

OPTICAL CHARACTER RECOGNITION

A Project Work Synopsis

Submitted in the partial fulfilment for the award of the degree of
BACHELOR OF ENGINEERING

**IN
ARTIFICIAL INTELLIGENCE
& MACHINE LEARNING**

Submitted by:

**VAIBHAV HARIRAMANI
18BCS6136**

**HRITHIK JOHRI
18BCS6145**

**KSHITIJ MURABATTE
18BCS6146**

**ADITYA GUPTA
18BCS6152**

**SAHIL KUMAR
18BCS6224**

Under the Supervision of:

PROF. SHEFALI SHARMA



**CHANDIGARH
UNIVERSITY**

Discover. Learn. Empower.

**CHANDIGARH UNIVERSITY, GHARUAN,
MOHALI-140413, PUNJAB**

SEPTEMBER 2020

Acknowledgments

I would like to express my deepest gratitude and thanks to my supervisor Prof. Shefali Sharma, who suggested the idea of the Optical Capture Recognition and who gave me the needed information to start working on the project. Also, I would like to thank her for being supportive and for her guidance through this semester and for giving me the necessary advices to be able to realize this project. I am really grateful to her contribution. Moreover, I would like to thank her for her supervising methodology that made my tasks easier and motivated me through this period.

Finally, I would like to thank my all teammates for their support. They have supported me during the difficult moments and encouraged me in carrying out this project work.

Abstract

The aim of this project is to develop an Optical Character Recognition (OCR) for Android based mobile devices. Scanned text documents, pictures stored in mobile phones having Android as operating system, and pictures taken by any Android device are the main focus of this application. The purpose of this application is to recognize text in scanned text documents, text images, and any picture taken by an Android based device in order to reuse it later. This application will allow its users to perform many actions in a few minutes, such as copy text from these aforementioned documents and modify it, instead of wasting time on retyping it from image.

Table of Contents

❖ Acknowledgment

❖ Abstract

1. Introduction

- 1.1 Problem Definition
- 1.2 Purpose
- 1.3 Project Scope
- 1.4 Overall description
- 1.5 Software Specification
- 1.6 Hardware Specification

2. Literature Survey

- 2.1 Existing System
- 2.2 Drawback of existing system
- 2.3 Proposed System
- 2.4 Benefit of Proposed System

3. Problem Formulation

- 3.1 Optical Character Recognition Working
- 3.2 Platform Used

4. Application Design

- 4.1 Architecture Diagram
- 4.2 Gantt Chart

5. Software Components and Technology Used

- 5.1 Technology Enablers
- 5.2 Operating Systems
- 5.2 Hardware Components

- 6. Testing**
- 7. Summary and Implementation**
- 8. Future Work**
- 9. References**

1. INTRODUCTION

1.1 Problem Definition

Optical character recognition

Optical character recognition (OCR) is a process of conversion of typed images, printed text into machine-encoded text, which means it will give us a text from images that contains the text.

It is a cycle of transformation of composed pictures, printed text into the machine-encoded text, which implies it will give us a text from images that contain the text.

Let's explore how *OCR recognizes text*. First, the program analyses the structure of the document picture. It separates the page into components, for example, blocks of texts, tables, pictures, and so on. The lines are isolated into words and then — into characters. When the characters have been singled out, the program contrasts them and a lot of example pictures.

It propels various speculations about what this character is. Based on these theories, the program examines various variations of breaking of lines into words and words into characters. In the wake of handling an enormous number of such probabilistic theories, the program at long last takes the choice, giving you the recognise

1.2 PURPOSE:

The main purpose of Optical Character Recognition (OCR) system based on a grid infrastructure is to perform Document Image Analysis, document processing of electronic document formats converted from paper formats more effectively and efficiently.

This improves the accuracy of recognizing the characters during document processing compared to various existing available character recognition methods. Here OCR technique derives the meaning of the characters, their font properties from their bit-mapped images. The primary objective is to speed up the process of character recognition in document processing. As a result, the system can process a huge number of documents with-in less time and hence saves the time.

Since our character recognition is based on a grid infrastructure, it aims to recognize multiple heterogeneous characters that belong to different universal languages with 3 different font properties and alignments.

1.3 PROJECT SCOPE:

The scope of our product Optical Character Recognition on a grid infrastructure is to provide an efficient and enhanced software tool for the users to perform Document Image Analysis, document processing by reading and recognizing the characters in research, academic, governmental and business organizations that are having large pools of documented, scanned images. Irrespective of the size of documents and the type of characters in documents, the product is recognizing them, searching them and processing them faster according to the needs of the environment.

1.4 OVERALL DESCRIPTION:

Optical Character Recognition, or OCR, is a technology that enables us to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera or phone into editable and searchable data. This technology is very useful since it saves time without the need of retyping the document. It can perform the action in a few minutes. It is able to recognize text in images and convert it into editable text by going throughout a simplified process as illustrated in figure 1.

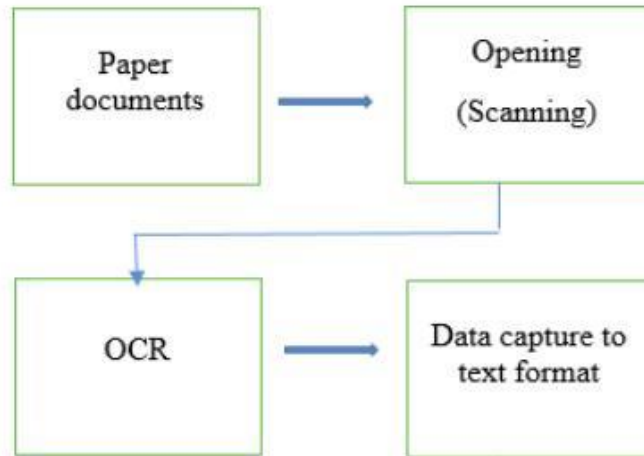


Figure 1: OCR Process

This process generally consists of three stages: Open (Scan) the document, Capture and recognize data and then Save in a convenient format.

1.5 Software Specification

Software Components and Technology used:

The software components and technology used in this project are:

Technology enablers:

1. Dart.
2. Flutter framework.
3. Android SDK (Software development Kit)
4. Android Studio.
5. Operating Systems: Windows, Android OS

1.6 Hardware Specification

Hardware Components:

Personal Computers

Android based mobile phone images with high resolution.

2. LITERATURE REVIEW

2.1 EXISTING SYSTEM:

In the running world there is a growing demand for the users to convert the printed documents into electronic documents for maintaining the security of their data. Hence the basic OCR system was invented to convert the data available on papers into computer process-able documents, so that the documents can be editable and reusable. The existing system/the previous system of OCR on a grid infrastructure is just OCR without grid functionality. That is the existing system deals with the homogeneous character recognition or character recognition of single languages.

2.2 DRAWBACK OF EXISTING SYSTEM:

The drawback in the early OCR systems is that they only have the capability to convert and recognize only the documents of English or a specific language only. That is, the older OCR system is unilingual.

2.3 PROPOSED SYSTEM:

Our proposed system is OCR on a grid infrastructure which is a character recognition system that supports recognition of the characters of multiple languages. This feature is what we call grid infrastructure which eliminates the problem of heterogeneous character recognition and supports multiple functionalities to be performed on the document. The multiple functionalities include editing and searching too whereas the existing system supports only editing of the document. In this context, Grid infrastructure means the infrastructure that supports a group of specific sets of languages. Thus, OCR on a grid infrastructure is multilingual.

2.4 BENEFIT OF PROPOSED SYSTEM:

The benefit of the proposed system that overcomes the drawback of the existing system is that it supports multiple functionalities such as editing and searching. It also adds benefit by providing heterogeneous character recognition.

3. Architecture & Working of OCR

3.1 OPTICAL CHARACTER RECOGNITION WORKING

Optical Character Recognition, or OCR, is a technology that enables you to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera into editable and searchable data.

- Photo OCR (Optical Character Recognition) Problem

- Given picture, detect location of text in the picture
- Read text at that location

- Photo OCR Pipeline

- Text detection
- Character segmentation
- Splitting “ADD” for example
- Character classification



- First character “A”, second “D”, and so on
- When you design a machine learning algorithm, one of the most important steps is defining the pipeline
 - A sequence of steps or components for the algorithms
 - Each step/module can be worked on by different groups to split the workload.

3.2 PLATFORM USED

3.2.1 About Flutter

Flutter is Google's UI toolkit for building beautiful, natively compiled applications form mobile, web, and desktop from a single codebase.

Flutter was released in May 2017. It is written using the C, C++, Dart, Skia Graphics Engine. Flutter works with existing code and ease with which it has been provided to create participants and organizations around the world. Flags also developed by Google. The flutter is the only framework with mobile SDK which provides reactive styles without using any JavaScript bridge. The SDK is free and launched as a source developer to explore and create powerful tracking applications around. It is the reason for apps and interfaces made with flutter. Flutter builds from a single codebase, compile directly to the native arm code, Use the GPU and access the platform APIs and services.

3.2.2 About Dart

Dart is a client-optimized programming language for apps on multiple platforms. It is developed by Google and is used to build mobile, desktop, server, and web applications. Dart is an object-oriented, class-based, garbage-collected language with C-style syntax. Dart can compile to either native code or JavaScript.

4. Application Design:

The design of the application is highly important in order to fulfil the requirements and functionalities of the project.

4.1 Architecture Diagram:

MyApp : This is the application

Material App : The app is a material app

CameraScreen : The main camera window in which the app is launched

Scaffold : The complete screen space (has a Stack and AppBar as child)

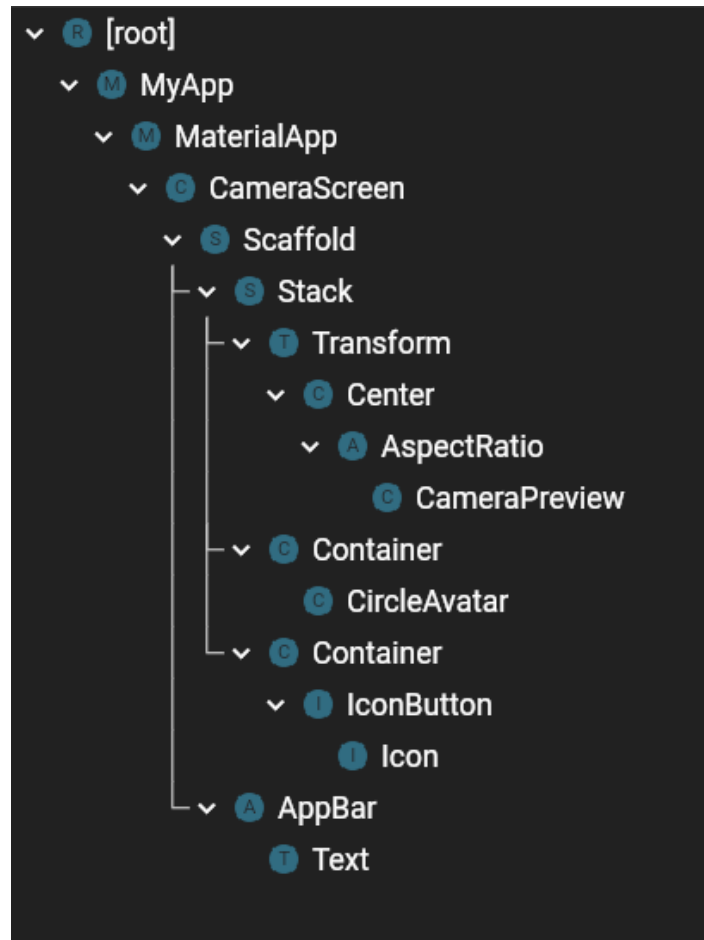
Stack : Used to stack widgets on top of each other

Transform : Display images in appropriate aspect ratio

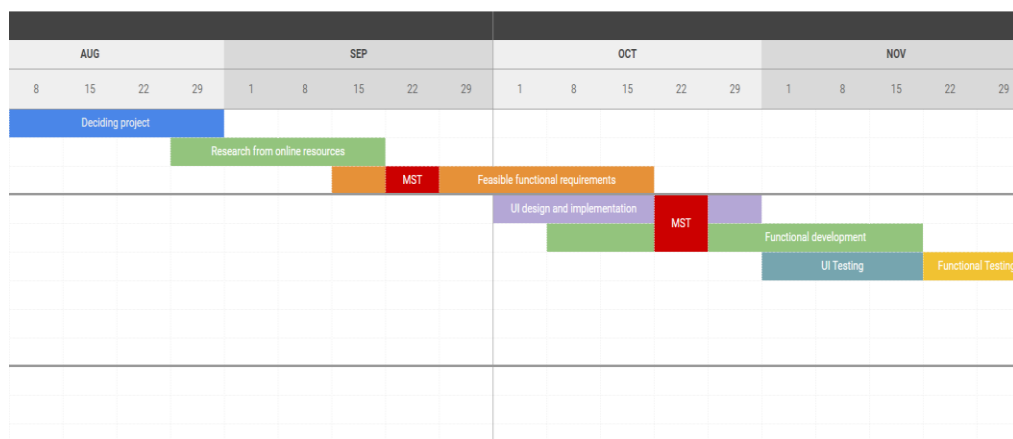
Centre : To centre the camera feed

Container : container for the camera capture button

AppBar : Shows the main appbar with text 'Text Scanner'



4.2 Gantt Chart:



4.1.1 Description of the Attributes:

- **User:** The one that will interact with the Optical Capture Recognition (OCR).
- **The Image:**
 - **Taken:** must have a certain resolution and must contain text.
 - **Selected from the phone:** must contain text.
- **The Text:** must have words.
- **Words:** English Alphabet.

- **Optical Capture Recognition Application.**

4.1.2 Description of the relationships between the modules:

1. **User-OCR Application:** It is one to one relationship because only one user interacts with the application at a time.
2. **OCR Application-Image:** Application is able to process one image at a time.
3. **User-Image:** User can either take a picture or choose one from phone directory.
4. **Image-Text:** The image must contain text.
5. **Text-Words:** The text can contain many words.
6. **Words-Characters:** Each word can have any number of English Alphabet's characters.
7. **User-Text:** User can copy, paste or select the whole text or just a part of it.

5. Software Components and Technology used:

The software components and technology used in this project are:

5.1. Technology enablers:

- Dart as an object-oriented programming language.
- Flutter SDK (Software development Kit)
- Android Studio and VS studio.

5.2. Operating Systems:

- Windows
- Android OS

5.3. Hardware Components:

- Personal Computers
- Android based mobile phone with high resolution.

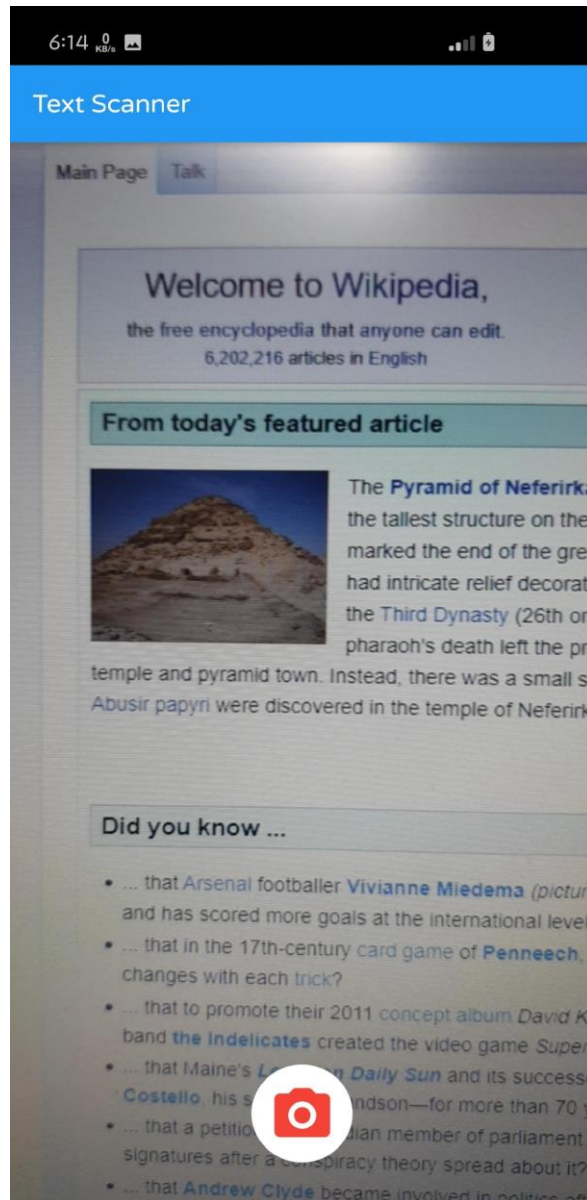
6. Testing:

Testing is an important step that helps to detect errors. Testing is a process of finding faults that might occur during the implementation phase. It is also a way to test if the product fulfils the requirements and to check the components functionalities. There exist many ways of testing where each one of them has a distinct requirement, but the only testing that we made is the acceptance testing.

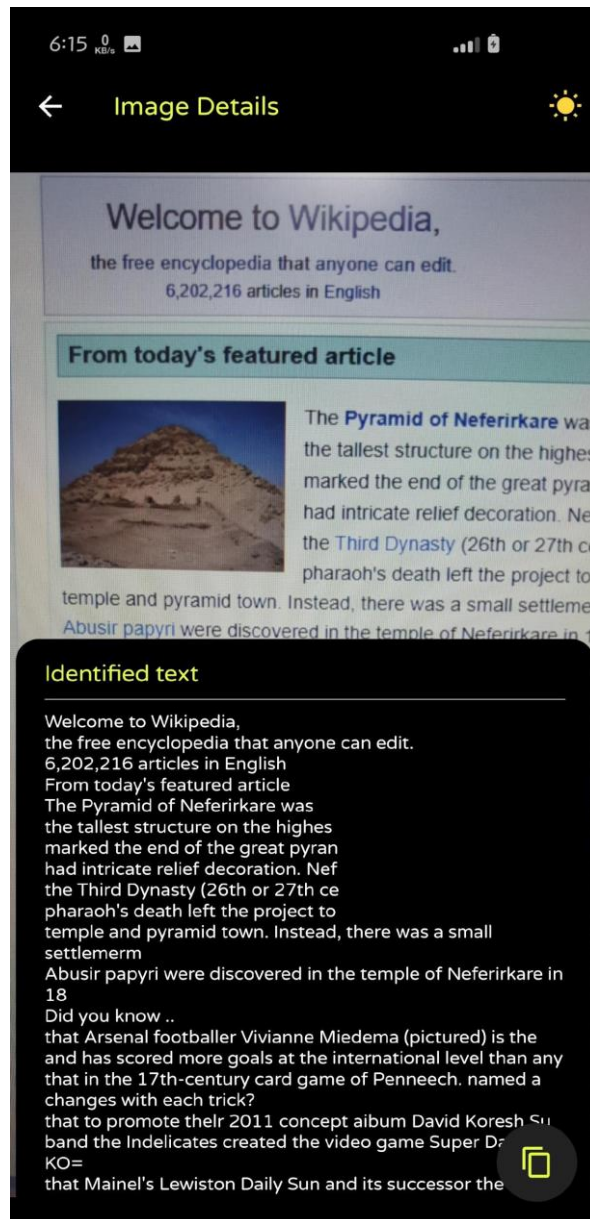
I have tested the application, Android based device, with the supervisor using an acceptance testing strategy.

7. Summary and Implementation results.

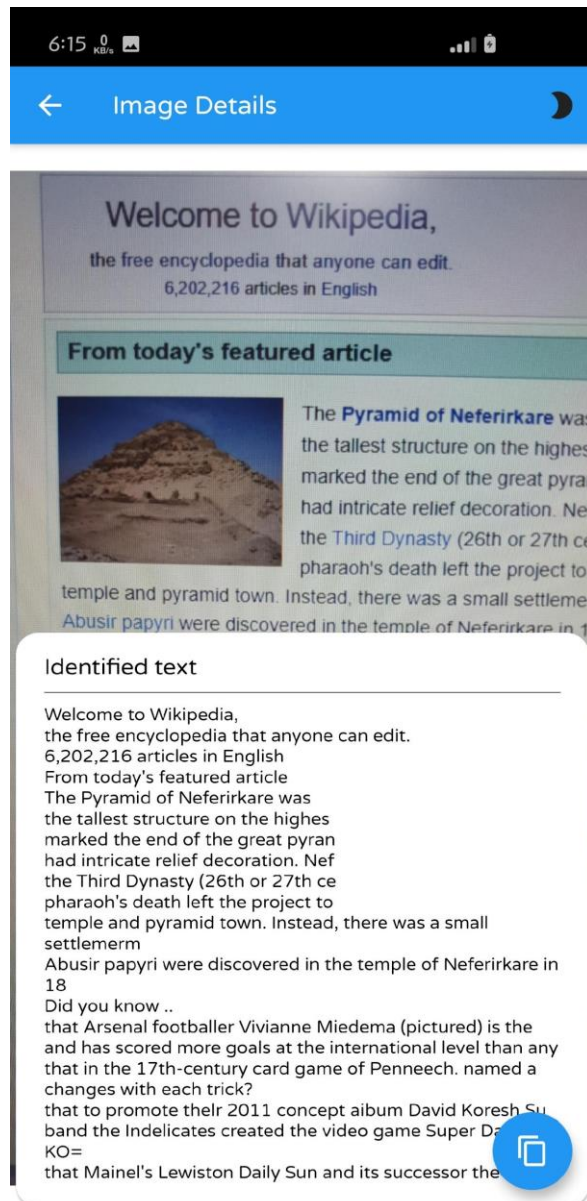
The implementation results show the end result of the project. It is an Android based device application that fulfils the requirements. These are some snapshots of the application



This is the home page of the application that is the finished project. From this page we can change the theme and take pictures of text that we want to be extracted.



This above picture shows the output of the application where the user picks a picture containing a text in English and it shows the text extracted by the application. The theme is Dark and can be changed.



This above picture shows the output of the application where the user picks a picture containing a text in English and it shows the text extracted by the application. This result is in Light theme mode.

8. Future Work:

For the future, we will try to integrate other languages such as Japanese, Chinese, Arabic, or any language that doesn't contain Latin alphabet.

Also, we will concentrate also on the text with skew to ensure that the application will be efficient and employed in further uses. We propose to extend this functionality to enable the accurate prediction of multiple characters simultaneously – thereby enabling truly real time Character Recognition. Also, we shall delve further into the implementation of Neural Networks and come up with methods to increase our accuracy levels. Last, but not the least, we shall develop a GUI which shall enable greater User usability and popularity

9. REFERENCES

- [1] Bhansali, M., & Kumar, P, 2013, An Alternative Method for Facilitating Cheque Clearance Using Smart Phones Application. International Journal of Application or Innovation in Engineering & Management (IJAIEM), 2(1), 211-217
- [2] Shen, H., & Coughlan, J.M, 2012, Towards A Real Time System for Finding and Reading Signs for Visually Impaired Users. Computers Helping People with Special Needs. Linz, Austria: Springer International Publishing.
- [3] Bhavani, S., & Thanushkodi, K, 2010, A Survey On Coding Algorithms In Medical Image Compression. International Journal on Computer Science and Engineering, 2(5), 1429-1434.
- [4] Bhammar, M.B., & Mehta, K.A, 2012, Survey of various image compression techniques. International Journal on Darshan Institute of Engineering Research & Emerging Technologies, 1(1), 85-90.

- [5] Lund, W.B., Kennard, D.J., & Ringger, E.K. (2013). Combining Multiple Thresholding Binarization Values to Improve OCR Output presented in Document Recognition and Retrieval XX Conference 2013, California, USA, 2013. USA: SPIE
- [6] Ciresan, D.C., Meier, U., Gambardella, L.M., & Schmidhuber, J, 2011, Convolutional neural network committees for handwritten character classification presented in International Conference on Document Analysis and Recognition, Beijing, China, 2011. USA: IEEE.
- [7] Mahmoud, S.A., & Al-Badr, B., 1995, Survey and bibliography of Arabic optical text recognition. *Signal processing*, 41(1), 49-77.
- [8] C. K. Roy, M. F. Zibran, and R. Koschke, "The vision of software clone management: Past, present, and future (Keynote paper)," in *2014 Software Evolution Week - IEEE Conference on Software Maintenance, Reengineering, and Reverse Engineering (CSMR-WCRE)*, 2014, pp. 18–33.
- [9] N. Singh, "An Efficient Approach for Handwritten Devanagari Character Recognition based on Artificial Neural Network," 2018 5th Int. Conf. Signal Process. Integr. Networks, SPIN 2018, pp. 894–897, 2018.
- [10] J. Cowell and F. Hussain, "Amharic character recognition using a fast signature based algorithm," *Proc. Int. Conf. Inf. Vis.*, vol. 2003–Janua, pp. 384–389, 2003.