server, the server will call the “pdftotext” program (developed by Xpdf [13]) to print the text elements into the text file. A text file will be generated and will be used to parse the specific field of the paper.

In the section, the result of our survey conducted in chapter 3 should be set as reference. To implement the parsing, we summed up for several methods of parse:

1) Search for the pattern of the target information.
2) Search for the keywords of the target information
3) Search for the start of the target, ie. what is expected to be written before the target.
4) Search for the end of the target, ie. what is expect to appear after the end of the target.

One or more above methods will be used in parsing different fields.

Now we would like to show the implementation one by one.

Title -

Using the result analyzed, there are seldom the "title" term to indicate the title. Also, the titles may not be first line of the paper and they may stay in more than one line. Extracting the first line or some lines are not possible, because it would be more difficult for the computer to identify whether the line following is the authors' name or just some technical proper noun.

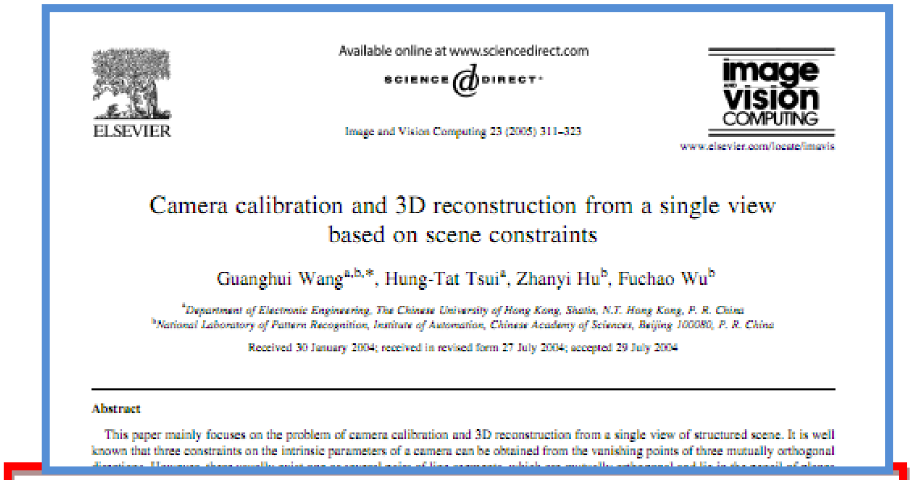
Based on these, we try to get the font size of all the text elements and compare them (implementation method 1 – pattern recognition). We can quite sure that the font size of title is different from the header and the author's information. We found that the title is usually in the biggest size. Moreover, to make sure that the size got is the title's one, few more rules should be set. For example. we compare the size of the fonts on the first page only.

- Pseudo code -

```
WHILE ( is on page one)
{       GET font size of each line
        GET the largest font size
        GET the line number(s) of the largest font appears}
OPEN the txt file
COPY the text in the line(s) of the largest font appears
UPDATE the database with title got
RETURN the title found
```

The title of the paper will be return in expect.

- Value Return - Example -



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Image and Vision Computing 23 (2005) 311–323

www.elsevier.com/locate/imavis

Camera calibration and 3D reconstruction from a single view based on scene constraints

Guanghui Wang[a,b,*], Hung-Tat Tsui[a], Zhanyi Hu[b], Fuchao Wu[b]

[a]Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong, P. R. China
[b]National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, P. R. China

Received 30 January 2004; received in revised form 27 July 2004; accepted 29 July 2004

Abstract

This paper mainly focuses on the problem of camera calibration and 3D reconstruction from a single view of structured scene. It is well known that three constraints on the intrinsic parameters of a camera can be obtained from the vanishing points of three mutually orthogonal

```
Image and Vision Computing 23 (2005) 311â€“323 www.elsevier.com/locate/imavis

Camera calibration and 3D reconstruction from a single view based on scene constraints
Guanghui Wanga,b,*, Hung-Tat Tsuia, Zhanyi Hub, Fuchao Wub
b a Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin,
N.T. Hong Kong, P. R. China National Laboratory of Pattern Recognition, Institute of
```

Camera calibration and 3D reconstruction from a single view based on scene constraints

Author -

According to the statistics, most of the papers are written by two or more authors. The authors' names are usually grouped in one line right after the title. There are often two to five lines descriptions and contact information follows the authors' names. We found that it is an extremely complex task to search for the next line of authors' names, because there is no any standard and pattern for the description. The description may include the title of the authors or the organization they belong to. Therefore, the implementation method 4 (search for the end) is not feasible. The lack of uniformity is also found in the authors' name, and hence the recognition of pattern for the names is not possible too.

Finally, we decided to make the method simple. We would like to simply parse the line after the title (implementation method 3 – search for start) because most of cases the authors' names appear right after the title.

- Pseudo code -

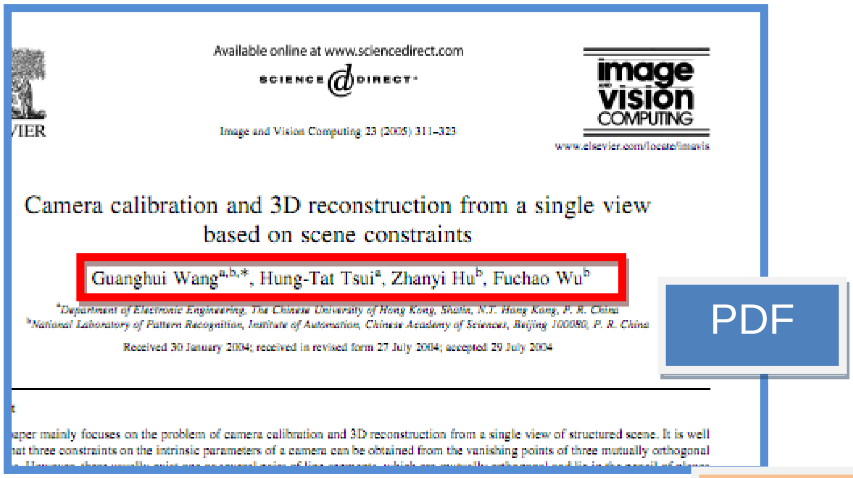
```
WHILE (title is found)

{        GET the next line AS authors }

UPDATE the database with authors got

RETURN authors
```

- Value Return - Example -



VIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Image and Vision Computing 23 (2005) 311–323

www.elsevier.com/locate/imavis

Camera calibration and 3D reconstruction from a single view based on scene constraints

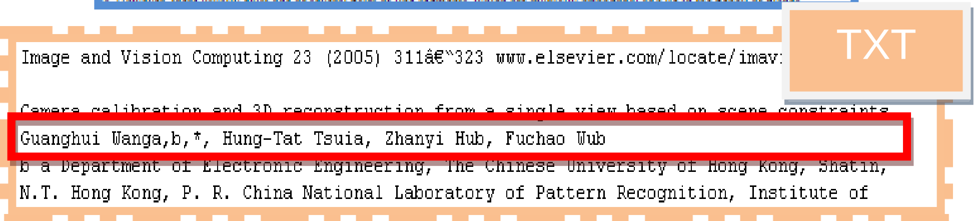
Guanghui Wang[a,b,*], Hung-Tat Tsui[a], Zhanyi Hu[b], Fuchao Wu[b]

[a]Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong, P. R. China
[b]National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, P. R. China

Received 30 January 2004; received in revised form 27 July 2004; accepted 29 July 2004

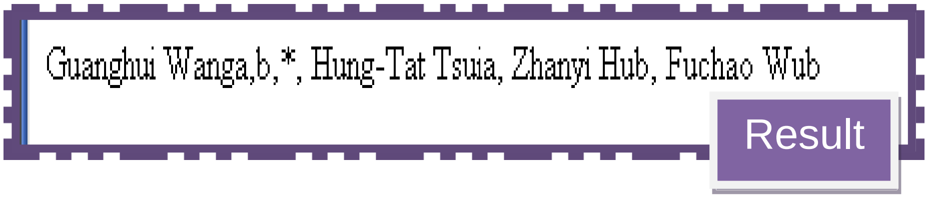
t

aper mainly focuses on the problem of camera calibration and 3D reconstruction from a single view of structured scene. It is well
at three constraints on the intrinsic parameters of a camera can be obtained from the vanishing points of three mutually orthogonal



```
Image and Vision Computing 23 (2005) 311â€“323 www.elsevier.com/locate/imav:

Camera calibration and 3D reconstruction from a single view based on scene constraints
Guanghui Wanga,b,*, Hung-Tat Tsuia, Zhanyi Hub, Fuchao Wub
b a Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin,
N.T. Hong Kong, P. R. China National Laboratory of Pattern Recognition, Institute of
```



Guanghui Wanga,b,*, Hung-Tat Tsuia, Zhanyi Hub, Fuchao Wub

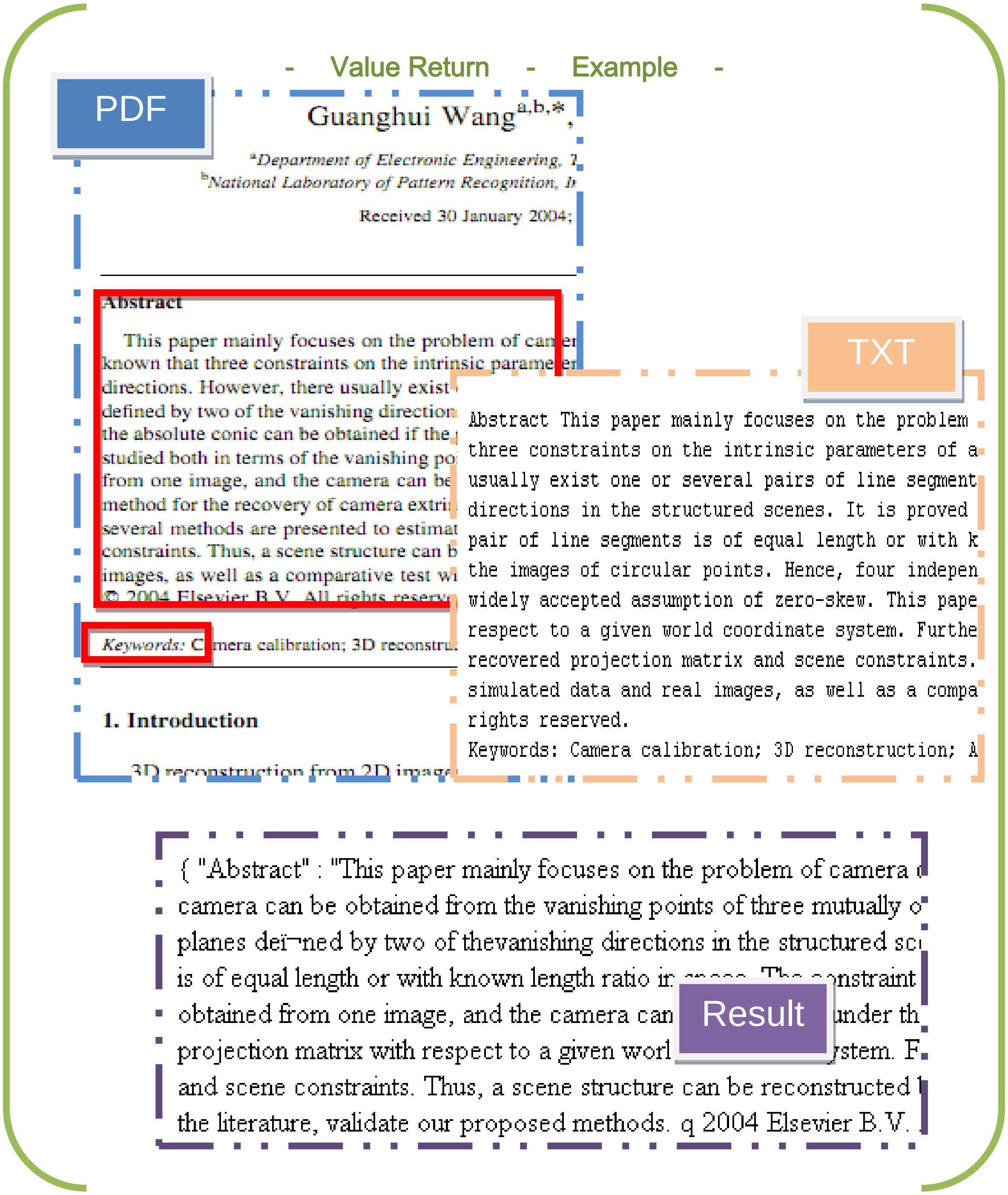
Abstract -

Usually the abstract is written after the appearance of the "Abstract" word, with reference to the survey result. Therefore it is easy to locate where the abstract starts (implementation method 2 – keywords location). However, sometimes the abstract are in more than a paragraph. To make sure the scope of abstract can be extract correctly, we may also check where is the end point of the abstract. Usually the introduction, sometimes the keywords column, follows the abstract. Therefore the scope ends before the line which contains the "keywords" or "introduction/summary" (implementation skill 4 – search for the end).

- Pseudo code -

```
WHILE (a line is parsed)
{       IF ("Abstract" is found)
        {       GET every line AS abstract
                UNTIL "Keywords" or "Introduction" appears } }
UPDATE the database with abstract got
RETURN abstract in a josn
```

Since the abstract may consist of several lines or paragraph, it will be encoded as json array to ensure correct data transmission.



Guanghui Wang[a,b,*],
[a]Department of Electronic Engineering, T
[b]National Laboratory of Pattern Recognition, In
Received 30 January 2004;

Abstract

This paper mainly focuses on the problem of camer
known that three constraints on the intrinsic parameter
directions. However, there usually exist
defined by two of the vanishing direction
the absolute conic can be obtained if the
studied both in terms of the vanishing poi
from one image, and the camera can be
method for the recovery of camera extri
several methods are presented to estimat
constraints. Thus, a scene structure can b
images, as well as a comparative test wi
© 2004 Elsevier B.V. All rights reserve

Keywords: Camera calibration; 3D reconstru

1. Introduction

3D reconstruction from 2D image

```
Abstract This paper mainly focuses on the problem
three constraints on the intrinsic parameters of a
usually exist one or several pairs of line segment
directions in the structured scenes. It is proved
pair of line segments is of equal length or with k
the images of circular points. Hence, four indepen
widely accepted assumption of zero-skew. This pape
respect to a given world coordinate system. Furthe
recovered projection matrix and scene constraints.
simulated data and real images, as well as a compa
rights reserved.
Keywords: Camera calibration; 3D reconstruction; A
```

```
{ "Abstract" : "This paper mainly focuses on the problem of camera c
camera can be obtained from the vanishing points of three mutually o
planes deï¬ned by two of thevanishing directions in the structured sc
is of equal length or with known length ratio in space. The constraint
obtained from one image, and the camera can                under th
projection matrix with respect to a given worl              ystem. F
and scene constraints. Thus, a scene structure can be reconstructed
the literature, validate our proposed methods. q 2004 Elsevier B.V.
```

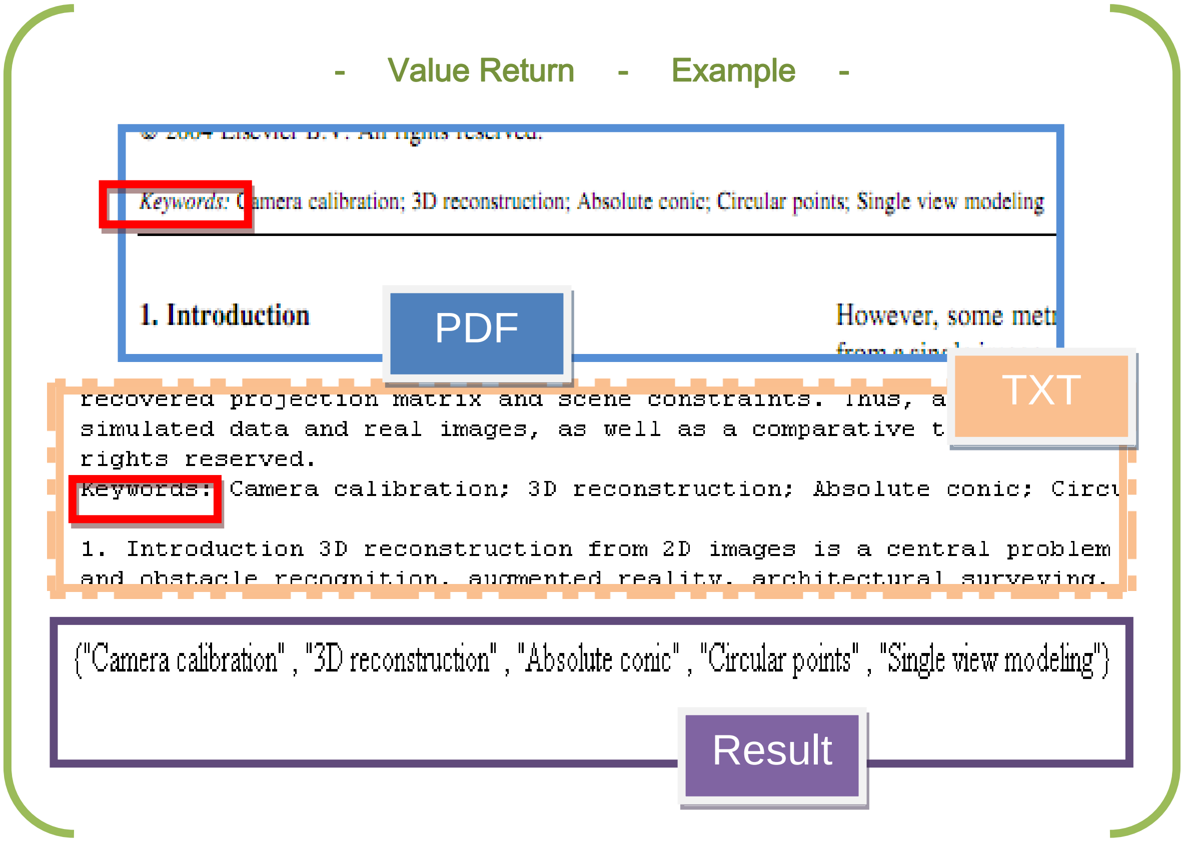
Keywords -

If there is any keyword, most of them appear after several terms indicate the keywords. The terms indicate keywords varies, like “Keywords”, “key terms”, “Index terms” and “General terms”. All of them should be checked. Similar to the case of “Abstract”, the keywords can be found by locating the specific terms and end by the presence of “Introduction” (implementation methods 2 & 4). More to note that, each keyword or key phrase are usually separated by “,” or “;”.

- Pseudo code -

```
WHILE (a line is parsed)
{       IF (“Keywords” is found)
        {       GET Keywords SPLIT BY “,” or “;”
                UNTIL “Introduction” appears } }
UPDATE the database with Keywords got
RETURN Keywords in json
```

Since there will be a list of keywords, it will be encoded as json array to ensure correct data transmission.



© 2004 Elsevier B.V. All rights reserved.

Keywords: Camera calibration; 3D reconstruction; Absolute conic; Circular points; Single view modeling

1. Introduction

However, some metr
from a sing

```
recovered projection matrix and scene constraints. Thus, a
simulated data and real images, as well as a comparative t
rights reserved.
Keywords: Camera calibration; 3D reconstruction; Absolute conic; Circu

1. Introduction 3D reconstruction from 2D images is a central problem
and obstacle recognition, augmented reality, architectural surveying,
```

{"Camera calibration" , "3D reconstruction" , "Absolute conic" , "Circular points" , "Single view modeling"}

References -

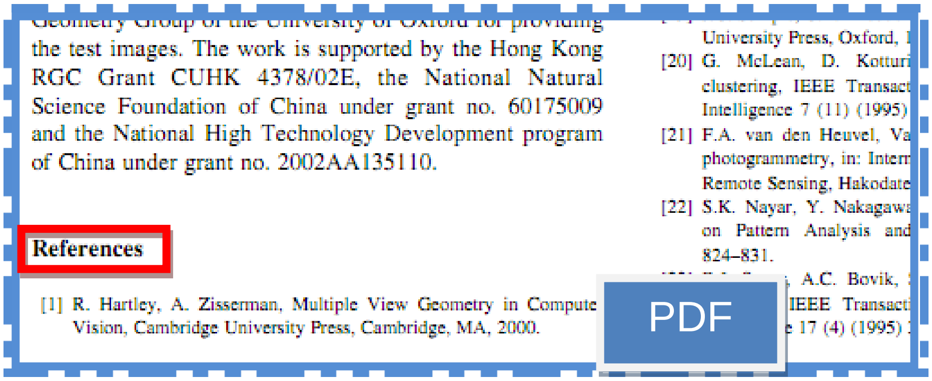
The references appear 100% after the word “REFERENCES” / “BIBLIOGRAPHY”. According to survey, 90% of references have shown in the index form [ ]. Some shows in the form “1. ” while the rest have only the reference items without index. In phase 1, we would like to parse the type of [ ] and those without indexes.

After locating the “References”, we will check if “[“ appears. If “[“ appears and is followed by an index (sometimes may not only digits but also alphabets) and “]”, we record the index with relevant reference item. If there is no “[“ appears within 10 alphabets, we record each line as one reference item.

- Pseudo code -

```
WHILE (a line is parsed)
{       IF (“References” / “Bibliography”is found)
        {       IF (“[“ is found within 10 alphabets]
                {       GET index within [ ]
                        GET relevant reference item }
                ELSE
                {       GET a line AS a reference item } }
UPDATE the database with References got
RETURN References in json
```

- Value Return - Example -

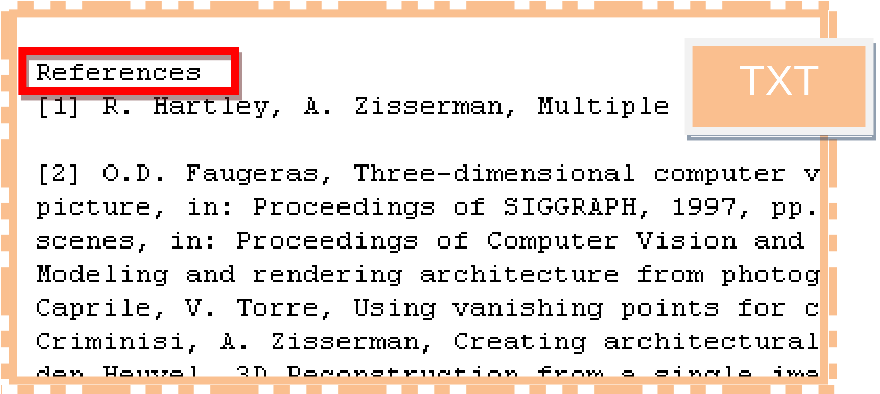


Geometry Group of the University of Oxford for providing the test images. The work is supported by the Hong Kong RGC Grant CUHK 4378/02E, the National Natural Science Foundation of China under grant no. 60175009 and the National High Technology Development program of China under grant no. 2002AA135110.

References

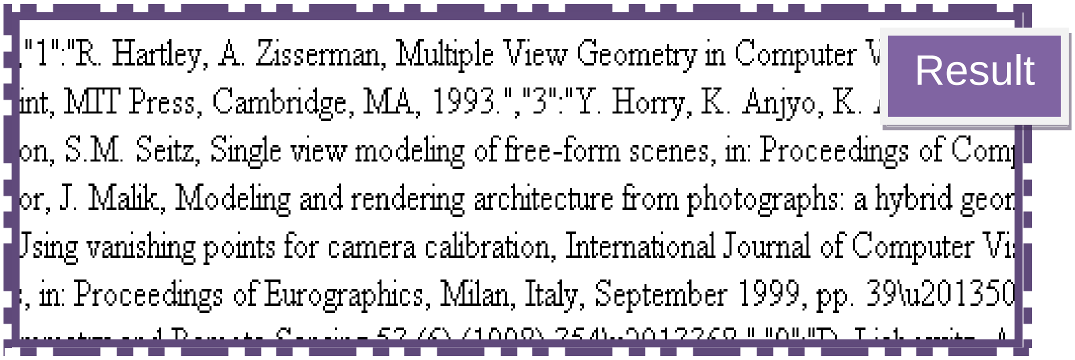
[1] R. Hartley, A. Zisserman, Multiple View Geometry in Compute
Vision, Cambridge University Press, Cambridge, MA, 2000.

University Press, Oxford,
[20] G. McLean, D. Kotturi
clustering, IEEE Transact
Intelligence 7 (11) (1995)
[21] F.A. van den Heuvel, Va
photogrammetry, in: Intern
Remote Sensing, Hakodate
[22] S.K. Nayar, Y. Nakagawa
on Pattern Analysis and
824–831.
, A.C. Bovik,
IEEE Transacti
e 17 (4) (1995)



```
References
[1] R. Hartley, A. Zisserman, Multiple

[2] O.D. Faugeras, Three-dimensional computer v
picture, in: Proceedings of SIGGRAPH, 1997, pp.
scenes, in: Proceedings of Computer Vision and
Modeling and rendering architecture from photog
Caprile, V. Torre, Using vanishing points for c
Criminisi, A. Zisserman, Creating architectural
den Heuvel, 3D Reconstruction from a single ima
```



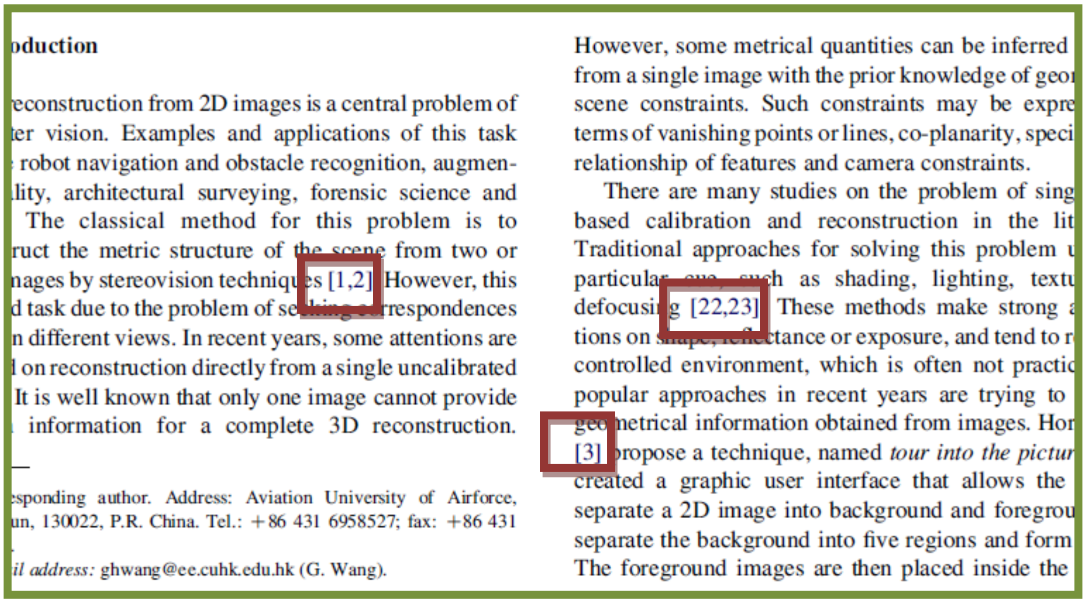
```
,"1":"R. Hartley, A. Zisserman, Multiple View Geometry in Computer V
int, MIT Press, Cambridge, MA, 1993.","3":"Y. Horry, K. Anjyo, K.
on, S.M. Seitz, Single view modeling of free-form scenes, in: Proceedings of Com
or, J. Malik, Modeling and rendering architecture from photographs: a hybrid geor
Jsing vanishing points for camera calibration, International Journal of Computer Vi
, in: Proceedings of Eurographics, Milan, Italy, September 1999, pp. 39\u201350
```

6.1.3. Referencing

Show the reference

A) SERVER SIDE –

To assist the search of reference indexes and reduce the workload of the client side, the server will parse the whole paper and get a list of reference indexes throughout the passage. For example:



oduction

econstruction from 2D images is a central problem of
ter vision. Examples and applications of this task
robot navigation and obstacle recognition, augmen-
lity, architectural surveying, forensic science and
The classical method for this problem is to
ruct the metric structure of the scene from two or
nages by stereovision techniques [1,2] However, this
d task due to the problem of seeking correspondences
n different views. In recent years, some attentions are
l on reconstruction directly from a single uncalibrated
It is well known that only one image cannot provide
information for a complete 3D reconstruction.

esponding author. Address: Aviation University of Airforce,
un, 130022, P.R. China. Tel.: +86 431 6958527; fax: +86 431

il address: ghwang@ee.cuhk.edu.hk (G. Wang).

However, some metrical quantities can be inferred
from a single image with the prior knowledge of geo
scene constraints. Such constraints may be expre
terms of vanishing points or lines, co-planarity, speci
relationship of features and camera constraints.

There are many studies on the problem of sing
based calibration and reconstruction in the lit
Traditional approaches for solving this problem u
particular cue, such as shading, lighting, textu
defocusing [22,23] These methods make strong a
tions on shape, reflectance or exposure, and tend to r
controlled environment, which is often not practic
popular approaches in recent years are trying to
geometrical information obtained from images. Hor
[3] propose a technique, named *tour into the pictur*
created a graphic user interface that allows the
separate a 2D image into background and foregrou
separate the background into five regions and form
The foreground images are then placed inside the

Figure 6.7 Photo showing the references inside the paper

Then the [1,2] ,[22,23] and [3] will be parse into the reference appearance list. The list will be returned to the client directly when client request.

B) CLIENT SIDE –

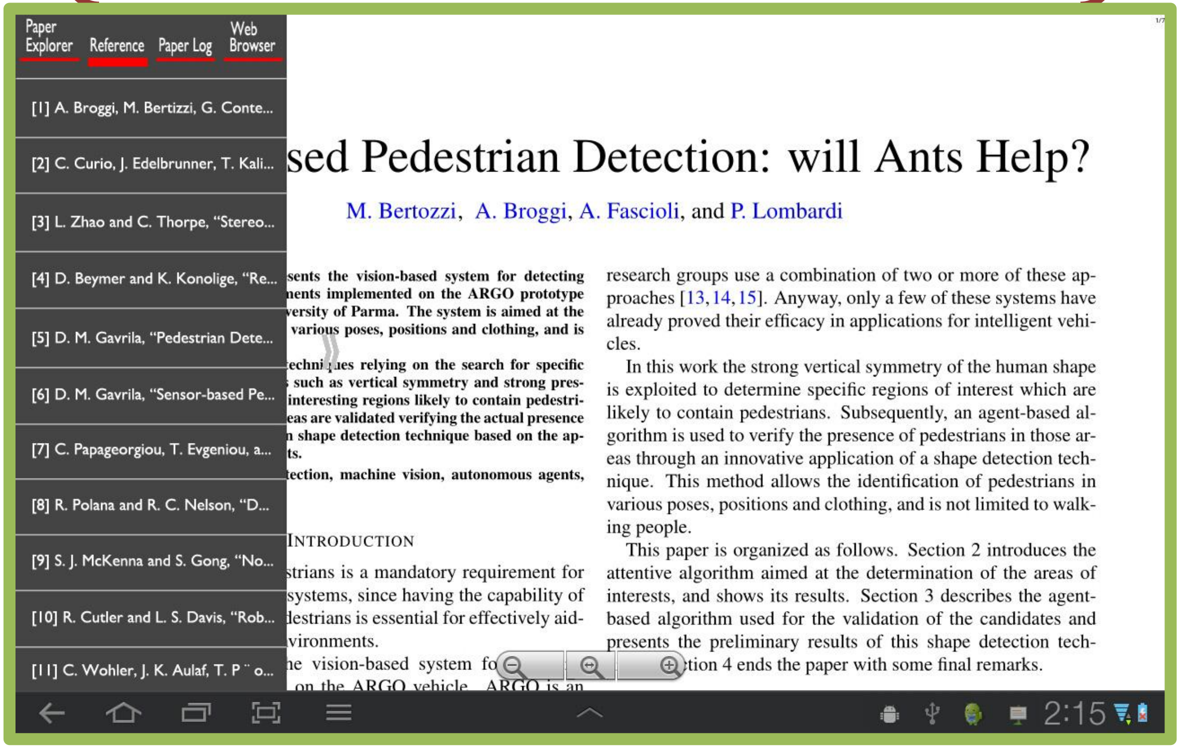
The client device will first ask the web server to send the list of references to it. The list of references will be shown in the transparent panel in the left hand side of the screen. Furthermore, the user can click on the buttons of references on the paper to show what the reference is. After clicking, the information of the reference will be shown.

- Pseudo code -

```
SHOW the list of references in the transparent panel

IF (choose one of reference buttons)

{  SHOW the information of that reference }        }
```



[1] A. Broggi, M. Bertizzi, G. Conte...

[2] C. Curio, J. Edelbrunner, T. Kali...

[3] L. Zhao and C. Thorpe, "Stereo...

[4] D. Beymer and K. Konolige, "Re...

[5] D. M. Gavrila, "Pedestrian Dete...

[6] D. M. Gavrila, "Sensor-based Pe...

[7] C. Papageorgiou, T. Evgeniou, a...

[8] R. Polana and R. C. Nelson, "D...

[9] S. J. McKenna and S. Gong, "No...

[10] R. Cutler and L. S. Davis, "Rob...

[11] C. Wohler, J. K. Aulaf, T. P ¨ o...

sed Pedestrian Detection: will Ants Help?

M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi

**sents the vision-based system for detecting
nents implemented on the ARGO prototype
ersity of Parma. The system is aimed at the
various poses, positions and clothing, and is**

**echniques relying on the search for specific
such as vertical symmetry and strong pres-
interesting regions likely to contain pedestri-
eas are validated verifying the actual presence
n shape detection technique based on the ap-
ts.**

ection, machine vision, autonomous agents,

INTRODUCTION

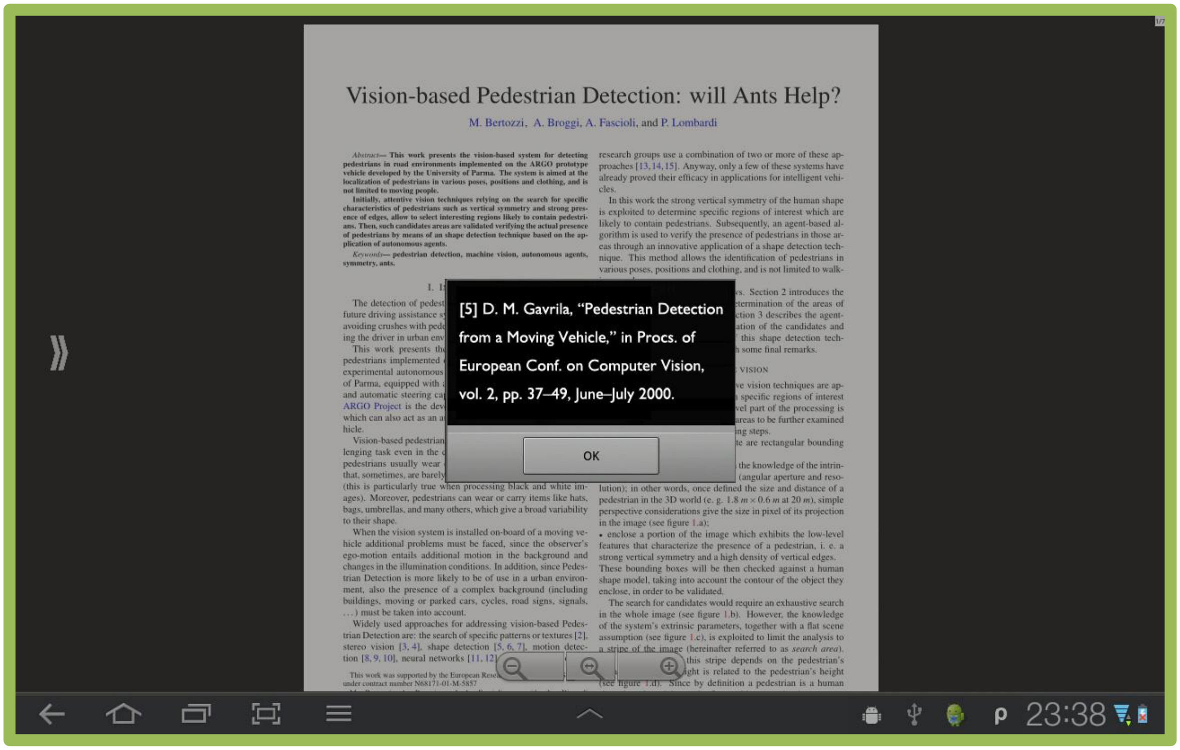
strians is a mandatory requirement for
systems, since having the capability of
lestrians is essential for effectively aid-
vironments.
he vision-based system fo
on the ARGO vehicle. ARGO is an

research groups use a combination of two or more of these approaches [13, 14, 15]. Anyway, only a few of these systems have already proved their efficacy in applications for intelligent vehicles.

In this work the strong vertical symmetry of the human shape is exploited to determine specific regions of interest which are likely to contain pedestrians. Subsequently, an agent-based algorithm is used to verify the presence of pedestrians in those areas through an innovative application of a shape detection technique. This method allows the identification of pedestrians in various poses, positions and clothing, and is not limited to walking people.

This paper is organized as follows. Section 2 introduces the attentive algorithm aimed at the determination of the areas of interests, and shows its results. Section 3 describes the agent-based algorithm used for the validation of the candidates and presents the preliminary results of this shape detection tech-
tion 4 ends the paper with some final remarks.

Figure 6.9 Screen capture of reference list



Vision-based Pedestrian Detection: will Ants Help?

M. Bertozzi, A. Broggi, A. Fascioli, and P. Lombardi

Abstract— This work presents the vision-based system for detecting pedestrians in road environments implemented on the ARGO prototype vehicle developed by the University of Parma. The system is aimed at the localization of pedestrians in various poses, positions and clothing, and is not limited to moving people.

Initially, attentive vision techniques relying on the search for specific characteristics of pedestrians such as vertical symmetry and strong presence of edges, allow to select interesting regions likely to contain pedestrians. Then, such candidates areas are validated verifying the actual presence of pedestrians by means of an shape detection technique based on the application of autonomous agents.

Keywords— pedestrian detection, machine vision, autonomous agents, symmetry, ants.

1. I

The detection of pedest
future driving assistance s
avoiding crushes with pede
ing the driver in urban env

This work presents th
pedestrians implemented
experimental autonomous
of Parma, equipped with
and automatic steering ca
ARGO Project is the dev
which can also act as an a
hicle.

Vision-based pedestrian
lenging task even in the
pedestrians usually wear
that, sometimes, are barely
(this is particularly true when processing black and white images). Moreover, pedestrians can wear or carry items like hats, bags, umbrellas, and many others, which give a broad variability to their shape.

When the vision system is installed on-board of a moving vehicle additional problems must be faced, since the observer's ego-motion entails additional motion in the background and changes in the illumination conditions. In addition, since Pedestrian Detection is more likely to be of use in a urban environment, also the presence of a complex background (including buildings, moving or parked cars, cycles, road signs, signals, ...) must be taken into account.

Widely used approaches for addressing vision-based Pedestrian Detection are: the search of specific patterns or textures [2], stereo vision [3, 4], shape detection [5, 6, 7], motion detection [8, 9, 10], neural networks [11, 12]

This work was supported by the European Rese
under contract number N68171-01-M-5857

research groups use a combination of two or more of these approaches [13, 14, 15]. Anyway, only a few of these systems have already proved their efficacy in applications for intelligent vehicles.

In this work the strong vertical symmetry of the human shape is exploited to determine specific regions of interest which are likely to contain pedestrians. Subsequently, an agent-based algorithm is used to verify the presence of pedestrians in those areas through an innovative application of a shape detection technique. This method allows the identification of pedestrians in various poses, positions and clothing, and is not limited to walk-

s. Section 2 introduces the
termination of the areas of
ction 3 describes the agent-
ation of the candidates and
this shape detection tech-
h some final remarks.

VISION

ve vision techniques are ap-
specific regions of interest
vel part of the processing is
areas to be further examined
ing steps.
te are rectangular bounding

the knowledge of the intrin-
(angular aperture and resolution); in other words, once defined the size and distance of a pedestrian in the 3D world (e. g. 1.8 $m \times 0.6$ m at 20 m), simple perspective considerations give the size in pixel of its projection in the image (see figure 1.a);

- enclose a portion of the image which exhibits the low-level features that characterize the presence of a pedestrian, i. e. a strong vertical symmetry and a high density of vertical edges.

These bounding boxes will be then checked against a human shape model, taking into account the contour of the object they enclose, in order to be validated.

The search for candidates would require an exhaustive search in the whole image (see figure 1.b). However, the knowledge of the system's extrinsic parameters, together with a flat scene assumption (see figure 1.c), is exploited to limit the analysis to a stripe of the image (hereinafter referred to as *search area*).
this stripe depends on the pedestrian's
ght is related to the pedestrian's height
(see figure 1.d). Since by definition a pedestrian is a human

[5] D. M. Gavrila, "Pedestrian Detection from a Moving Vehicle," in Procs. of European Conf. on Computer Vision, vol. 2, pp. 37–49, June–July 2000.

Figure 6.8 Screen capture of showing reference information

6.1.4. Library renewal

Manage the library manual

A) SERVER SIDE –

During the early development, a viewable managing GUI platform is provided for the administrator to use.

There is a drop box for the admin to drag the PDF file and upload it. Once the file is dropped and successfully uploaded, a new row will be inserted at the top. The txt file will be automatically generated for parsing. The newly uploaded file will be parsed too. On the panel, admin can delete the paper too.

6.2. Module 2 : Communication

6.2.1. Posting comments

Add comments

A) SERVER SIDE –

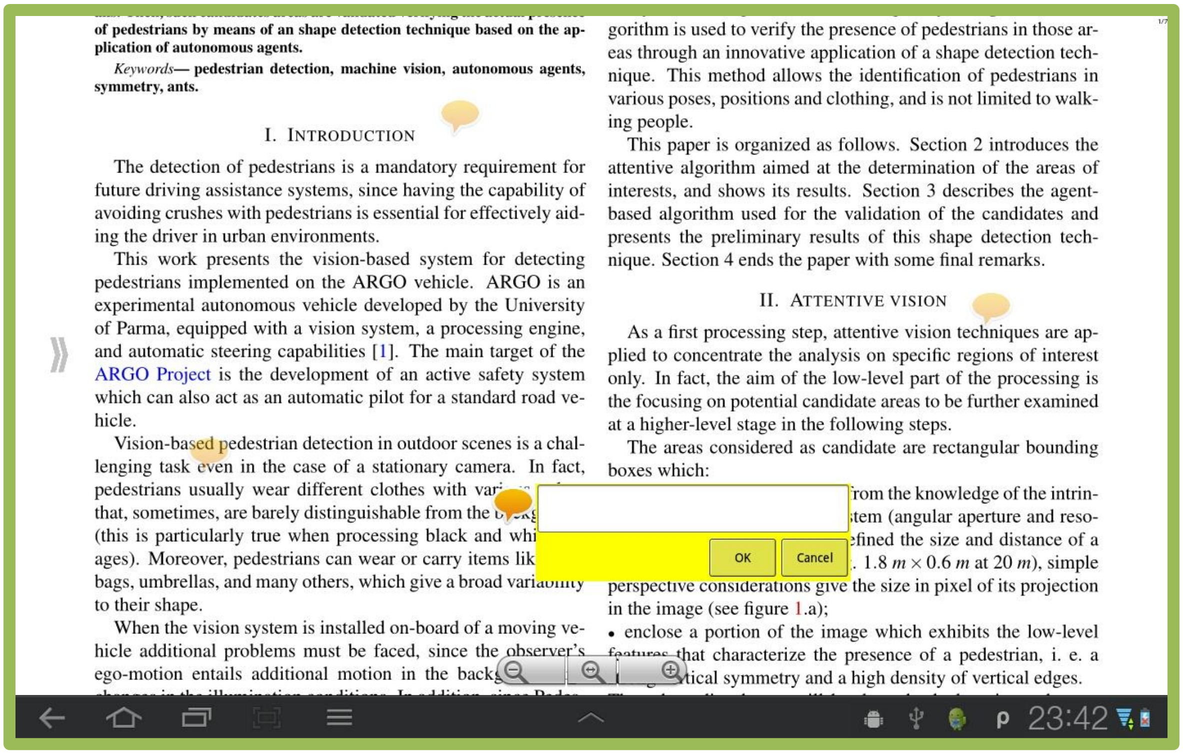
On adding the comment, the client sends the user ID, paper ID, page number, positions and the comment to the server. Server will automatically generate a comment ID and the update time for inserting record into database.

B) CLIENT SIDE –

The user can add comments on the paper by long-pressing on the screen. The client device will then create a comment button at that location, and ask the user to input the comment. After the user input the comment, the client device will send some information, including location of the comment, content of the comment, paper ID, user ID, etc. , to the web server to update the database.

- Pseudo code -

```
IF (add comment)
{  CHECK the location
   ASK the user to input the comment
      CREATE a new button for comment
   SEND the information to the web server }
```



of pedestrians by means of an shape detection technique based on the application of autonomous agents.

Keywords— **pedestrian detection, machine vision, autonomous agents, symmetry, ants.**

I. INTRODUCTION

The detection of pedestrians is a mandatory requirement for future driving assistance systems, since having the capability of avoiding crushes with pedestrians is essential for effectively aiding the driver in urban environments.

This work presents the vision-based system for detecting pedestrians implemented on the ARGO vehicle. ARGO is an experimental autonomous vehicle developed by the University of Parma, equipped with a vision system, a processing engine, and automatic steering capabilities [1]. The main target of the ARGO Project is the development of an active safety system which can also act as an automatic pilot for a standard road vehicle.

Vision-based pedestrian detection in outdoor scenes is a challenging task even in the case of a stationary camera. In fact, pedestrians usually wear different clothes with var that, sometimes, are barely distinguishable from the (this is particularly true when processing black and whi ages). Moreover, pedestrians can wear or carry items lik bags, umbrellas, and many others, which give a broad variability to their shape.

When the vision system is installed on-board of a moving vehicle additional problems must be faced, since the observer's ego-motion entails additional motion in the back

gorithm is used to verify the presence of pedestrians in those areas through an innovative application of a shape detection technique. This method allows the identification of pedestrians in various poses, positions and clothing, and is not limited to walking people.

This paper is organized as follows. Section 2 introduces the attentive algorithm aimed at the determination of the areas of interests, and shows its results. Section 3 describes the agent-based algorithm used for the validation of the candidates and presents the preliminary results of this shape detection technique. Section 4 ends the paper with some final remarks.

II. ATTENTIVE VISION

As a first processing step, attentive vision techniques are applied to concentrate the analysis on specific regions of interest only. In fact, the aim of the low-level part of the processing is the focusing on potential candidate areas to be further examined at a higher-level stage in the following steps.

The areas considered as candidate are rectangular bounding boxes which:

rom the knowledge of the intrin stem (angular aperture and reso efined the size and distance of a . 1.8 $m \times 0.6$ m at 20 m), simple perspective considerations give the size in pixel of its projection in the image (see figure 1.a);

- enclose a portion of the image which exhibits the low-level features that characterize the presence of a pedestrian, i. e. a tical symmetry and a high density of vertical edges.

Figure 6.10 Screen capture of adding comment

Show comments

A) SERVER SIDE –

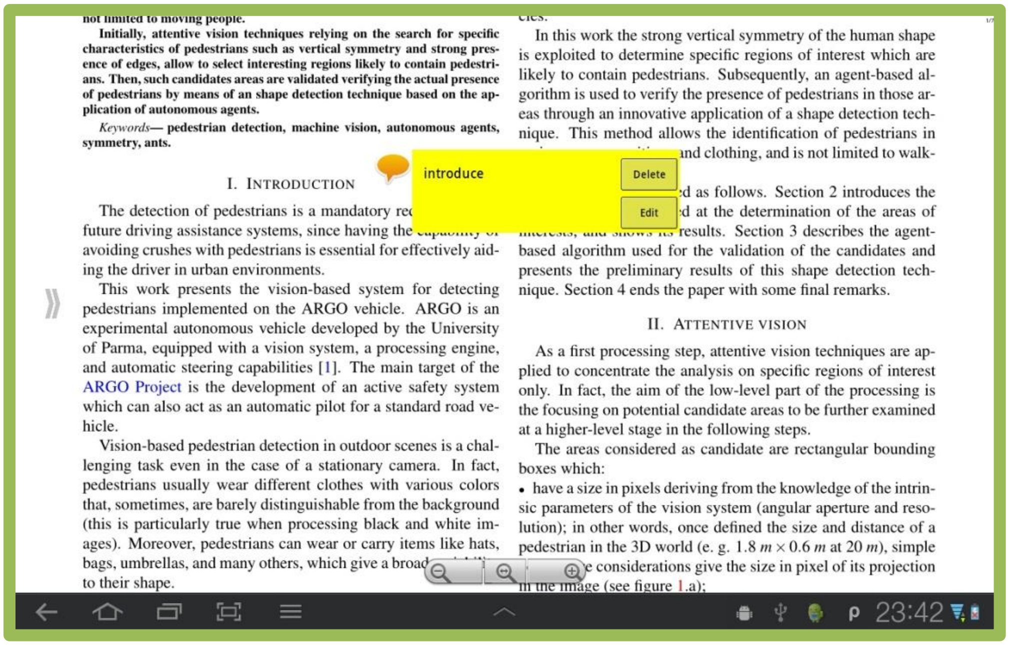
On deleting the comment, the client sends the paper ID, page number and position as request. Server will check return a list of usernames, comments and comment IDs from database to the client.

B) CLIENT SIDE –

Before showing a paper, the client device will ask the web server for the previous comments of that paper. The client device will then create the comment button at the corresponding location, and load the previous comment content. The user can read the comment added by himself or others. Besides, the comment buttons are translucent so that not to block the text of the paper.

- Pseudo code -

```
SHOW paper
CREATE buttons for all previous comments
IF (click on a comment button)
{  SHOW the previous comment contents
   SHOW the button of reply and delete     }
```



not limited to moving people.

Initially, attentive vision techniques relying on the search for specific characteristics of pedestrians such as vertical symmetry and strong presence of edges, allow to select interesting regions likely to contain pedestrians. Then, such candidates areas are validated verifying the actual presence of pedestrians by means of an shape detection technique based on the application of autonomous agents.

Keywords— **pedestrian detection, machine vision, autonomous agents, symmetry, ants.**

I. INTRODUCTION

The detection of pedestrians is a mandatory re
future driving assistance systems, since having the
avoiding crushes with pedestrians is essential for effectively aiding the driver in urban environments.

This work presents the vision-based system for detecting pedestrians implemented on the ARGO vehicle. ARGO is an experimental autonomous vehicle developed by the University of Parma, equipped with a vision system, a processing engine, and automatic steering capabilities [1]. The main target of the ARGO Project is the development of an active safety system which can also act as an automatic pilot for a standard road vehicle.

Vision-based pedestrian detection in outdoor scenes is a challenging task even in the case of a stationary camera. In fact, pedestrians usually wear different clothes with various colors that, sometimes, are barely distinguishable from the background (this is particularly true when processing black and white images). Moreover, pedestrians can wear or carry items like hats, bags, umbrellas, and many others, which give a broad
to their shape.

In this work the strong vertical symmetry of the human shape is exploited to determine specific regions of interest which are likely to contain pedestrians. Subsequently, an agent-based algorithm is used to verify the presence of pedestrians in those areas through an innovative application of a shape detection technique. This method allows the identification of pedestrians in
and clothing, and is not limited to walk-

d as follows. Section 2 introduces the
d at the determination of the areas of
results. Section 3 describes the agent-based algorithm used for the validation of the candidates and presents the preliminary results of this shape detection technique. Section 4 ends the paper with some final remarks.

II. ATTENTIVE VISION

As a first processing step, attentive vision techniques are applied to concentrate the analysis on specific regions of interest only. In fact, the aim of the low-level part of the processing is the focusing on potential candidate areas to be further examined at a higher-level stage in the following steps.

The areas considered as candidate are rectangular bounding boxes which:

- have a size in pixels deriving from the knowledge of the intrinsic parameters of the vision system (angular aperture and resolution); in other words, once defined the size and distance of a pedestrian in the 3D world (e. g. 1.8 $m \times 0.6$ m at 20 m), simple
e considerations give the size in pixel of its projection in the image (see figure 1.a);

Figure 6.10 Screen capture of showing comment

Reply comments

A) SERVER SIDE –

The client sends the user ID, paper ID, page number, positions and the comment to the server. By checking the database, server will recognize that there is a discussion invoked on the same point before. It will hence add the comment with an additional entry "ReplyTo", which records the comment ID the new comment replies to.

B) CLIENT SIDE –

User can reply on any comment added by himself or others. After replying, the message will be sent to the web server to update the database.

- Pseudo code -

```
IF (click reply button)
{  ASK user to input the reply message
   SEND the information to the web server }
```

Delete comments

A) SERVER SIDE –

The client sends the comment ID to the server. With reference to the database, if the CID is the first comment invoking the whole discussion, all the comments of the same discussion will be deleted. Yet, only the comment of that CID will be deleted. Meanwhile, the other comment which have “ReplyTo” points the deleted CID with be redirected to the previous comment that the reply to point to.

B) CLIENT SIDE –

User can delete his own comments and replies, the information will be sent to the web server to update the database.

- Pseudo code -

```
IF (delete comment)

{  SEND the information to the web server }

IF (delete reply)

{  SEND the information to the web server }
```

6.3. Module 3 : Research Log

6.3.1. User accounts

Login to system

A) SERVER SIDE –

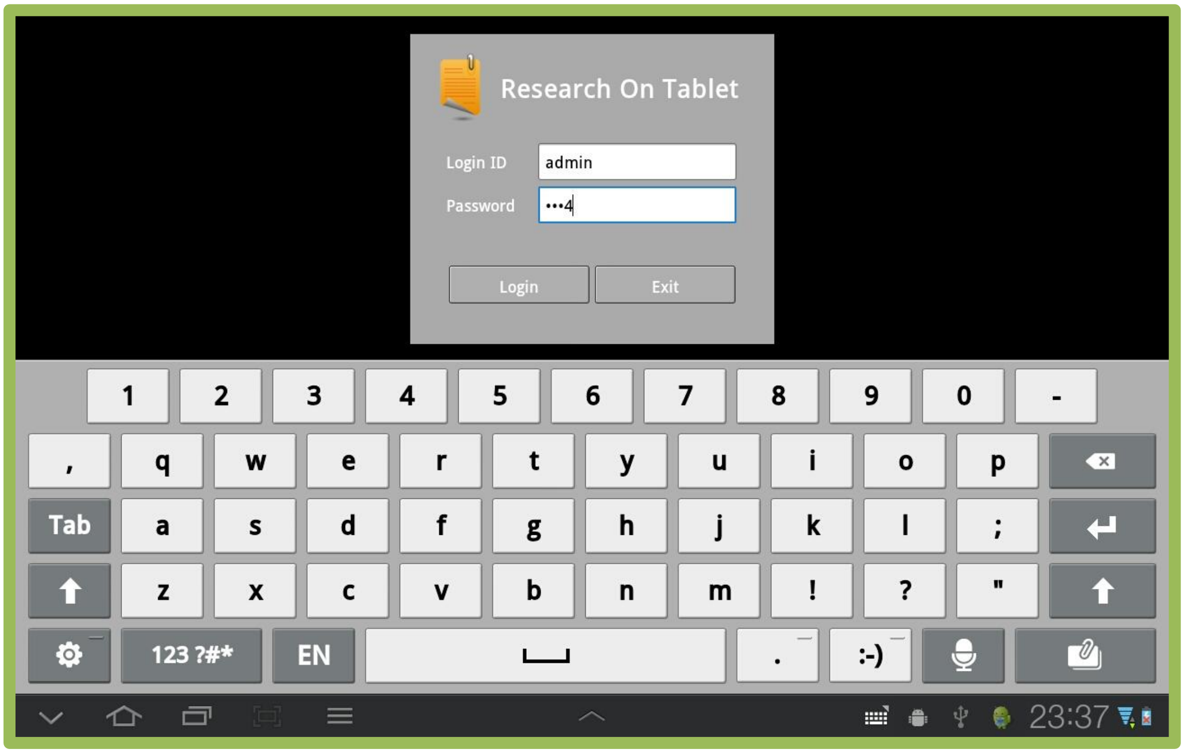
When the user logins to system, the client sends the user ID and password to the server. The server will check if the login succeeds, the user name will be returned to the client.

B) CLIENT SIDE –

User is required to login to use this application, which is required to input the login ID and password, the client devices will then send these information to the web server to check if the login validates. If it succeeds, the user is able to choose a paper to read and use the application normally, otherwise, an error message about login failure will be shown, and the client device asks the user to input the information again.

- Pseudo code -

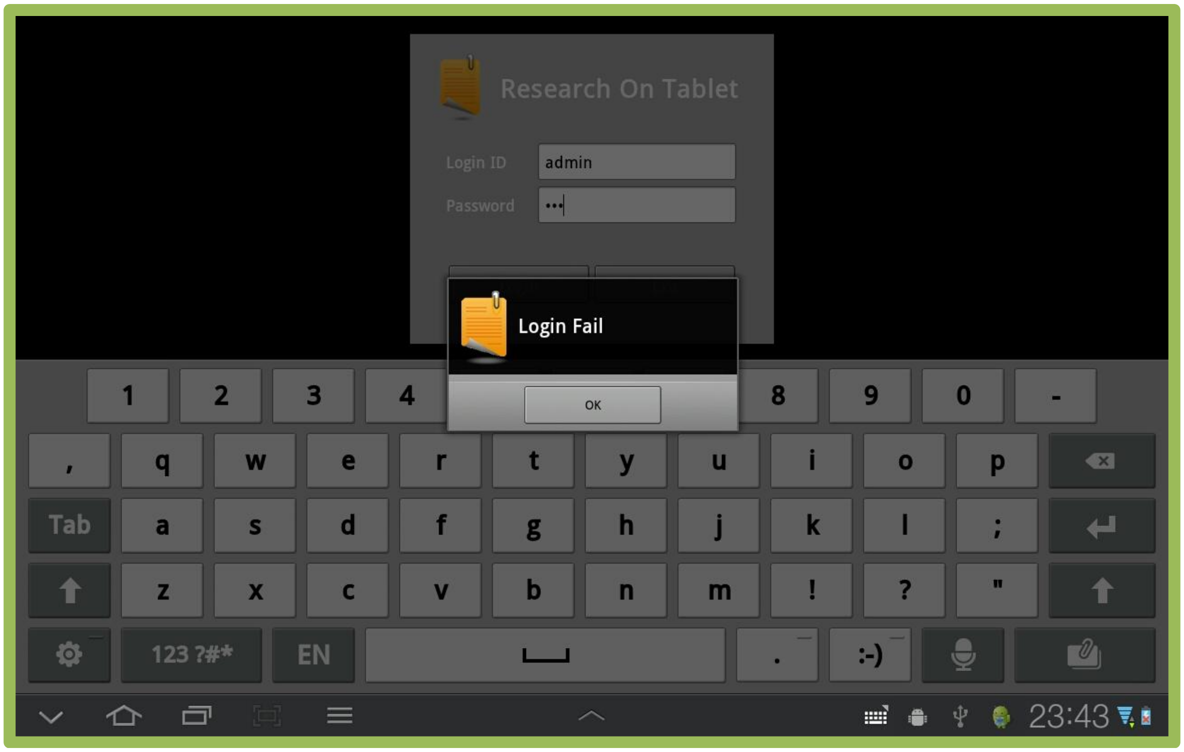
```
IF (open the application)
{  Ask the user to input login ID and password
   Send the login ID and password to web server
   GET the result from the web server
   IF (validate)
   {        ALLOW the user to user the application  }
   Else
   {        SHOW error message
            ASK the user to login again        }        }
```


Research On Tablet

Login ID: admin

Password: •••4

Figure 6.11 Screen capture of login



Login ID admin

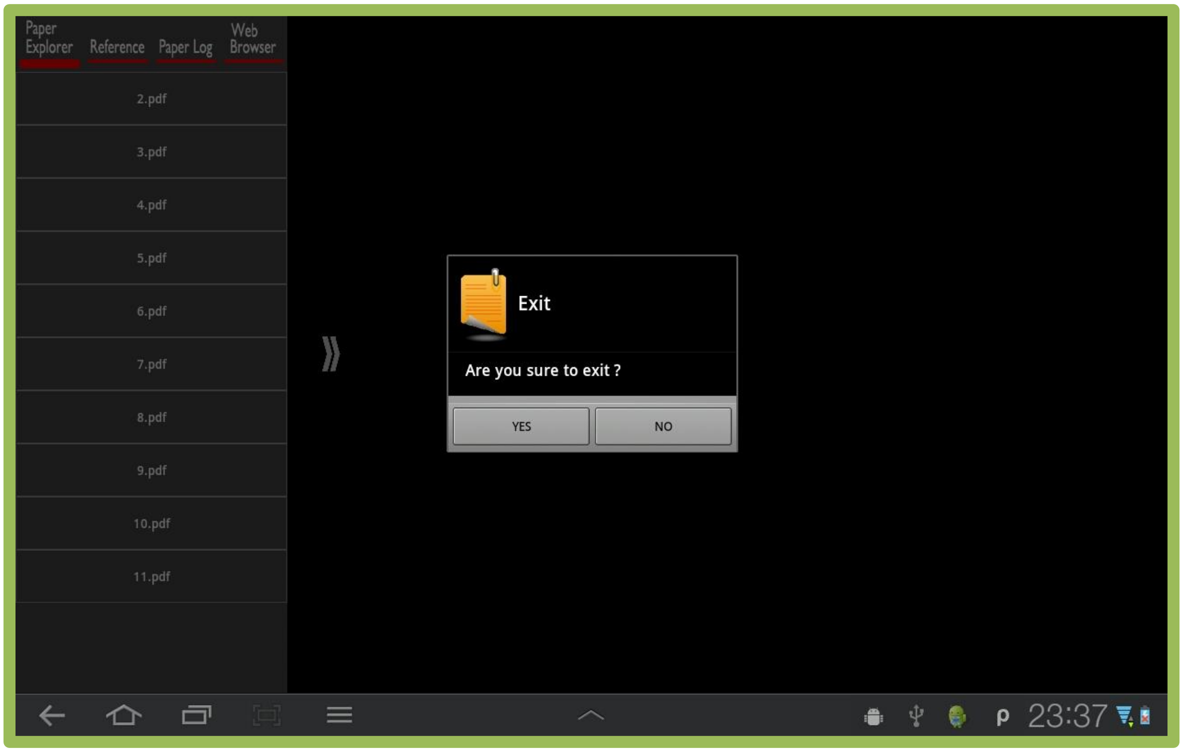
Password •••

Figure 6.12 Screen capture of login fail

If the user want to exit the application, the client device will ask the user to confirm his exiting. If the user confirms to exit, he will logout from the application and the application will close. Otherwise, the client device will return to the previous view and let the user continue to use.

- Pseudo code -

```
IF (choose exit)
{  ASK the user to confirm
   IF (confirm)
   {        EXIT the application}
   ELSE
   {        STAY on that page }        }
```



Exit

Are you sure to exit ?

Figure 6.13 Screen capture of exiting

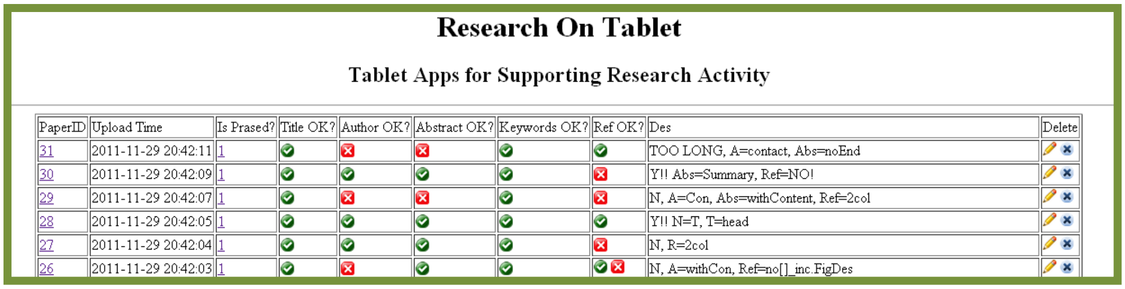
7. Testing and Evaluation

In this section, we are going to show the testing result and evaluation on them. We will test the server and client side separately. Both of them focus on different aspects. For the server, the testing focuses on the parsing performance of the papers. While for the client, the testing focuses on the functional performance of the application in the tablet devices.

7.1. Server Side

7.1.1. Testing Methodology

To test the parsing performance, we modified the managing panel designed for the administrator to capture the data easily. With the panel, the data call be represented in the table form, so as to make everything clearly.



Research On Tablet

Tablet Apps for Supporting Research Activity

PaperID	Upload Time	Is Prased?	Title OK?	Author OK?	Abstract OK?	Keywords OK?	Ref OK?	Des	Delete
31	2011-11-29 20:42:11	1	✓	☒	☒	✓	✓	TOO LONG, A=contact, Abs=noEnd	✎ ✖
30	2011-11-29 20:42:09	1	✓	✓	✓	✓	☒	Y!! Abs=Summary, Ref=NO!	✎ ✖
29	2011-11-29 20:42:07	1	✓	☒	☒	✓	☒	N, A=Con, Abs=withContent, Ref=2col	✎ ✖
28	2011-11-29 20:42:05	1	✓	✓	✓	✓	✓	Y!! N=T, T=head	✎ ✖
27	2011-11-29 20:42:04	1	✓	✓	✓	✓	☒	N, R=2col	✎ ✖
26	2011-11-29 20:42:03	1	✓	☒	✓	✓	✓ ☒	N, A=withCon, Ref=no[]_inc.FigDes	✎ ✖

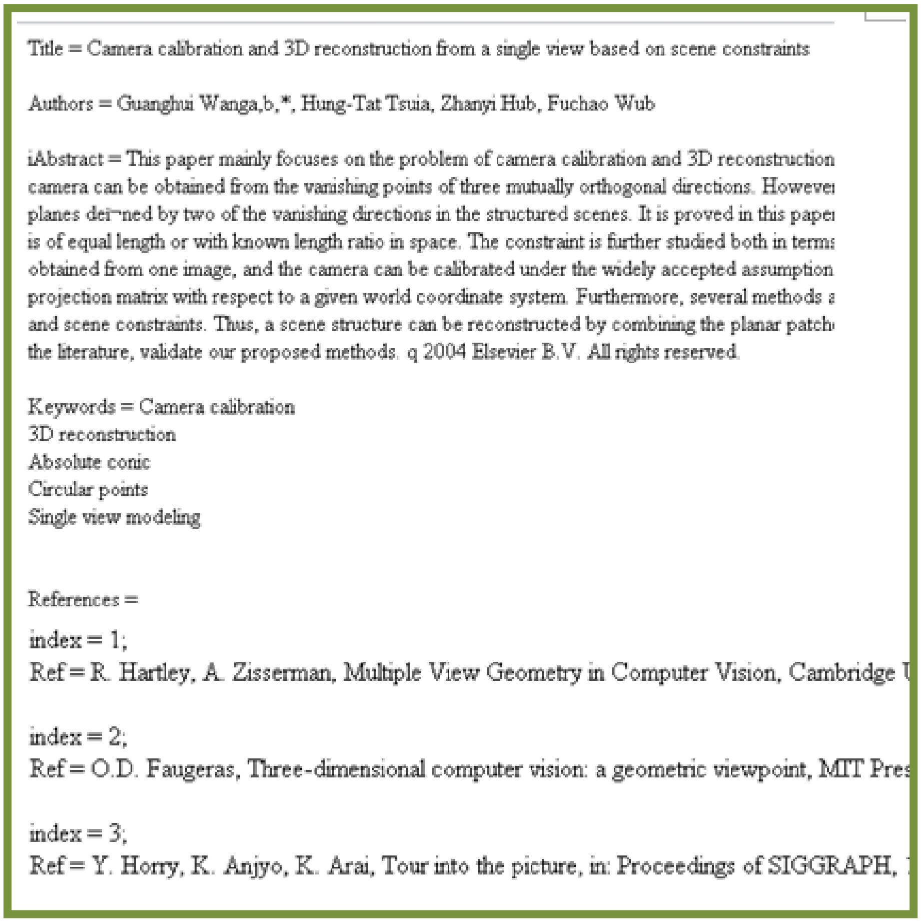
Figure 7.1 Admin Panel on web server

.

There are eight columns designed for the testing.

Data Viewing

The paper file can also be view by clicking the Paper ID, while the “IsPrased” column provides a link to view the parsing results. The page shows all the entries that parsed into the database.



Title = Camera calibration and 3D reconstruction from a single view based on scene constraints

Authors = Guanghui Wanga,b,*, Hung-Tat Tsuia, Zhanyi Hub, Fuchao Wub

iAbstract = This paper mainly focuses on the problem of camera calibration and 3D reconstruction
camera can be obtained from the vanishing points of three mutually orthogonal directions. Howeve
planes dei¬ned by two of the vanishing directions in the structured scenes. It is proved in this pape
is of equal length or with known length ratio in space. The constraint is further studied both in term
obtained from one image, and the camera can be calibrated under the widely accepted assumption
projection matrix with respect to a given world coordinate system. Furthermore, several methods
and scene constraints. Thus, a scene structure can be reconstructed by combining the planar patch
the literature, validate our proposed methods. q 2004 Elsevier B.V. All rights reserved.

Keywords = Camera calibration
3D reconstruction
Absolute conic
Circular points
Single view modeling

References =

index = 1;
Ref = R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge

index = 2;
Ref = O.D. Faugeras, Three-dimensional computer vision: a geometric viewpoint, MIT Pre

index = 3;
Ref = Y. Horry, K. Anjyo, K. Arai, Tour into the picture, in: Proceedings of SIGGRAPH,

Result Recording

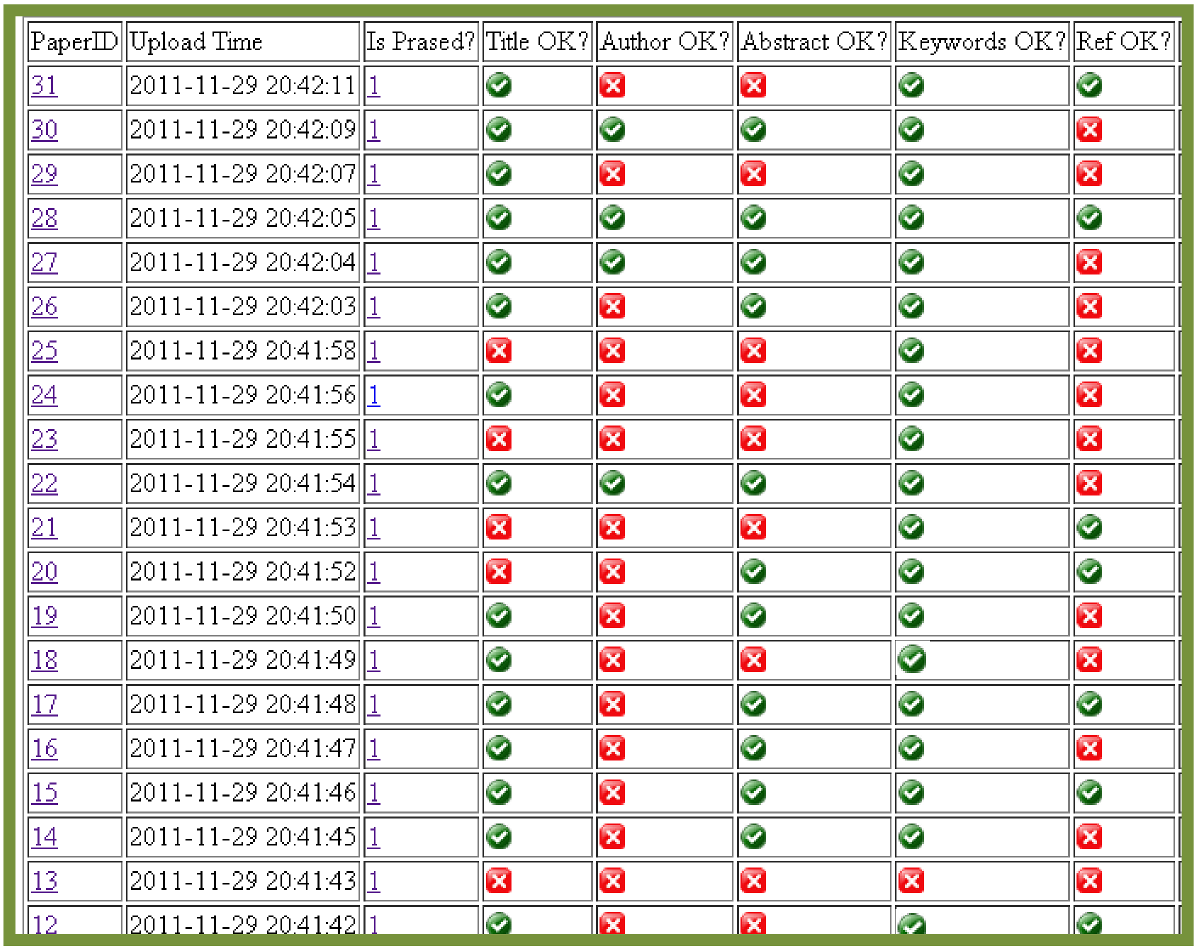
There are five columns available to mark the tick or cross for each entry. After checking the results, we mark down whether the results given are successful or not.

Remarks Marking

There is a “Des” to show the remarks. We check the results after parsing, then try to find out the reason of failure and mark it down. The remarks can be edited by clicking the pencil icon on the right.

7.1.2. Data Description

We have parsed twenty PDF files for testing the parsing performance. The following figure shows our results:



PaperID	Upload Time	Is Prased?	Title OK?	Author OK?	Abstract OK?	Keywords OK?	Ref OK?
31	2011-11-29 20:42:11	1	✓	✗	✗	✓	✓
30	2011-11-29 20:42:09	1	✓	✓	✓	✓	✗
29	2011-11-29 20:42:07	1	✓	✗	✗	✓	✗
28	2011-11-29 20:42:05	1	✓	✓	✓	✓	✓
27	2011-11-29 20:42:04	1	✓	✓	✓	✓	✗
26	2011-11-29 20:42:03	1	✓	✗	✓	✓	✗
25	2011-11-29 20:41:58	1	✗	✗	✗	✓	✗
24	2011-11-29 20:41:56	1	✓	✗	✗	✓	✗
23	2011-11-29 20:41:55	1	✗	✗	✗	✓	✗
22	2011-11-29 20:41:54	1	✓	✓	✓	✓	✗
21	2011-11-29 20:41:53	1	✗	✗	✗	✓	✓
20	2011-11-29 20:41:52	1	✗	✗	✓	✓	✓
19	2011-11-29 20:41:50	1	✓	✗	✓	✓	✗
18	2011-11-29 20:41:49	1	✓	✗	✗	✓	✗
17	2011-11-29 20:41:48	1	✓	✗	✓	✓	✓
16	2011-11-29 20:41:47	1	✓	✗	✓	✓	✗
15	2011-11-29 20:41:46	1	✓	✗	✓	✓	✓
14	2011-11-29 20:41:45	1	✓	✗	✓	✓	✗
13	2011-11-29 20:41:43	1	✗	✗	✗	✗	✗
12	2011-11-29 20:41:42	1	✓	✗	✗	✓	✓

Percentage of successful TITLE Parsing = 15/20 = 75%

Percentage of successful AUTHOR Parsing = 5/20 = 25%

Percentage of successful ABSTRACT Parsing = 11/20 = 55%

Percentage of successful KEYWORDS Parsing = 3/20 = 85%

Percentage of successful REFENENCE Parsing = 7/20 = 35%

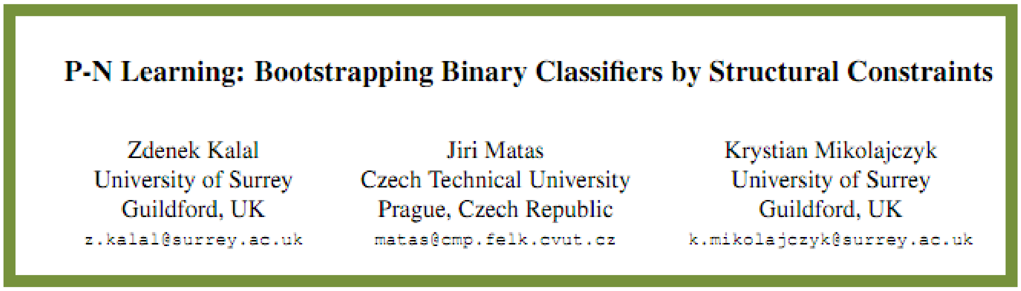
7.1.3. Evaluation

Title

The percentage of successful cases (75%) is satisfactory. Throughout the fail cases, most of them resulted in the malfunction of getting font size function in PHP. Within 20 papers, there are totally 4 papers with header goes before the title and only one of them success. The fail case returns the header as the title. Besides, one of failure (PID=25) resulted because of the wrong arrangement of the title and author. The author in this paper goes before the title. The paper with PID=13 fails because the “pdftotext” function does not support the UTF-8 encoding.

Author

The percentage of successful cases (25%) is disappointing. Part of the failure is caused because of the failure of parsing title (since our implementation method of the authors is parsing the “authors” right after the “title”.).For the majority of the fail cases, the failures are resulted because of the format of PDF file. Take the following case as examples:



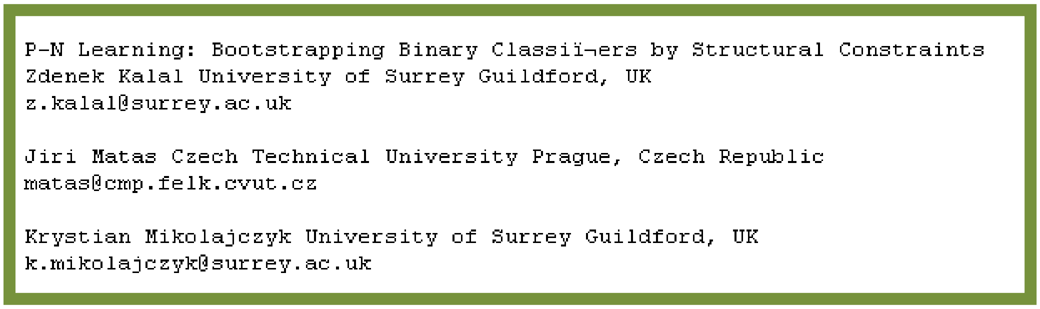
P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints

Zdenek Kalal
University of Surrey
Guildford, UK
z.kalal@surrey.ac.uk

Jiri Matas
Czech Technical University
Prague, Czech Republic
matas@cmp.felk.cvut.cz

Krystian Mikolajczyk
University of Surrey
Guildford, UK
k.mikolajczyk@surrey.ac.uk

In this case, the authors are parsed to plain text as:



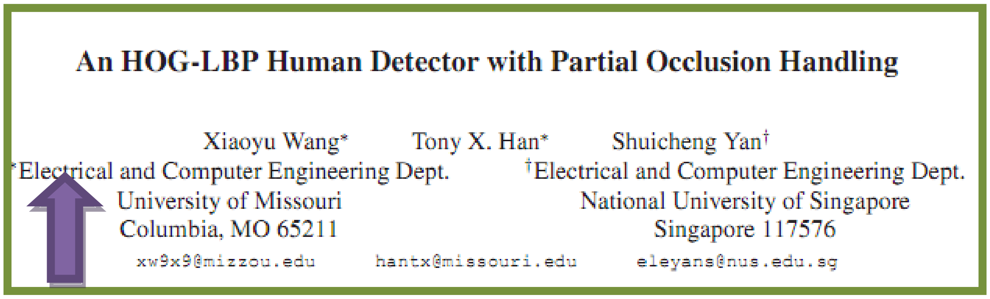
```
P-N Learning: Bootstrapping Binary Classiï¬ers by Structural Constraints
Zdenek Kalal University of Surrey Guildford, UK
z.kalal@surrey.ac.uk

Jiri Matas Czech Technical University Prague, Czech Republic
matas@cmp.felk.cvut.cz

Krystian Mikolajczyk University of Surrey Guildford, UK
k.mikolajczyk@surrey.ac.uk
```

And this comes with difficulties that have been mentioned in the previous chapters: there has no standard to capture the difference between the authors' name and the descriptions like the authors' titles and contact information.

There is another case with the authors' descriptions. Let's take a look at the following case:



An HOG-LBP Human Detector with Partial Occlusion Handling

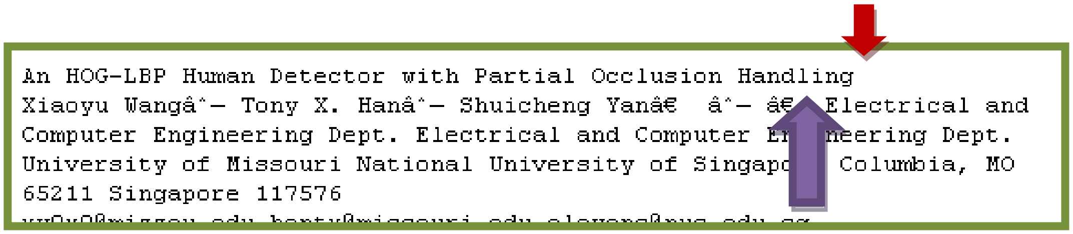
Xiaoyu Wang* Tony X. Han* Shuicheng Yan†

*Electrical and Computer Engineering Dept.
University of Missouri
Columbia, MO 65211

†Electrical and Computer Engineering Dept.
National University of Singapore
Singapore 117576

xw9x9@mizzou.edu hantx@missouri.edu eleyans@nus.edu.sg

And we have the following text parsed:

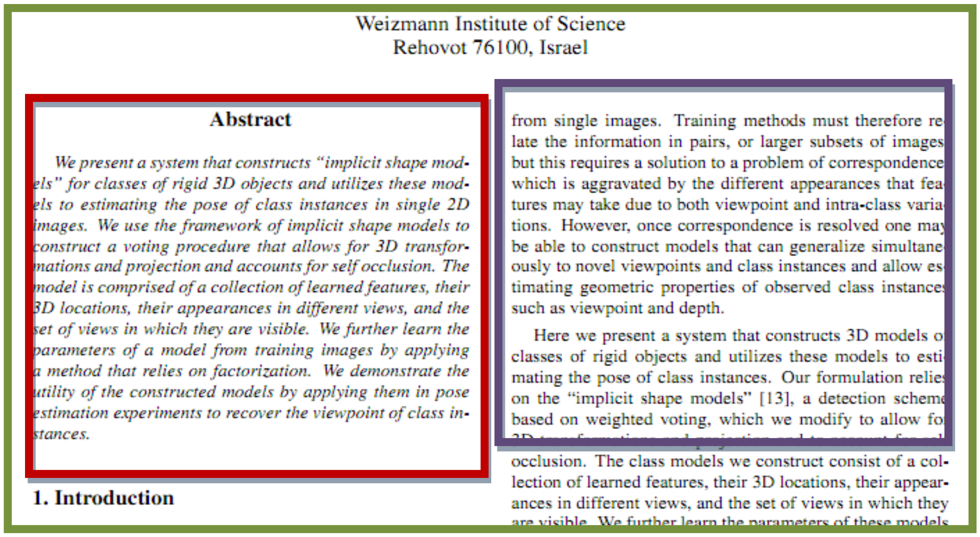


```
An HOG-LBP Human Detector with Partial Occlusion Handling
Xiaoyu Wangâˆ— Tony X. Hanâˆ— Shuicheng Yanâ€  âˆ— â€ Electrical and
Computer Engineering Dept. Electrical and Computer En  neering Dept.
University of Missouri National University of Singapo  Columbia, MO
65211 Singapore 117576
```

It is clearly that the new line character is not parsed after the end of the last author's name. This results that the authors' names are recorded with their titles.

Abstract

The percentage of successful cases (55%) is just fair. The most significant failure the abstract includes some part of content or the figure caption. The reason that makes the parsing of abstract unsuccessful is the malfunction of the parsing program converting the PDF file to plain text (pdftotext). One of the case is shown as follow:



Weizmann Institute of Science
Rehovot 76100, Israel

Abstract

We present a system that constructs "implicit shape models" for classes of rigid 3D objects and utilizes these models to estimating the pose of class instances in single 2D images. We use the framework of implicit shape models to construct a voting procedure that allows for 3D transformations and projection and accounts for self occlusion. The model is comprised of a collection of learned features, their 3D locations, their appearances in different views, and the set of views in which they are visible. We further learn the parameters of a model from training images by applying a method that relies on factorization. We demonstrate the utility of the constructed models by applying them in pose estimation experiments to recover the viewpoint of class instances.

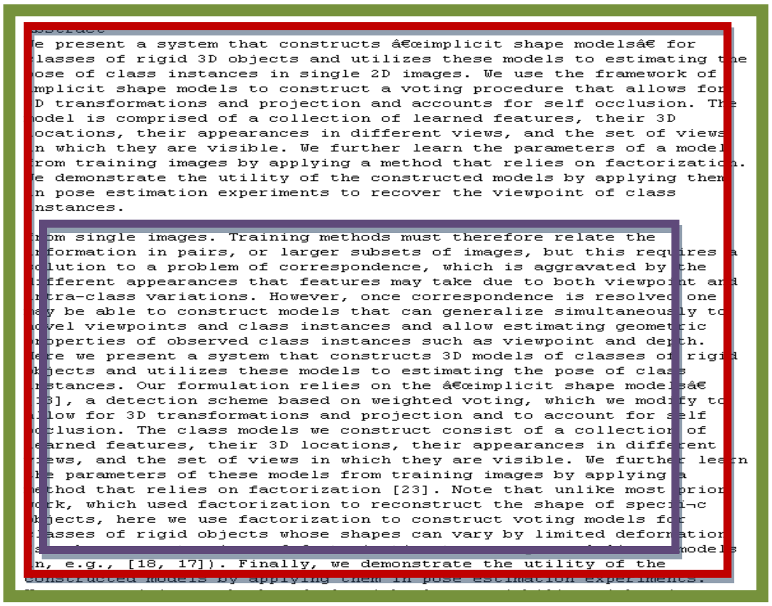
1. Introduction

from single images. Training methods must therefore re late the information in pairs, or larger subsets of images but this requires a solution to a problem of correspondence which is aggravated by the different appearances that fea tures may take due to both viewpoint and intra-class varia tions. However, once correspondence is resolved one may be able to construct models that can generalize simultane ously to novel viewpoints and class instances and allow es timating geometric properties of observed class instances such as viewpoint and depth.

Here we present a system that constructs 3D models o classes of rigid objects and utilizes these models to esti mating the pose of class instances. Our formulation relies on the "implicit shape models" [13], a detection scheme based on weighted voting, which we modify to allow fo

occlusion. The class models we construct consist of a collection of learned features, their 3D locations, their appearances in different views, and the set of views in which they

Let’s see the text parsed:



```
e present a system that constructs â€œimplicit shape modelsâ€ for
lasses of rigid 3D objects and utilizes these models to estimating the
ose of class instances in single 2D images. We use the framework of
mplicit shape models to construct a voting procedure that allows for
D transformations and projection and accounts for self occlusion. The
odel is comprised of a collection of learned features, their 3D
ocations, their appearances in different views, and the set of views
n which they are visible. We further learn the parameters of a model
rom training images by applying a method that relies on factorization.
e demonstrate the utility of the constructed models by applying them
n pose estimation experiments to recover the viewpoint of class
nstances.

rom single images. Training methods must therefore relate the
nformation in pairs, or larger subsets of images, but this requires a
olution to a problem of correspondence, which is aggravated by the
ifferent appearances that features may take due to both viewpoint and
ntra-class variations. However, once correspondence is resolved one
ay be able to construct models that can generalize simultaneously to
ovel viewpoints and class instances and allow estimating geometric
roperties of observed class instances such as viewpoint and depth.
ere we present a system that constructs 3D models of classes of rigid
bjects and utilizes these models to estimating the pose of class
nstances. Our formulation relies on the â€œimplicit shape modelsâ€
13], a detection scheme based on weighted voting, which we modify to
llow for 3D transformations and projection and to account for self
cclusion. The class models we construct consist of a collection of
earned features, their 3D locations, their appearances in different
iews, and the set of views in which they are visible. We further learn
he parameters of these models from training images by applying a
ethod that relies on factorization [23]. Note that unlike most prior
ork, which used factorization to reconstruct the shape of speci¬c
bjects, here we use factorization to construct voting models for
lasses of rigid objects whose shapes can vary by limited deformation
s                                                            models
n, e.g., [18, 17]). Finally, we demonstrate the utility of the
constructed models by applying them in pose estimation experiments.
```

The above case shows that the content of the nest column goes before the “Introduction”. This results in wrong parsing of the abstract.

There are still few problems of parsing. One is that some paper in report-like format that contains table of content would make the server cannot locate the end of the scope of abstract. Another problem is that some abstract are stated without the word “Abstract”.

Keywords

The percentage of successful cases (95%) is almost satisfied. One reason is that not every paper has the keywords column. Even if the paper has keywords part, the keywords are few that avoid the case happened in the abstract – the probability of wrong parsing into fragments is relatively. The failure case is because of the wrong encoding of the PDF file to text file.

References

The percentage of successful cases (35%) is not satisfied. The reason of failure is quite similar to the cases as shown previously. Most of the papers are divided into two columns for better spacing. Yet, the parser sometimes mistakenly parses the right column first and make the order of the text got mess. Some of the reference items is parsed before the term “reference” and hence results in missing items.

There are also some problems for those reference items stated without the [ ] index. Since the “[ ]“ are missing, we must rely on the new line character to break the reference item into pieces. However the new line characters are not always successfully parsed. This makes parse the reference items fail.

7.2. Client Side

7.2.1. Evaluation

For application in the client devices, we do testing on every functions, including login, opening and showing a paper, functions about references, functions about comments, zooming, finding texts, going to certain page and exiting the application.

Login

	Case	Expectation	Result
1	Empty Login ID and Password	Error Message	Same
2	Wrong Login ID	Error Message	Same
3	Wrong Password	Error Message	Same
4	Correct Login ID and Password	Login Succeed	Same

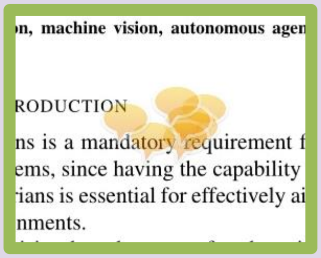
Open and Show a Paper

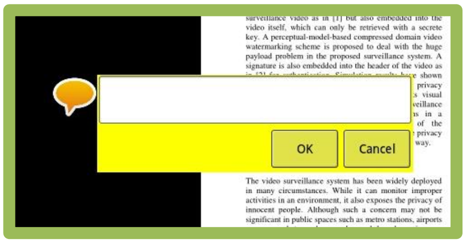
	Case	Expectation	Result
1	Choose one paper	Open the paper	Same
2	Choose another paper while reading one	Close the original paper and open the new one	Same
3	Show every part of the paper when moving it	Show every part normally	The last page cannot be read normally, which suddenly flick to the previous page

Reference

	Case	Expectation	Result
1	Show the reference list after opening a paper	Show the reference list in the transparent panel	Same
2	Show the reference button after opening a paper	Show the reference button beside every references in the paper	Same
3	Show the information of the reference by clicking reference button	Show the information of that reference in an alert box	Same

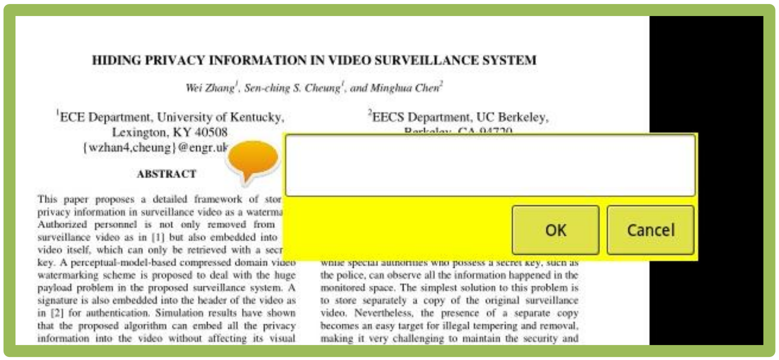
Comment

	Case	Expectation	Result
1	Show the previous comments after opening a paper	Show the buttons of comments at the correct positions	Same
2	Add new comment within the paper	Ask the user to input the comment, and add a new button	Same
3	Add new comment very close to a previous one	Unable to add a new comment, and read the content of the previous one	Add a new comment  on, machine vision, autonomous agen RODUCTION ns is a mandatory requirement f ems, since having the capability rians is essential for effectively ai nments.

4	Add new comment out of the paper	Unable to add a new comment	Add a new comment  surveillance video as in [1] but also embedded into the video itself, which can only be retrieved with a secrete key. A perceptual-model-based compressed domain video watermarking scheme is proposed to deal with the huge payload problem in the proposed surveillance system. A signature is also embedded into the header of the video as e shown privacy s visual veillance s in a of the privacy way. The video surveillance system has been widely deployed in many circumstances. While it can monitor improper activities in an environment, it also exposes the privacy of innocent people. Although such a concern may not be significant in public spaces such as metro stations, airports
5	Show the comment content by clicking the comment button	Show the comment content besides the button without being out of screen	Same
6	Hide the comment content by clicking the comment button	Hide the comment content, and make the button translucent	Same
7	Reply comment	Show the reply at the bottom of the comment	Same
8	Cancel replying a comment	No new reply will be added	Same
9	Delete a comment	Delete that comment and all its replies, and also the button	Same
10	Delete a reply	Delete that reply	Same

Zoom

	Case	Expectation	Result
1	Click the zoom in button	The paper is zoomed in, show all the view with the new scale	The buttons cannot place at the correct positions
2	Click the zoom out button	The paper is zoomed out, show all the view with the new scale	The buttons cannot place at the correct positions
3	Click the zoom fit-in-width button	The paper is zoomed to fit the width of the paper, show all the view with the new scale	The buttons cannot place at the correct positions
4	Double click the screen	The paper is zoomed in, show all the view with the new scale	The buttons cannot place at the correct positions



HIDING PRIVACY INFORMATION IN VIDEO SURVEILLANCE SYSTEM

Wei Zhang[1], Sen-ching S. Cheung[1], and Minghua Chen[2]

[1]ECE Department, University of Kentucky,
Lexington, KY 40508
{wzhan4,cheung}@engr.uk

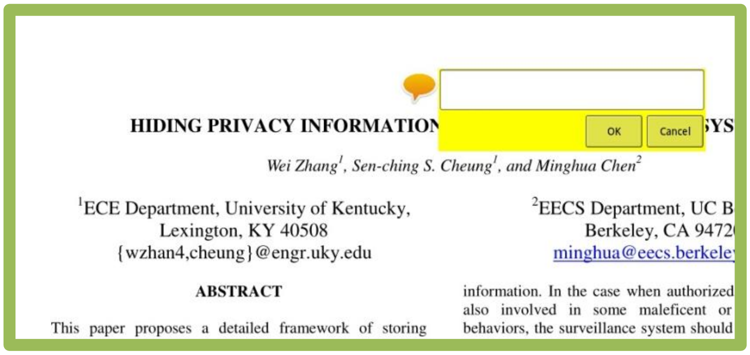
[2]EECS Department, UC Berkeley,

ABSTRACT

This paper proposes a detailed framework of stor
privacy information in surveillance video as a waterma
Authorized personnel is not only removed from
surveillance video as in [1] but also embedded into
video itself, which can only be retrieved with a secr
key. A perceptual-model-based compressed domain video watermarking scheme is proposed to deal with the huge payload problem in the proposed surveillance system. A signature is also embedded into the header of the video as in [2] for authentication. Simulation results have shown that the proposed algorithm can embed all the privacy information into the video without affecting its visual

the police, can observe all the information happened in the monitored space. The simplest solution to this problem is to store separately a copy of the original surveillance video. Nevertheless, the presence of a separate copy becomes an easy target for illegal tempering and removal, making it very challenging to maintain the security and

Figure 7.3 Screen capture of original position of comment



HIDING PRIVACY INFORMATION **SYS**

Wei Zhang[1], Sen-ching S. Cheung[1], and Minghua Chen[2]

[1]ECE Department, University of Kentucky,
Lexington, KY 40508
{wzhan4,cheung}@engr.uky.edu

[2]EECS Department, UC B
Berkeley, CA 9472
minghua@eecs.berkele

ABSTRACT

This paper proposes a detailed framework of storing

information. In the case when authorized
also involved in some maleficent or
behaviors, the surveillance system should

Figure 7.4 Screen capture of position of comment after zooming

Find Text

	Case	Expectation	Result
1	Empty string	Nothing will be found	The application has an error
2	A text does not exist in the paper	Nothing will be found	Same
3	A text existing in the paper	The texts in the paper will be bounded by three red lines	Same
4	Choose previous result	Move the view to the location of the previous result	Same
5	Choose next result	Move the view to the location of the next result	Same
6	Choose hiding all the result	Return to the normal view without showing the result of finding	Same

Go to Certain Page

	Case	Expectation	Result
1	Empty String	Error Message	Same
2	A number which is out of bound	Error Message	Same
3	A number which is within the bounding	Move the view to that certain page	Same
4	Choose going to the first page	Move the view to the first page	Same
5	Choose going to the last page	Move the view to the last page	Same

Exit the Application

	Case	Expectation	Result
1	Click exit button	Ask the user to confirm	Same
2	Confirm exiting	Exiting the application	Same
3	Cancel exiting	Return to the view before	Same

8. Conclusion and Future Work

8.1. Difficulties and Suggested Solutions

We are going to summarize all the difficulties we have encountered during the whole working period in phase 1.
This part will be divided into the server and client side.

8.1.1. Server Side

Difficulty

On summarizing, we have found there are quite a lot problems raised in the parsing functions. There are significant problem should be highlighted.

Firstly, we face great difficulties in parsing the authors' name and information. It is not standardized and extremely various in format, which makes us hard to check where to start to record the author's name and where to end. Thought there are also unstandardized problem in the other parsing fields, they can be addressed by doing a more comprehensive survey and gather the pattern. Yet it is nearly impossible to apply the same method onto the author's part, because it is so varied that enlarging the sampling spare still cannot help.

Secondly, there is problem in the parser which we currently adopted. The performance of Xpdf pdftotext program varies so much that makes the parsing fails frequently. The uncertainty often occurs when the two column formatted paper is parsed. In fact the percentage of fail parsing of the two column paper is quite large.

Thirdly, the format of paper varies. Especially for those scholarly paper is so long that it is written in a report-like format. Such format usually comes with a table of content. The format of those papers is quite different from our original design parsing method. Such paper is not easy to be successfully parsed in our design.

Suggested Solution

This project requires huge amount of data and statistic to support. After all, the paper is written for human beings to read but not the computer. The variation in format is ordinary. To reduce the influence the variation bringing to our system, a larger sample space and more comprehensive research data are required.

Nonetheless, besides improving the research result, some solutions are suggested to make the progress run smoothly.

Firstly, to ease the parsing of the author field, the introduction of PDFmeat to our system should be seriously considered. In fact we have investigated that the PDFmeat is really powerful in acquiring the metadata and is in high precision in parsing the title, authors and year etc., as we mentioned in chapter 3. The reason why we have not implemented the PDFmeat in phase 1 is that we cannot access the CLI interface successfully, which prevent us from using it. After the a long term of trying to parse the data by ourselves, we think that it would be wiser to adopt the PDFmeat. We have to emphases again that our goals to make the management simpler, and high accuracy in data parsing help our goal very much.

Secondly, we would like to seek for other PDF parser to replace the Xpdf which is very unstable. If the plain text parser functions smoothly, most of the parsing problem can be solved with our originally designed implementation methods.

Thirdly, we thought it is not easy to solve the problem of the report-like formatting papers. In fact the report-like formatting paper should be parsed with a brand new implementation method different from what we are writing. We considered that parsing papers in such format should be suspended until most of the functions of the application are fully-developed, since the proportion of report-like formatting paper is relatively minor.

8.1.2. Client Side

Difficulty

It takes time to understand the design of the API since its design is very complicated. It results in being very hard to modify it to fit out project. Besides, the API has some bugs originally, which have to be fixed. For example, in our testing on the application on the client device, we have found that the API has a problem of reading the last page of a PDF file. These bugs are very difficult to be fixed due to its complicated design.

Another difficulty is about the calculation of the position of buttons which we need to add on the paper. Again, due to the complex design, the calculation of the view is not simple. We need to be very clear about how the API works and how it calculates the view of the paper. Some build-in functions like zooming will also affect the calculation, and make it much more intricacy for us to add other features to it, e.g. buttons of references and comments. Because of this problem, our calculation is something wrong that results in erroneous position of buttons after zooming, which we have mentioned in the part of testing on application on the client devices

Suggested Solution

One of our problems is the API cannot work properly when viewing the last page of any papers. Owning to the last page of a paper is quiet important that it shows the references of the paper, we would work on it by finding out the original bugs in the API, and try to fix it.

In addition, we would be much more careful to calculate the views of different elements. By fully understanding the computing logic of the API, it is possible to make all the views in their corresponding positions. However, it needs a large amount of time to completely understand its complex design. We would spend more time on it to realize the complicated design.

Moreover, we have found some bugs during our testing on the application on the client devices. We would try our best to fix them. For the problems of adding comments, we would set some limitation so that the user cannot add comments at some invalid positions. For the case of inputting empty string to find texts, the issue can be solved by add a restriction to it, so that the application would not have a error.

8.2. Conclusion

Our objective of the project is to make things simpler. Tablet devices take rising importance in the age. Although it is impossible for the tablet to fully substitute the traditional laptops, its raising prevalence should be valued, and we believe that tablet device can be introduced to academic aspects.

Yet, It is not easy to develop the application. No matter how much researches we have done to assist the implementation of the whole project, many obstacles still appears during the progress made.

For the server side, we learnt much concerning the format of a formal scholarly paper so as to teach the computer to read the paper. Many parsing problem aroused from the execution. Some of them are solvable, while the others depend on the tools support. We hope that the parsing problem can be addressed by the introducing the PDFmeat and new PDF to plain text parser next semester.

For the client side, most of them are executing problem on the interface. Several solutions are suggested and hoped to alleviate the problems. There are still numerous points can be fine-tuned and we envisage the interface can be further improved.

8.3. Future Work

8.3.1. Sharpen the performance of the parser

It is suggested to bring new tools back to the parser, like PDFmeat and a new “PDFtoText” tool. We will keep making effort on enhancing the performance and accuracy of parsing in order to minimize the workload of maintenance in the future. The more accurate the system is, the more users show their confidence to the whole implementation of the system.

8.3.2. Implement advanced features

We have designed many features to achieve our objectives and some of them have not been implemented, such as the searching functions and the user log functions. Implementing advanced features will therefore be our major concern of phase 2.

8.3.3. Improve the interface

On developing the new functions, we will try to notice whether the existing interface design suits the user’s need or it is really user-friendly. A good user interface can greatly improve the efficiency of the working progress, and the interface is not easily quantified for assessment. Improving the interface should be achieved by gathering more experiences for users.

9. Acknowledgement

We would like to thank the supervisor of our project, Professor Michael Lyu. who gives us valuable advices for our project and remainders to present well. Professor Lyu arranges a meeting with us every week for 1 hour so as to keep our progress up. We are not sure whether the final product can be achieved perfectly, but we feel confident to overcome the obstacles with the help of Professor Lyu.

Besides, we would like to thank Mr.Edward Yan in VIEW Lab. He provides us with facilities and technical support when we are setting up the server and gives us many suggestions on how to make the application more feasible.

10. Reference

[1] KI MAE HEUSSNER and BECKY WORLEY : "Apple iPad: Steve Jobs Unveils the New Apple Tablet ", abcnews http://abcnews.go.com/Technology/apple-ipad-steve-jobs-unveils-apple-tablet/story?id=9667578#.TtDPW7LTpQA

[2] Android application: “gScholarReader” : http://www.androidzoom.com/android_applications/lifestyle/gscholarreader_mdcf.html

[3] Android and iPhone Apps for Selected Research Resources: http://www.lincolnlibraries.org/databases/Android_and_iPhone_Apps_for_Selected_Research_Resources.htm

[4] “Research paper” – WIKIPEDIA : http://en.wikipedia.org/wiki/Research_paper

[5] Behrooz Parhami: "Research Paper Guidelines", University of California, Santa Barbara :http://ece.ucsb.edu/~parhami/rsrch_paper_gdlns.htm

[6] "Writing Research Papers", Rice University: http://www.ruf.rice.edu/~bioslabs/tools/report/reportform.html

[7] "Portable Document Format", WIKIPEDIA : http://en.wikipedia.org/wiki/Portable_Document_Format

[8] Adobe Systems Incorporated: "PDF Referebce" (sixth edition)

[9] "pdfmeat": http://code.google.com/p/pdfmeat/

[10] "PDFlib": http://www.pdflib.com/products/pdflib-family/

[11] "PHP Manual": http://www.php.net/manual/en/function.pdf-set-parameter.php

[12] “pdfbox”: http://pdfbox.apache.org/download.html

[13] “Xpdf”: http://foolabs.com/xpdf/about.html

[14] “SAMSUNG GALAXY Tab 10.1”: http://www.samsung.com/hk/consumer/mobile/mobile-phones/mobile-tablet/GT-P7500FKDTGY/index.idx?pagetype=prd_detail

Appendix : Research on Paper Format

No.	Title	No. of col.	Is title the 1st line	Prsence of					Content after Ref,	Remarks
				author des.	author contact	Abstract	Key.	Ref.		
1	Interactive Light Field Display from a Cluster of Projectors	2	N	Y	N	Y	N	Y	N	
2	The GPU as Numerical Simulation Engine	2	Y(url)	N	N	Y	Keywords	R	N	
3	- WRONG CODING -	1	N	Y	Y	N	N	R	N	CODING
4	Data Fusion Development Concepts within complex Surveillance Systems	2	N	Y	Y	Y	Keywords	Y	N	
5	Development and characterization of a high-spatial-temporal-resolution foot-sole-pressure measurement system	1	N	N	N	Y	N	Y	PAGE NO	REPORT
6	Counting People in Crowds with a Real-Time Network of Simple Image Sensors	2	N	Y	Y	Y	N	Y	PAGE NO	
7	HIDING PRIVACY INFORMATION IN VIDEO SURVEILLANCE SYSTEM	2	N	Y	Y	Y	N	Y	N	
8	A Linear-Time Component-Labeling Algorithm Using Contour Tracing Technique	1	N	Y	Y	Y	Keywords	Y	PAGE NO	
9	Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns	MERGE	Y(Header)	N	N	Y	INDEX TERMS	Y	AUTHOR INFO	
10	Disc In-dash CD Changer Deck	MERGE	N	N	N	Y	N	1)	AUTHOR INFO	Japanese
11	Simple and efficient hash-based verifiable mixing for remote electronic voting Francesc Sebé a,*, Josep M. Miret a, Jordi Pujolàs a, Jordi Puiggalí b	Complex	?	Y	N	Y	Keywords	Y	?	Complex Layout
12	Weighted Sampling for Large-Scale Boosting	1	N	Y	Y	Y	N	Y	N	
13	Online learning of robust object detectors during unstable tracking	MERGE	Y(Header)	Y	Y	Y	N	Y	N	
14	P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints	MERGE	Y(Header)	Y	Y	Y	N	Y	N	
15	FACE-TLD: TRACKING-LEARNING-DETECTION APPLIED TO FACES	MERGE	N	Y	N	Y	INDEX TERMS	Y	REMARKS	

16	Forward-Backward Error: Automatic Detection of Tracking Failures	MERGE	Y(Header)	Y	Y	Y	N	Y	PAGE NO	
17	Robust Recognition of Specific Human Behaviors in Crowded Surveillance Video Sequences	MERGE	Y(Header)	Y	Y	?	N	Y	EXTRA PAGE	Complex Layout
18	Vanishing Point Detection without Any A Priori Information	2	Y(Header)	N	N	Y	INDEX TERMS	Y	REMARKS	
19	E-voting 14 PUBLISHED BY THE IEEE COMPUTER SOCIETY ! 1540-7993/08/$25.00 © 2008 IEEE ! IEEE SECURITY & PRIVACY A Three-Ballot-Based Secure Electronic Voting System	-	-	-	-	-	-	-	-	Article
20	Design and Implementation of Secure Internet Based Voting System with User Anonymity using Identity Based Encryption System	MERGE	Y(Header)	Y	Y	Y	KEY WORDS	Y	PAGE NO	
21	Counting people in crowded environments by fusion of shape and motion information	MERGE	Y(Header)	Y	Y	Y	N	Y	N	
22	FAST HUMAN DETECTION USING MI-SVM AND A CASCADE OF HOG-LBP FEATURES	MERGE	Y(Header)	Y	Y	Y	INDEX TERMS	Y	PAGE NO	
23	View-insensitive Gender Recognition Using Local Binary Patterns	-	-	-	-	-	-	-	-	
24	An HOG-LBP Human Detector with Partial Occlusion Handling	MERGE	N	Y	Y	Y	N	Y	N	
25	A mean-shift tracker: implementations in C++ and Hume	1	N	Y	Y	N	N	Bibliography	MANY PAGES	REPORT
26	A probabilistic framework for real-time head shape detection and tracking	MERGE	PAGE	Y	N	WITH MANY DES	N	Bibliography	N	REPORT
27	A Robust Algorithm for Characterizing Anisotropic Local Structures	1	N	Y	Y	Y	N	1. a	N	
28		-	-	-	-	-	-	-	-	PPT
29	Advanced Data Fusion for Airport Surveillance	MERGE	PAGE	N	N	Y	N	R	N	REPORT
30	Automatic E-Comic Content Adaptation	1	Y(Header)	Y	Y	Y	Keywords	1. a	FOOTER	
31	Background Cut	1	N	Y	Y	Y	N	1. a	N	
32	Homography-based 2D Visual Tracking and Servoing	2	AUTHOR	Y	Y	Y	KEY WORDS	R	FOOTER	Complex Layout
33	Bokode: Imperceptible Visual tags for Camera Based Interaction from a Distance	MERGE	N	Y	Y	Y	Keywords	R	N	

34	Vision-based Pedestrian Detection: will Ants Help?	MERGE	N	N	N	Y	Keywords	Y	N	
35	Computing commonsense	MERGE	N	N	N	DIRECT	N	1	AUTHOR INFO	Complex Layout
36	N/A	-	-	-	-	-	-	-	-	DIAGRAM
37	Shoes Customization Design Tools for the “Diabetic Foot”	1	Y(Header)	Y	Y	Y	Keywords	Y	FOOTER	
38	Real-time Volumetric Reconstruction and Tracking of Hands in a Desktop Environment	1	N	Y	N	Y	KEY WORDS	1. a	DIAGRAM	
39	Camera calibration and 3D reconstruction from a single view based on scene constraints	MERGE	HEADER	Y	Y	Y	Keywords	R	N	
40	Coercion-Resistant Electronic Elections	1	N	Y	Y	Y	KEY WORDS	1. a	MANY PAGES	
41	Dynamic Depth of Field on Live Video Streams: A Stereo Solution	MERGE	HEADER	N	N	Y	Keywords	1. a	N	
42	N/A	N/A	-	-	-	-	-	-	-	CNS1 / UNKNOWN DOC
43	Footwear Fit Categorization	1	N	Y	N	DIRECT	N	Y	AUTHOR INFO	
44	Computational Re-Photography	2	N	Y	N	DIRECT	GENERAL TERMS & KEY WORDS	R	DIAGRAM	
45	N/A	-	-	-	-	-	-	-	-	BOOK
46	N/A	-	-	-	-	-	-	-	-	BOOK
47	Constructing Implicit 3D Shape Models for Pose Estimation	MERGE	N	Y	N	Y	N	Y	N	
48	Footwear Fit Categorization	-	-	-	-	-	-	-	-	CNS1 / UNKNOWN DOC
49	Yet another improvement over the Mu–Varadharajan e-voting protocol	?	PAGE	Y	Y	Y	Keywords	Y	AUTHOR INFO	COMPLEX LAYOUT
50	- WRONG CODING -	1	N	N	Y	Y	Keywords	R	PAGE NO	CODING
51	Histograms of Oriented Gradients for Human Detection	MERGE	N	Y	Y	Y	N	Y	N	
52	N/A	-	-	-	-	-	-	-	-	NO TEXT ELEMENT
53	Data Mining for Surveillance Applications	1	N	Y	Y	Y	N	N	N	
54	ON THE APPLICATION OF A THEORY FOR MOBILE SYSTEMS TO BUSINESS PROCESS MANAGEMENT	1	PAGE	Y	Y	Y	N	Y	N	COMPLEX LAYOUT
55	N/A	-	-	-	-	-	-	-	-	UNKNOWN
56	Robust Real-Time Visual Tracking using Pixel-Wise Posteriors	1	N	Y	Y	Y	N	1. a	N	
57	MapReduce Online	MERGE	PAGE	Y	N	Y	N	Y	PAGE NO	COMPLEX LAYOUT
58	Efficient Indexing for Large Scale Visual Search	MERGE	N	Y	Y	Y	N	Y	N	

59	Efficient particle filtering using RANSAC with application to 3D face tracking	MERGE	Y(Header)	Y	N	Y	Keywords	Y	N	
60	Enhancing the Accessibility for All of Digital Comic Books	1	N	Y	Y	Y	N	R	PAGE NO	
61	Ensemble Tracking	MERGE	Y(Header)	N	N	Y	INDEX TERMS	Y	AUTHOR INFO	
62	Estimating the Number of People in Crowded Scenes by MID Based Foreground Segmentation and Head-shoulder Detection	MERGE	N	Y	Y	Y	N	R	N	
63	Estimation of Number of People in Crowded Scenes Using Perspective Transformation	MERGE	Y(Header)	N	N	Y	INDEX TERMS	R	AUTHOR INFO	
64	FaceSeg: Automatic Face Segmentation for Real-Time Video	MERGE	Y(Header)	N	N	Y	INDEX TERMS	Y	AUTHOR INFO	
65	Fast automatic single-view 3-d reconstruction of urban scenes	MERGE	N	Y	Y	Y	N	1. a	N	
66	Fast automatic single-view 3-d reconstruction of urban scenes	MERGE	N	Y	Y	Y	N	1. a	N	
67	- WRONG CODING -	1	Y(Header)	Y	Y	Y	Keywords	R	PAGE NO	
68	N/A	-	-	-	-	-	-	-	-	BOOK
69	A new fast track-fit algorithm based on broken lines	MERGE	N	Y	N	Y	KEY WORDS	Y	PAGE NO	
70	From Video Sequences to Motion Panoramas	MERGE	N	Y	Y	Y	N	R	FOOTER	
71	- WRONG CODING -	-	-	-	-	-	-	-	-	UNKNOWN
72	View-dependent exploration of massive volumetric models on large-scale light field displays	MERGE	Y(Header)	N	N	Y	Keywords	1. a	FOOTER	
73	Histograms of Oriented Gradients for Human Detection	MERGE	N	Y	Y	Y	N	R	FOOTER	
74	Hive – A Petabyte Scale Data Warehouse Using Hadoop	MERGE	N	Y	N	Y	N	R	N	
75	N/A	-	-	-	-	-	-	-	-	**PPT**
76	Putting Objects in Perspective	MERGE	N	Y	Y	Y	N	(R = M)	N	
77	N/A	-	-	-	-	-	-	-	-	**PPT**
78	N/A	-	-	-	-	-	-	-	-	UNKNOWN
79	Homography Estimation	1	N	Y	N	Y	N	R	PAGE NO	COMPLEX LAYOUT
80	Homography Estimation	MERGE	Y(Header)	N	N	Y	INDEX TERMS	Y	AUTHOR INFO	
81	Human Detection Using Oriented Histograms of Flow and Appearance	1	N	Y	Y	Y	N	Y	N	
82	N/A	-	-	-	-	-	-	-	-	PPT

83	Image interpolation for virtual sports scenarios	MERGE	Y(Header)	N	N	Y	Keywords	1. a	N	AUTHOR COME FIRST
84	- WRONG CODING -	1	Y(Header)	Y	Y	Y	KEY WORDS	Y	EXTRA PAGE	(CODING & COMPLEX
85	Indexing Key Positions between Multiple Videos	MERGE	N	Y	Y	Y	N	Y	REMARKS	
86	Mining High-Dimensional Data for Information Fusion: A Database-Centric Approach	MERGE	N	Y	Y	Y	Keywords	Y	N	