MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
TERNOPIL IVAN PULUJ NATIONAL TECHNICAL UNIVERSITY
FACULTY OF FOREIGN STUDENT
Department of Computer Science

**ISA SADEEQ AKANNI**

UDC 004.10.82

# INFORMATION SYSTEM DEVELOPMENT FOR OPTICAL CHARACTER RECOGNITION USING FOURIER DESCRIPTORS

8.05010101 "Information Control System and Technologies "

master’s degree

The thesis has carried out at the Computer Science Department in Ternopil Ivan Puluj National Technical University, Ministry of Education and Science of Ukraine

**Supervisor:** Ph.D., Senior Lecturer of computer science
**Nazarevych Oleg,**
Ternopil Ivan Puluj National Technical University,

**Reviewer:** Ph.D., assistant professor of Informatic and Mathematical Modelling Department,
**Baran Ihor**
Ternopil Ivan Puluj National Technical University,

Defence will be held in February 26, 2017 at 10.00 am at the meeting of the examination board №32 of Ternopil Ivan Puluj National Technical University at 46001, Ternopil, st. Ruska 56, educational building № 2, Aud. 702

## GENERAL DESCRIPTION OF WORK

**Actuality of thesis:** Thesis aimed to understand, utilize and improve the open source Optical Character Recognizer (OCR) software, OCRopus, to better handle some of the more complex recognition issues such as unique language alphabets and special characters ABSTRACT This paper presents performance evaluation of thresholding algorithms in the context of document analysis and character recognition systems, Several thresholding.

**The purpose of the work:** To better handle some of the more complex recognition issue such as unique image, language alphabets, and special character, to perfom evaluation of thresholding algorithms in the conext of documents analysis and character recognition systems.

**Object, methods and sources of research.**The main object of study is given technological process of images and character recognition. Methods of work: economic and statistical, graphical, comparative, neural network; theoretical and empirical.

**Scientific novelty of the results:**

- The research features of the method of optical character to optimize the layout of image recognition equipment;
- design and analysis service purpose facility production, the analysis of adaptability;
- explored ways of making similar parts;
- completed development process of manufacture of specified parts, for which the selected equipment, tooling, cutting and measuring tools, designed, cutting conditions and standards of time;
- chosen and designed the necessary technological equipment;
- completed a feasibility study of the decisions;
- The question of the use of information technology, safety, safety in emergencies is ecology;
- station designed machine shop for manufacturing case.

**The practical significance of the results.**

Developed in process that can be implemented in a real production. The method of optical recognition can be apply in airport for instance in other to recognice passport.

**Thesis approbation.** The results are multitude of applications where novelty detection is extremely important including signal processing, computer vision, pattern recognition, data mining, and robotics in this paper, this technique is adopted for the optical character quality control in the machine vision inspection application.

**The structure of the work.** The work consists of an explanatory optical and graphical part. Imge explanatory note consists of an introduction, 5 parts, conclusions, list of references and appendixs. Scope of work: settlement and explanatory note – 111 pages. A4, graphic part - 8 sheets A1

## MAIN CONTENTS

**In the introduction** a review of the current state of the industries and describes the main challenges that need to be addressed on Fourier descriptor.

**In the main part** analysis of the issue according to the literature and other sources, the urgency of work done on the formulation of the problem thesis.

**In the research part** The research features of the method utilize and improve the open source optical character recognizer.

**In the technological part** shows the characteristics of the object of production, analysis drawing details and specifications for production, analysis of technological parts, conclusions and formulated the basic tasks of designing, designed single technological process of testing valid documents.

**In the design of** made the choice and design of technological equipment for making the set details

**As a special part** The research capabilities of the package CHARACTER RECOGNITION USING FOURIER, the peculiarities of using computer-aided design technology to solve problems, using appropriate software designed alternative process.

**As part of the project** held designing manufacturing site for the implementation of the developed technological process implemented refine program production at the site, the calculation complexity and Ocropus manufacturing products based on the developed processes, determine annual needs for technological equipment, preparation of summary information equipment, identification of quantitative structure of workers in the softwaer department, determination the size of the main and auxiliary space station and identifying image sizes and a choice of type and design of the output layout plan designed equipment layout.

**As part of the "Substantiation of economic efficiency"** The question of production and calculations conducted technical and economic efficiency of design solutions.

**In terms of** "Health and safety in emergency situations" considered planning work on a labor safety at the site of the projected legal foundations of security in emergencies, and scheme of the protective switch mechanism in the event of tension on cases of equipment or if accidentally touches the conductive parts .

**As part of "Ecology"** analyzes the current status of Ukraine, the issues of pollution arising from the implementation process and proposed measures to reduce environmental pollution.

Annexed to the explanatory note given information sheets, a set of technical documentation for HOST3.1404-86.

In the graphic part of the drawing shows the details of the designation of the coordinate axes and surfaces, drawing blanks, naladok technological schemes, assembly drawings of technological equipment and layout of equipment at the site mehobrobky.

## CONCLUSIONS

A feature-set based on Fourier Descriptors was developed. The features employed a new method for handling images which contain multiple curves. In addition, a set

of software components were developed which allow these features to be extracted

from an image and used for classification. The feature extraction software was made available as an extension to the Gamera open-source document processing framework. Two classier were implemented using the feature-set, one based on a fuzzy-knn and the other based on a neural-network. In addition, software was developed for training and testing the classier in parallel on a beowulf cluster. The software was used to evaluate the classier performance on a data set containing newsprint. The fuzzy-knn classier was tested using a range of parameter values in order to determine their optimal values. The neural-network classier was made available through the python api in the Gamera system.

**List of papers published by the author of thesis**

1. Isa Sadeeq Akanni, B.B.Mlynko Ph.D., Assoc.Prof. Information system development for optical character recognition using fourier descriptors / Proc. Of V International scientific conference of young scientists and students <<Actual tasks of modern technologies>>, Ternopil, November 17-18, 2016, vol.2. – Ternopil, TNTU Press. 9 p.

## SUMMARY

Optical character recognition has many different practical applications. The main areas where OCR has been of importance, are text entry (office automation), data entry (banking environment) and process automation (mail sorting).

**Keywords**: Technology, devices, machines, learning parameters, Momentum, Optical character recognition, Scripts, digital image processing.