

Рецензия

върху дисертационен труд за получаване на образователната и научна степен “доктор” по научна специалност: **"Автоматизирани системи за обработка на информация и управление"**

в професионално направление: **5.3 Комуникационна и компютърна техника**

Автор на дисертационния труд: **Салма Бумиза**

Тема на дисертационния труд: **Разработване на модели за аудио-визуален контрол и мониторинг на реакцията на студента чрез електронно обучение**

Изготвил: **доц. д-р инж. Агата Христова Манолова**, Факултет по Телекомуникации, Технически Университет-София

1. Актуалност на разработвания проблем

Съвременните комуникационни и информационни технологии осигуряват до потребителя предоставят изключителни възможности за обмен и усвояване на знания, опит, повишаване и придобиване на нова квалификация. Те позволяват провеждането на високоефективен учебен процес във виртуални среди, наречен електронно обучение. Проблемите на електронното обучение, свързани с автоматизирането на процеса на обучение без да се налага постоянното наблюдение от преподавателя, намира отражение в дисертационния труд на маг. Салма Бумиза. Научната област на представения за рецензия дисертационен труд, без съмнение, е изключително актуална, особено в условията на Covid-19 пандемията и е свързана с изследването и разработването на модели за аудио-визуален контрол и наблюдение на реакцията на ученика чрез електронно обучение.

Темата на дисертационния труд пряко касае разработването на модели за аудио-визуален контрол и наблюдение на реакцията на ученика чрез електронно обучение.

2. Познаване на състоянието на проблема и творческа интерпретация на литературния материал

Дисертационния труд показва, че докторантът познава добре актуалните системи за обучение в отдалечена среда и някои съпътстващи софтуерни приложения. Литературният обзор обхваща 50 източници, книги, научни публикации, интернет сайтове от 2003 до 2019 г..

След направен обстоен анализ на най-популярните методи за разпознаване на лица на Viola-Jones; за проследяване на лица на Kanade, Lucas и Tomasi (KLT) и за разпознаване на емоции чрез Principal Component Analysis (PCA); методите за разпознаване на глас, основани на Mel Frequency Cepstral Coefficient (MFCC) и Dynamic Time Warping (DTW), и за класификация и вземане на решение с помощта на алгоритъма Support Vector Machine (SVM), в края на Глава 1, маг. Бумиза обосновава формулирането на **основните задачи** на дисертационния труд:

- **Разработване на теоретични модели за анализ на аудио и визуални реакции на учащи за интерактивен контрол в електронното обучение;**
- **Разработване на алгоритми на базата на математическите модели, предложени за аудиовизуален контрол и мониторинг на реакцията на учащия чрез електронно обучение;**

- **Експериментални изследвания за разработване на алгоритми за аудиовизуален контрол и мониторинг на реакциите на учениците в средата за електронно обучение.**

Решаването на така представените задачи на дисертационния труд е представено в следващите 4 глави.

3. Съответствие на избраната методика на изследване с поставената цел и задачи на дисертационния труд

В глава 1 на дисертационния труд са дефинирани 3 основни задачи и получените експериментални резултати предполагат, че в основната си част така формулираните задачи са изпълнени. Разработени и представени в дисертационния труд са два модела за анализ на аудио и визуалните реакции на обучаемите с помощта на системата за разпознаване на емоции, интегрирана в платформата за електронно обучение Moodle. От представените резултати и научни публикации следва, че разработените модели за изследване, основана на методите за лицево и гласово разпознаване с помощта на алгоритмите Viola-Jones, KLT, DTW, PCA, MFCC и SVM, съответстват на поставените задачи на дисертационния труд.

4. Кратка аналитична характеристика на естеството и оценка на достоверността на материала, върху който се градят приносите на дисертационния труд

Трудът съдържа 126 страници с фигури и таблици и 6 приложения. Той включва общо 4 глави. В края на работата са формулирани приносите и перспективите за бъдеща работа, и са приложени списъци на използваната литература, на публикациите по дисертацията, с участие на докторанта. Не са предстани използвани съкращения и означения, но им отделно описание на фигурите.

В края на Глава 1 въз основа на направения обзор, авторът си поставя няколко цели и задачи за изпълнение, на които се дава отговор в следващите глави.

В Глава 2 се разработват два модела за анализ на аудио и визуални реакции на учащи, базирани на разпознаване на емоции. Първият модел се основава на интелигентна мрежа, която използва емоциите на обучаващия се като входни неврони, филтрира тези емоции и взема решение за нивото на усвояване на материала от страна на студента. В съответствие с втория модел, системата за разпознаване на емоции е интегрирана в модула на от страна на преподавателя, който може да проследи реакциите на всеки студент. Предложена е и възможност да се автоматизира ролята на преподавателя: „WorkflowE-learningProcessAction“ за електронно обучение с използване на аудио-визуално разпознаване на реакциите на студентите.

В глава 3 се разработват алгоритми на базата на математически модели, предложени за аудиовизуален контрол и мониторинг на реакцията на учащия чрез електронно обучение.

В глава 4 се представя експериментално изследване на разработваните алгоритми за аудиовизуален контрол и мониторинг на реакциите на учениците чрез електронно обучение и заключение върху представените резултати. Представените алгоритми са разработени в среда Matlab, като не са направени цялостна оценката на алгоритмите и сравнение на предложената система с други вече съществуващи.

В приложенията са дадени кодовете на алгоритмите на Matlab за детектиране и проследяване на лица от видеоклипове и на алгоритмите MFCC, DTW, SVM за разпознаване на гласови емоции; използваните тестови изображения с основни 6 лицеви емоции; кратко описание на платформата Moodle и кода на Matlab програмата за интегриране в Moodle.

5. Научни и/или научноприложни приноси на дисертационния труд

Дисертацията има приложени характер и принос в посока разработката на автоматизирана система за електронно обучение.

Основните приноси на дисертацията имат основно научноприложен и приложен характер и могат да се обобщят по следния начин:

а) Научноприложните приноси

- разработени са два нови модела за анализ на аудио и визуалните данни на студенти с помощта на система за разпознаване на емоции, интегрирана в платформата Moodle, съгласно фиг. 2.2, 2.3 и 2.4;
- предложен е модел на автоматичен преподавател (WEPA) в среда за електронно обучение чрез аудиовизуално разпознаване на реакциите на студентите на фиг. 2.7 и 2.8;
- разработени са 4 алгоритъма, основани на двата модела за аудиовизуален контрол и проследяване на поведението на студентите в процеса на електронно обучение (на фиг. 3.1, 3.2, 3.3 и 3.4) и 2 алгоритъма: за вземане на решение в зависимост от разпознатата емоция (фиг. 3.7) и за избор на студент от студентската база данни (фиг. 3.8).

б) Приложни приноси:

- Разработени са Matlab програми за визуално и гласово разпознаване на емоции (на фиг. 4.1 и 4.2) със съответен графичен интерфейс (на фиг. 4.3 и 4.4);
- Интегрирана е системата за разпознаване на емоции в платформите за онлайн обучение Moodle и за видеоконференции Cisco Webex (на фиг. 4.5 и 4.6).

6. Оценка за степента на личното участие на дисертанта в приносите

Предоставеният ми за рецензия дисертационен труд безспорно съдържа научно-приложни и приложни приноси, като те са представени на научни форуми в България и чужбина и са получили необходимата публичност и са били рецензирани от независими рецензенти. По дисертацията са направени общо 7 публикации в съавторство, от които 1 на френски език, останалите на английски език. От тези публикации, 5 са индексирани в базата Scopus, като след справка в нея, някои от тях вече са генерирали и множество цитати, което доказва приложимостта на дисертационния труд и оценката на международната научна общост за приноса на дисертационния труд и работата на докторанта.

7. Преценка на публикациите по дисертационния труд

По дисертацията няма направени самостоятелни публикации. Направените 7 публикации по дисертационния труд са в периода 2015 - 2020 г.: 2 статии в списания "International Journal of Computer Engineering and Technology" и "International Journal of Advance Research in Computer Science and Management Studies"; 5 статии са в сборници доклади, представени на международни научни конференции. 5 от тези публикации са индексирани в базата Scopus. Посочените трудове по брой и място на публикуване покриват изцяло изискванията за придобиване на ОНС „Доктор”.

8. Използване на резултатите от дисертацията в научната и социалната практика

Няма

9. Мнения, препоръки и бележки

Дисертантът е достигнал до интересни научно-приложни резултати. Очевидно е голямото поле за тяхното практическо приложение във всекидневния живот.

Една от препоръките към маг. Бумиза е да представи разработката си за тестване и използване за нуждите на избран от нея факултет и да валидира резултатите от дисертацията в реални условия.

На предварителното обсъждане на труда на заседанието на докторантското училище представих рецензия с въпроси и забележки, част от които в окончателния текст на дисертацията са коригирани. Затова по отношение на представената финална версия на дисертацията имам следните бележки и въпроси:

- а) Как представените модели с интеграция към Moodle дават възможност в реално време да се детектират, отделят и идентифицират лица и разпознават експресии на един или няколко учащи?
- б) В глава 2, точка II.3.2. Моделът за разпознаване на гласови емоции (*Le modèle de reconnaissance des émotions vocales*) е обявено, че ще се използват коефициентите на Мел (*Mel Frequency Cepstral Coefficient*) за разпознаване на емоции от речеви сигнали. В [М. E. Ayadi, M. S. Kamel, and F. Karray, "Survey on speech emotion recognition: Features, classification schemes, and databases", *Pattern Recognition*, vol. 44, issue 3, pp. 572-587, 2011.], авторите доказват, че правилният избор на признаци за разпознаване на емоция в речта силно зависи от естеството на конкретната класификационна задача. Например, ако разгледаме класификацията на афективни състояния с високи нива на активиране спрямо такива с ниско ниво, обикновено непрекъснати функции, базирани на тон или височина на гласа се използват за описание. Смята се, че по-сложните задачи изискват подход, надграден върху комбинация от различни видове функции. На какво основание се предлага да се използват само Мел коефициентите? Защо не се използват кепстрални коефициенти, спектрални честоти, енергия и др.
- в) В дисертационния труд не се представят количествени резултати от разпознаването на експресии и мултимодалната комбинацията от глас и лице като напр. *confusion matrix*, *precision*, *recall*, *F-measure*, и др.

Направените забележки не омаловажават значимостта на резултатите.

10. Заключение

Дисертацията на Салма Бумиза се отличава с редица достойнства в експериментално-приложен план. Като цяло дисертационният труд издава добра компетентност по поставената тематика и след направените забележки и коментати ще отговаря на изискванията на ЗРАСРБ и на Правилника на ТУ-София за прилагането му. Въз основа на горепосочените достойнства и приноси на дисертационния труд, считам, че са достатъчно основание за присъждане на ОНС „Доктор“ по научната специалност „Автоматизирани системи за обработка на информация и управление“ на маг. инж. Салма Бумиза. Затова предлагам на уважаемите членове на Научното жури да се присъединят към моята положителна оценка и да гласуват с „ДА“.

27.08.2021 г.

Подпис:

/доц. д-р Агата Манолова/

Review

on the dissertation work for obtaining the educational and scientific degree „Doctor”

Author of the thesis: MSc. Eng. Salma Boumiza

Title of the thesis: DEVELOPMENT OF MODELS FOR AUDIO-VISUAL CONTROL AND MONITORING OF THE STUDENT'S REACTION THROUGH E-LEARNING

Reviewer: assoc. prof. **Agata Hristova Manolova**, PhD, Faculty of Telecommunications, Technical University of Sofia

1. Actuality of the problem investigated in the PhD thesis and tasks, developed in the dissertation

Modern communication and information technologies provide the user with exceptional opportunities for exchange and acquisition of knowledge, experience, improvement and acquisition of new qualifications. They allow for a highly effective learning process in virtual environments called e-learning. The problems of e-learning, related to the automation of the learning process without the need for constant monitoring by the teacher, is reflected in the dissertation of Salma Bumiza. The scientific field of the dissertation submitted for this review is undoubtedly extremely relevant, especially in the context of the Covid-19 pandemic and is related to the research and development of models for audio-visual control and monitoring of student response through e-learning.

The topic of the dissertation directly concerns the development of models for audio-visual control and monitoring of the student's reaction through e-learning.

2. Degree of knowledge of the problem

The dissertation shows that the doctoral student is well acquainted with current e-learning systems and some related software applications. The literature review covers 50 sources, books, scientific publications, websites from 2003 to 2019.

After a thorough analysis of the most popular face recognition method Viola-Jones; face tracking of Kanade, Lucas and Tomasi (KLT) and expression recognition through Principal Component Analysis (PCA); voice recognition methods based on Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW), and for classification and decision making based on the Support Vector Machine (SVM) algorithm, at the end of Chapter 1, miss Bumiza justifies the formulation of the main tasks of the dissertation:

- Development of theoretical models for analysis of audio and visual reactions of students for interactive control in e-learning environment;
- Development of algorithms based on mathematical models proposed for audiovisual control and monitoring of the learner's reaction through the e-learning process;
- Experimental research of the developed algorithms for audiovisual control and monitoring of students' reactions in the e-learning environment.

The solutions of the presented tasks of the dissertation are described in the next 4 chapters.

3. Correspondence of the chosen research methodology and the set objective and tasks of the thesis with the achieved contributions

Chapter 1 of the dissertation defines 3 main tasks and the obtained experimental results suggest that in general terms the tasks formulated in this way are fulfilled. Developed and presented in the dissertation are two models for analysis of audio and visual reactions of

students using the emotion recognition system integrated into the e-learning platform Moodle. From the presented results and scientific publications it follows that the developed research models based on the methods of facial and voice recognition using the algorithms Viola-Jones, KLT, DTW, PCA, MFCC and SVM, correspond to the tasks of the dissertation.

4. Short analytical characteristics of the thesis

The dissertation contains 126 pages with figures and tables and 6 appendices. It includes a total of 4 chapters. At the end of the dissertation, the contributions and perspectives for future work are formulated, and lists of the used literature, of the dissertation's publications, with the participation of the doctoral student are attached. Abbreviations and notations used are not presented.

At the end of **Chapter 1**, based on the review, the author sets several goals and objectives for implementation, which are answered in the following chapters.

Chapter 2 two models for analyzing audio and visual reactions of learners based on emotion recognition are presented. The first model is based on an intelligent network that uses the learner's emotions as input neurons, filters these emotions and decides on the student's level of content understanding. According to the second model, the emotion recognition system is integrated in the module by the teacher, who can track the reactions of each student. It is also possible to automate the role of the teacher: "WorkflowE-learningProcessAction" for e-learning using audio-visual recognition of students' reactions.

Chapter 3 algorithms based on mathematical models proposed for audiovisual control and monitoring of the learner's reaction through e-learning are described.

Chapter 4 presents an experimental study of the developed algorithms for audiovisual control and monitoring of students' reactions through e-learning and a conclusion on the experimental results. The presented algorithms are developed in Matlab environment, but a complete evaluation of the algorithms and comparison of the proposed system with other existing ones have not been made.

The appendices provide the codes of Matlab's algorithms for detecting and tracking faces from videos and of the MFCC, DTW, SVM algorithms for recognizing voice emotions; used test images with basic 6 facial emotions; a brief description of the Moodle platform and the code of the Matlab program for integration into Moodle.

5. Scientific and applied contributions of the thesis

The dissertation has an applied character and contribution towards the development of an automated e-learning system.

The main contributions of the dissertation are mainly scientifically applied and applied in nature and can be summarized as follows:

a) Scientifically applied contributions

- Two new models for analysis of audio and visual data of students have been developed with the help of an emotion recognition system integrated in the Moodle platform, according to fig. 2.2, 2.3 and 2.4;

- a model of an automatic teacher (WEPA) in an e-learning environment through audiovisual recognition of the reactions of the students in fig. 2.7 and 2.8 is described;

- 4 algorithms have been developed, based on the two models for audiovisual control and tracking of students' behavior in the e-learning process (in Fig. 3.1, 3.2, 3.3 and 3.4) and

2 algorithms: for decision making depending on the recognized emotion (Fig. 3.7) and for selecting a student from the student database (Fig. 3.8).

b) Applied contributions:

- Matlab programs for visual and voice recognition of emotions have been developed (in Fig. 4.1 and 4.2) with a corresponding graphical interface (in Fig. 4.3 and 4.4);
- The system for recognizing emotions in the online learning platforms Moodle and for video conferencing Cisco Webex is integrated (in Fig. 4.5 and 4.6).

6. Evaluation of the personal participation of the applicant in the contributions

The dissertation submitted to me for review undoubtedly contains scientific-applied and applied contributions, as they were presented at scientific forums in Bulgaria and abroad and received the necessary publicity and were reviewed by independent reviewers. A total of 7 co-authored publications were made on the dissertation, of which 1 in French language, the rest in English. Of these publications, 5 are indexed in the Scopus database, and after reference in it, some of them have already generated numerous citations, which proves the applicability of the dissertation and the assessment of the international scientific community for the contribution of the dissertation and doctoral work.

7. Evaluation of the publications on the dissertation

No independent publications have been made on the dissertation. The 7 publications on the dissertation were made in the period 2015 - 2020: 2 articles in the journals “International Journal of Computer Engineering and Technology” and “International Journal of Advance Research in Computer Science and Management Studies”; 5 articles are in collections of papers presented at international scientific conferences. 5 of these publications are indexed in the Scopus database. The indicated works by number and place of publication fully cover the requirements for acquiring title educational and scientific degree „Doctor”.

8. Use of the achieved results of the thesis in the scientific and social practice

None

9. Notes and suggestions

The dissertation has reached interesting scientific and applied results. Obviously, there is a large field for their practical application in everyday life.

One of the recommendations to Miss Bumiza is to present her developed system for testing and use for the needs of a faculty of her choice and to validate the results of the dissertation in real conditions.

At the preliminary discussion of the work at the meeting of the doctoral school I’ve presented a review with questions and comments, some of which in the final text of the dissertation are corrected. Therefore, regarding the presented final version of the dissertation I have the following notes and questions:

a) How do the presented models with Moodle integration allow real-time detection, tracking and identification of faces and recognition of expressions of one or more learners?

b) In Chapter 2, point II.3.2. The Voice Emotion Recognition Model (Le modèle de reconnaissance des émotions vocales) is announced that Mel Frequency Cepstral Coefficient will be used to recognize emotions from speech signals. In [M. E. Ayadi, M. S. Kamel, and F. Karray, “Survey on speech emotion recognition: Features, classification schemes, and databases”, Pattern Recognition, vol. 44, issue 3, pp. 572-587, 2011.], the authors prove that the correct choice of features for recognizing emotion in speech strongly depends on the nature of the specific classification task. For example, if we consider the classification of affective states with high levels of activation versus those with low levels, usually continuous

functions based on tone or pitch are used for features. It is believed that more complex tasks require an approach built on a combination of different types of functions. On what grounds is it proposed to use only Mel coefficients? Why not use cepstral coefficients, spectral frequencies, energy, etc.

c) The dissertation does not present quantitative results from the recognition of expressions and the multimodal combination of voice and face such as. confusion matrix, precision, recall, F-measure, etc.

The remarks made do not diminish the significance of the results.

10. Conclusion

Salma Bumiza's dissertation has a number of merits in the experimental application. In general, the dissertation issues good competence on the set topic and after the remarks and comments will meet the requirements of the Law on the Development of the Academic Staff in the Republic of Bulgaria and the Regulations of TU-Sofia for its implementation. Based on the above-mentioned merits and contributions of the dissertation, I consider that they are a good reason for the assigning the educational and scientific degree „Doctor” to M. Sc. Eng. Salma Boumiza. This is why, I propose the reputable members of the Scientific Jury to join my positive evaluation and to vote „YES“.

27.08.2021
Sofia

Member of scientific jury:
/assoc. prof. Agata Manolova, PhD/