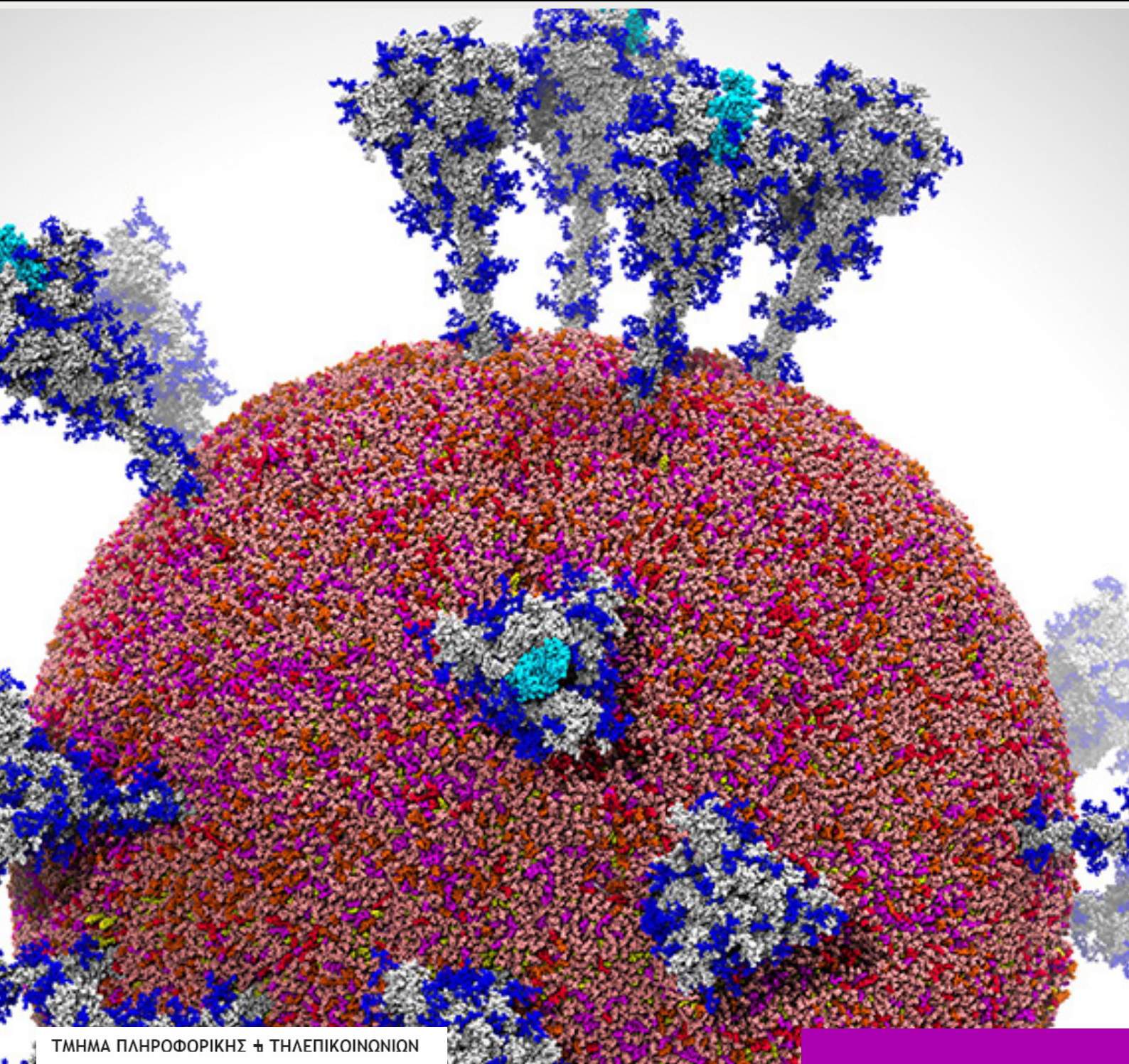


STUDBOOK

ΕΠΙΛΕΓΜΕΝΕΣ ΠΤΥΧΙΑΚΕΣ ΚΑΙ ΔΙΠΛΩΜΑΤΙΚΕΣ ΕΡΓΑΣΙΕΣ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ + ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών
— ΙΔΡΥΘΕΝ ΤΟ 1837 —



ΤΟΜΟΣ 18
2021

STUDENT BOOK

Εκδίδεται μία φορά το χρόνο από το:

**Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών,
Πανεπιστημιούπολη, 15784 Αθήνα**

Επιμέλεια έκδοσης:

Επιτροπή Ερευνητικών και Αναπτυξιακών Δραστηριοτήτων

Θ. Θεοχάρης (υπεύθυνος έκδοσης), Καθηγητής, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Η. Μανωλάκος, Καθηγητής, Τμήμα Πληροφορικής και Τηλεπικοινωνιών

Γραφιστική επιμέλεια - Επιμέλεια κειμένων:

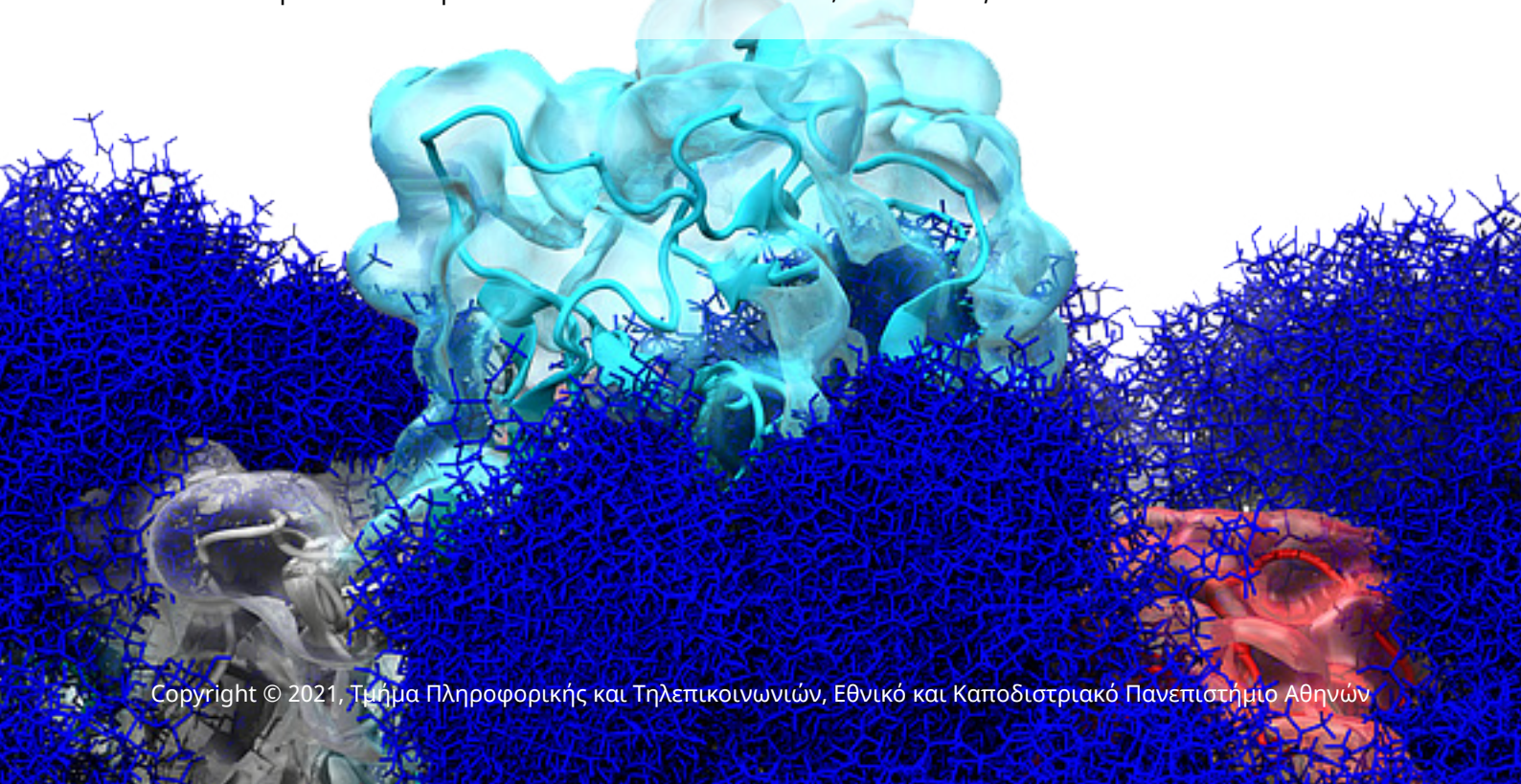
Λ. Χαλάτση, ΕΤΕΠ, Τμήμα Πληροφορικής και Τηλεπικοινωνιών

ISSN 1792-8826

Εικόνα εξωφύλλου: A snapshot of a visualization of the SARS-CoV-2 viral envelope comprising 305 million atoms. Image Credit: Rommie Amaro, University of California San Diego; Arvind Ramanathan, Argonne National Laboratory.

A team led by Rommie Amaro, professor and endowed chair of chemistry and biochemistry at the University of California San Diego, and Arvind Ramanathan, computational biologist at Argonne National Laboratory, has been exploring the movement of the virus's spike protein to understand how it behaves and gains access to the human cell. Now, the team has built a first-of-its-kind workflow based on artificial intelligence (AI) and has run it on the Oak Ridge Leadership Computing Facility's (OLCF's) 200-petaflop IBM AC922 Summit supercomputer to simulate the spike in numerous environments, including within the SARS-CoV-2 viral envelope comprising 305 million atoms—the most comprehensive simulation of the virus performed to date.

Εικόνα εσωφύλλου: Glycans (dark blue) coat the SARS-CoV-2 spike protein (light blue), which is anchored in the viral envelope. Credit: Adapted from ACS Central Science 2020, DOI: 10.1021/acscentsci.0c01056



Περιεχόμενα

ΠΡΟΛΟΓΟΣ	4
ASSESSING THE URXD 3D FACE RECOGNITION ALGORITHM ON SYNTHETIC FACIAL DATA, PANTELEIMON D. KANELIS	6
ΕΙΣΑΓΩΓΗ ΣΤΗ ΘΕΩΡΙΑ ΜΗΤΡΟΕΙΔΩΝ, ΜΕΡΚΟΥΡΗΣ ΠΑΠΑΜΙΧΑΗΛ	18
ΣΧΗΜΑ ΟΔΗΓΟΣ-ΑΚΟΛΟΥΘΟΣ ΚΙΝΟΥΜΕΝΩΝ ΡΟΜΠΟΤ ΜΕ ΤΗ ΧΡΗΣΗ RGB-D ΚΑΜΕΡΑΣ ΚΑΙ MOBILENETS, ΠΑΝΑΓΙΩΤΗΣ Α. ΠΕΤΡΟΠΟΥΛΑΚΗΣ	33
A MIXED REALITY DASHBOARD BASED ON A DISTRIBUTED DATA STREAMING ARCHITECTURE, DELIGIANNAKIS N. NEKTARIOS	52
ΑΞΙΟΠΟΙΗΣΗ ΤΗΣ ΠΛΑΤΦΟΡΜΑΣ ARDUINO ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ: ΣΧΕΔΙΑΣΜΟΣ ΜΑΘΗΣΙΑΚΩΝ ΔΡΑΣΤΗΡΙΟΤΗΤΩΝ ΜΕ ΒΑΣΗ ΤΟ ΠΛΑΙΣΙΟ ECLIP, ΙΟΡΔΑΝΗΣ Δ. ΚΟΥΣΗΣ	66

Πρόλογος

Ο τόμος αυτός περιλαμβάνει περιλήψεις επιλεγμένων διπλωματικών και πτυχιακών εργασιών που εκπονήθηκαν στο Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών κατά το διάστημα **01/01/2020 - 31/12/2020**. Πρόκειται για τον **18^ο τόμο** στη σειρά αυτή. Στόχος του θεσμού είναι η ενθάρρυνση της δημιουργικής προσπάθειας και η προβολή των πρωτότυπων εργασιών των φοιτητών του Τμήματος.

Η έκδοση αυτή είναι ψηφιακή, έχει δικό της ISSN και αναρτάται στην επίσημη ιστοσελίδα του Τμήματος ώστε να έχει μεγάλη προσβασιμότητα. Για το στόχο αυτό, σημαντική ήταν η συμβολή της Λήδας Χαλάτση που επιμελήθηκε και φέτος την ψηφιακή έκδοση και πέτυχε μια ελκυστική ποιότητα παρουσίασης, ενώ βελτίωσε και την ομοιογένεια των κειμένων.

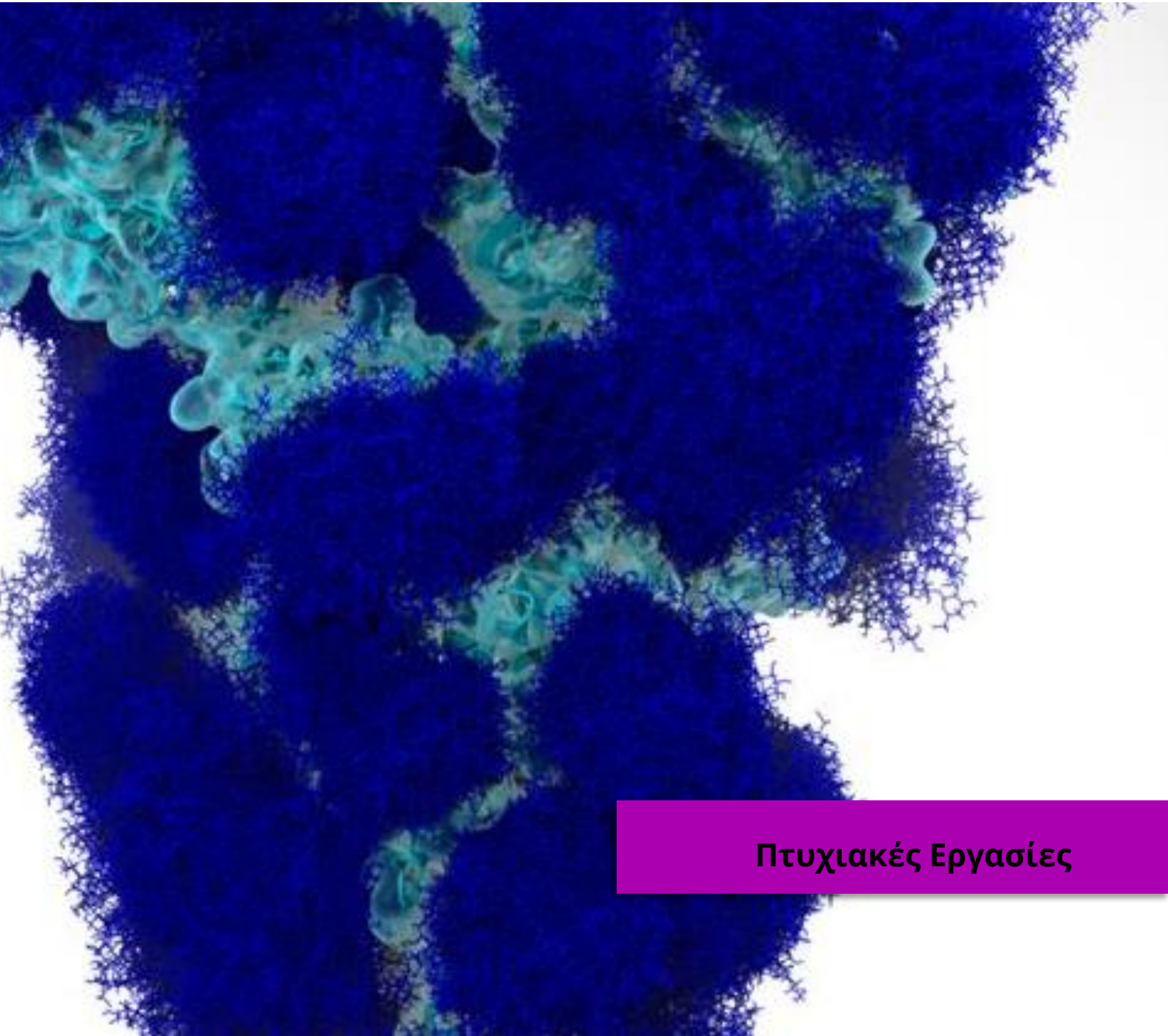
Η στάθμη των επιλεγμένων εργασιών είναι υψηλή και κάποιες από αυτές έχουν είτε δημοσιευθεί είτε υποβληθεί για δημοσίευση.

Θα θέλαμε να ευχαριστήσουμε τους φοιτητές για το χρόνο που αφιέρωσαν για να παρουσιάσουν τη δουλειά τους στα πλαίσια αυτού του θεσμού και να τους συγχαρούμε για την ποιότητα των εργασιών τους. Ελπίζουμε η διαδικασία αυτή να προσέφερε και στους ίδιους μια εμπειρία που θα τους βοηθήσει στη συνέχεια των σπουδών τους ή της επαγγελματικής τους σταδιοδρομίας.

Η Επιτροπή Ερευνητικών και Αναπτυξιακών Δραστηριοτήτων

Θ. Θεοχάρης (υπεύθυνος έκδοσης), Η. Μανωλάκος

Αθήνα, Ιούνιος 2021



Πτυχιακές Εργασίες

Assessing the URxD 3D Face Recognition Algorithm on Synthetic Facial Data

Panteleimon D. Kanellis

ABSTRACT

The subject of face recognition has gone through extensive research for the past three decades. The current thesis examines the URxD project, a pioneering 3D face recognition algorithm presented in [1]. Recently, synthetic 3D face generation has achieved a high degree of realism and the idea behind this thesis is to explore how a state of art face recogniser can benefit from synthetic 3D face generation. Instead of testing the algorithm with a dataset of real 3D faces (something that was already done before), this thesis explores the performance of URxD in a database consisting of 100 computer generated 3D faces. Next, the results are compared with those extracted from the real faces and several conclusions and remarks are made.

Subject Area: Face Recognition

Keywords: synthetic faces, 3D database, facial expressions, face recognition algorithm, recognition rate

ADVISORS

Theoharis Theoharis, Professor NKUA

Antonios Danelakis, PhD

1 INTRODUCTION

1.1 Defining the Problem

It is safe to assume that most (if not all) face recognition algorithms are built with the purpose of being used on real human faces. While this is a fairly reasonable motive, what will happen if the same recognition techniques are used for synthetic faces instead of real ones? Imagine a room of 100 people sitting next to each other. The chances of two people looking alike is low and even if they do, there are distinct facial features that can uniquely identify each individual. Now imagine if we ask these people to perform a random facial expression of their liking. The similarity rate between two people is further diminished, since the facial expression creates a "noise" in our dataset (for example a different bone structure can alter the look of the same expression). The same cannot be guaranteed for synthetic faces. Even if the computer generates faces randomly, there is still a chance that a random face will be very close to a previously created face (the parameters might be close). Random facial expressions don't help much, because the random generator might choose a similar expression. This is why face recognition for computer generated faces is a harder problem to solve.

In this thesis, a 3D database of 100 random synthetic faces was constructed with neutral facial expression. For each face, a random expression was chosen with random intensity. The entire dataset was given as input to the state of art 3D face recogniser URxD and signatures were produced for both the neutral and the random expression of the 3D face. The goal of the experiment is to observe how successfully can URxD correlate the neutral signature (also called input face) with the signature of the random expression of the same face (also called query face). We count the correct matches and try to explain why some of them failed.

1.2 Related Work

Blanz et al. [2], introduced the 3D Morphable Model (3DMM). In this concept, each face is divided in four segments as seen in Figure 1. Every other face is expressed as a linear combination of these segments with coefficients. The recognition stage between two faces is achieved by comparing the coefficients.

In a future publication [3], the same authors expanded upon the 3DMM by taking a 2D image and creating a 3DMM.

Georgios Passalis et al.[1] developed a face recognition algorithm based on the Annotated Face Model (AFM). The AFM as seen in Figure 2, is a 3D geometry with distinctly marked facial features. The AFM is fitted to each face in the database, producing a unique signature for each face. Comparing two 3D faces is reduced to comparing the signatures. URxD managed to score over 97% success rate on the Full FRGC v2 Database containing real faces.

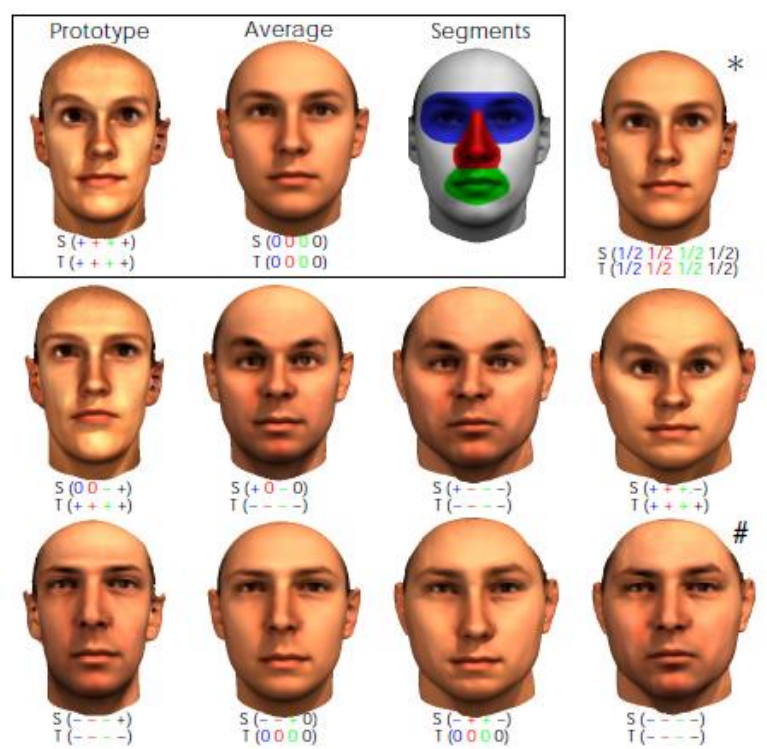


Figure 1: The prototype and the average face are used to generate new faces. The amount of deformation for each segment can influence the end result.



Figure 2: The Annotated Face Model.
Each facial feature is marked as a distinct area.

2 CHOOSING A RANDOM FACE GENERATOR

Among the several face reconstruction and synthetic face generation systems, we chose to explore two of the most prevalent ones: EOS and FaceGen. EOS [4] is a lightweight library written in C++11/14, and offers functionalities such as face reconstruction from a 2D image using a 3DMM and camera manipulation. The EOS source code needs to be compiled and requires additional dependencies. The results do not meet our quality standards, since most of them have incomplete and stretched texture maps during the reconstruction.

FaceGen [5] comes in two forms: the Demo version with a friendly interface but limited functionalities. In order to reconstruct a face from a 2D image, the landmarks must be placed manually (as opposed to EOS, where the landmarks are xy coordinates read from a file). Figure 4 describes the landmark placing stage and the reconstructed 3D face. The standard FaceGen version, the one we used for the thesis, offers a great variety of tools for image and mesh manipulation. The program functions with commands written by the user through a command line.



Figure 4: Manual placement of landmarks in a 2D face image and 3D face reconstruction using FaceGen.

3 CREATING A SYNTHETIC 3D FACE DATABASE WITH FACEGEN

A 3D database in this context is a collection of 3D objects in a specific file format. Several 3D facial databases with real faces exist, such as **Face Recognition Grand Challenge (FRGC)** [6], **Bosphorus** [7] and the **Binghamton University 3D Facial Expression (BU-3DFE) database** [8], to name a few.

We shall create our own 3D database consisting of synthetic faces instead of real ones. For each face, we store an .obj file for the neutral expression (input face) along with its texture and a frontal and side 2D image. The same face with an expression (query face) is stored in a different folder, with the .obj file and texture and a frontal 2D image. The structure for a single face is visualized in Figure 5. A batch script is responsible for generating the aforementioned structure for 100 faces.

FaceGen offers random values for the following parameters:

Gender, Age, Ethnicity, Asymmetry, Caricature Shape, Caricature Color.

A random face can have intermediate value for a specific category. For example, the gender of a person might be between male and female. FaceGen provides a set of 22 facial expressions with intensity that can be given as a parameter by the user. In this work, both the expression and its intensity are chosen randomly.

In order to generate the database, we must first understand the workflow of creating a single face with FaceGen. Figure 7 briefly describes the workflow : FaceGen depends on the 3DMM concept and as such it includes a set of average faces called csam. For our purposes, we choose the animate csam to include expressions. Every other face will be generated as a linear combination of the average face. The shape of the average face with fixed topology is stored in a Statistical Shape Model (SSM), represented by the .EGM extension. The Statistical Color Model (SCM), stored as a file with .EGT extension, describes the mean texture map. In order to use a csam file, we need facial coordinates (included in a .fg file). These coordinates combined with the EGM and EGT will transform the average face and will output the base geometry corresponding to the new face (.tri file, which can be easily exported to .obj).

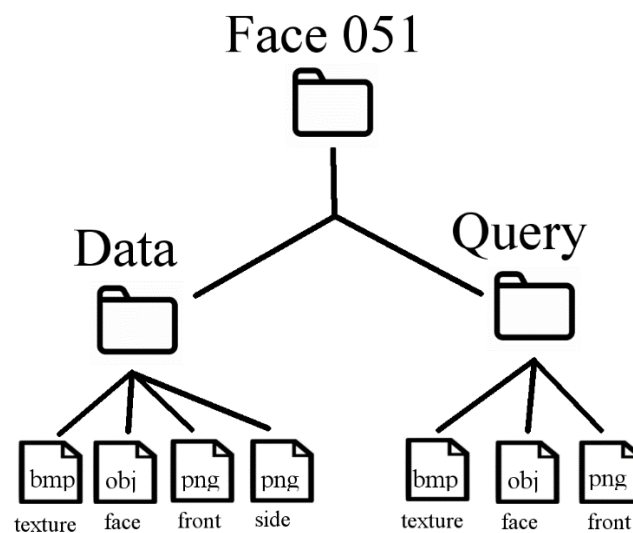


Figure 5: The structure of the database for a single facial mesh.

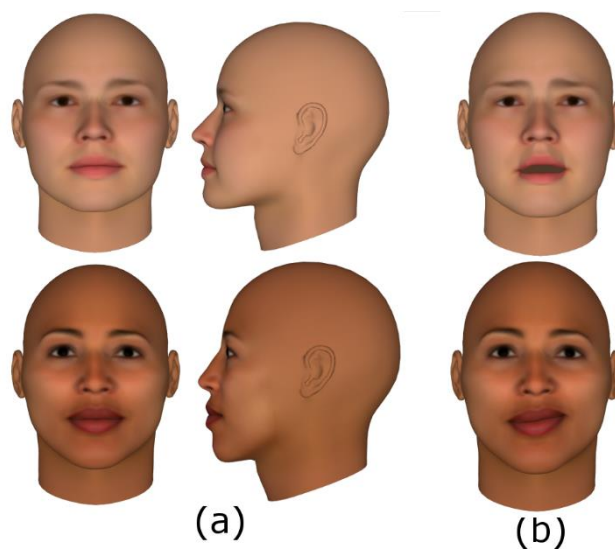


Figure 6: Random faces created by Facegen.

a) Neutral expressions b) Random expression with random intensity

4 UNDERSTANDING THE URxD WORKFLOW

Before we test our newly created database to yield recognition rates, it is crucial to understand how the state of the art 3D face recognition system URxD works. The source code of URxD (consisting of 72 projects) was modified and recompiled

from scratch, in order to be able to run in modern x64 machines. In total, 55 total executable programs were produced, but we will only need two of them in order to run experiments. The workflow mentioned below was also implemented in a database consisting of real faces, solely for comparison purposes.

Fimgen takes as input the 3D faces of the database, performs a registration, alignment and fitting process, and outputs a unique signature of each face with fixed size (121 KB). As seen in Figure 7, the program is parametrized by a configuration file (config.ini). Inside config.ini, we specify the Annotated Face Model that will be aligned and later wrapped around the 3D face in the alignment and fitting process respectively, in order to create the signature.

Simutil collects all the fwv files produced by fimgen and creates a $N \times N$ similarity matrix, where N is the number of fwv files. In this thesis, N is 200, since 200 faces (100 query plus 100 database faces) in the FaceGen database produce 200 signatures with fimgen. A similarity matrix contains the names of the signatures as rows and columns and each cell describes how similar two signatures are. An example similarity matrix is seen in Figure 8. The cell located in the first row and third column, suggests that the distance from face 001 to face 002 is 234.296.

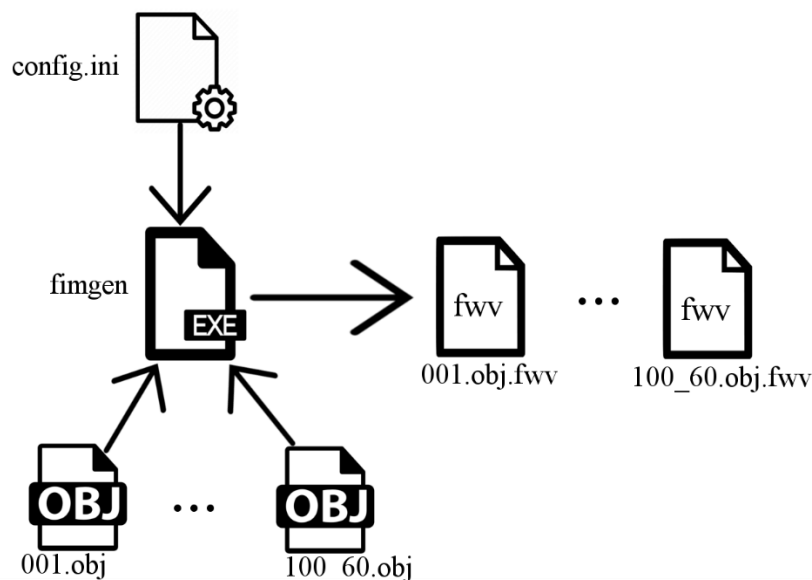


Figure 7: Fimgen workflow.

	001	001_68	002
001	0	135.637	234.296
001_68	135.637	0	253.699
002	234.296	253.699	0

Figure 8: Similarity matrix.

5 FACE RECOGNITION WITH URxD

A similarity matrix alone is not enough to extract recognition rates. A Python script was created to interpret the similarity matrix visually, by counting the correct results and outputting a graph. The results seen in Figure 10, are represented in Cumulative Match Characteristics (CMC) curves. In machine learning, a CMC curve is a method of showing the performance of recognition precision for each rank. Table 1 aggregates the success rates for each available sas file for the two databases. A sas file represents an Annotated Face Model that will wrap around the 3D face in the fitting phase.

Table 1: URxD success rate (%) for different sas files

sas file	FaceGen	BU-3DFE
half_face	67	91
mean_face	43	91
quarter_face	38	91
slim_face2	40	90
slim_full_face2	42	91
slim_full_face_bones	42	91

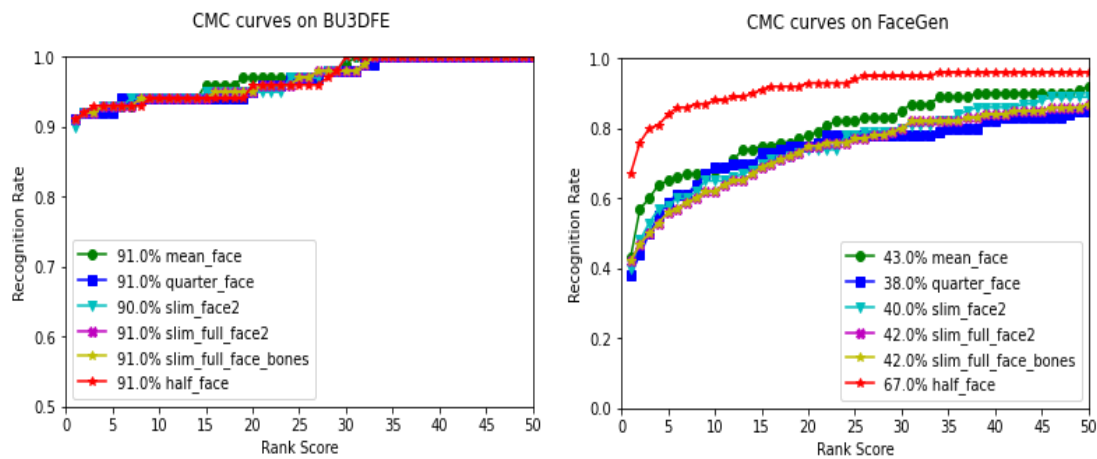


Figure 10: CMC curves for the BU-3DFE (left), and the FG database (right) for each available sas file

6 DISCUSSIONS AND CONCLUSIONS

As seen in the previous chapter, URxD has a low recognition rate for artificial faces. Despite the innate difficulty of identifying faces that have similar shape and texture, another culprit is the registration phase. Many faces (and that only happened with the FaceGen DB), were extremely rotated in the registration phase and as a result, the corresponding procedure tried to fit the AFM into the rotated face, and the algorithm produces a distorted fwv signature. Figure 11 depicts this phenomenon. An attempt to counter this effect was to rotate manually a face along a certain axis and then provide the rotated face as input for fimgen. This process is described in Figure 12, but fimgen is very sensitive to rotations and a definitive rotation value and axis orientation is difficult to achieve for each face.

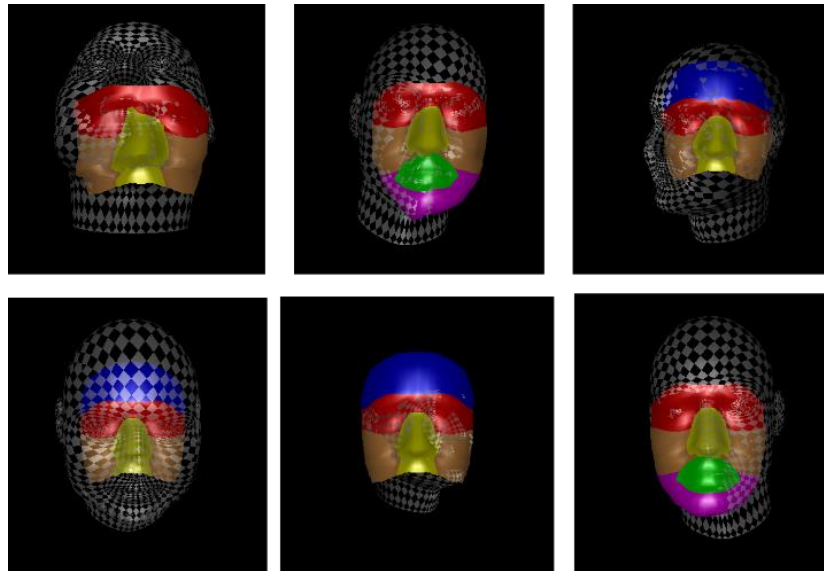
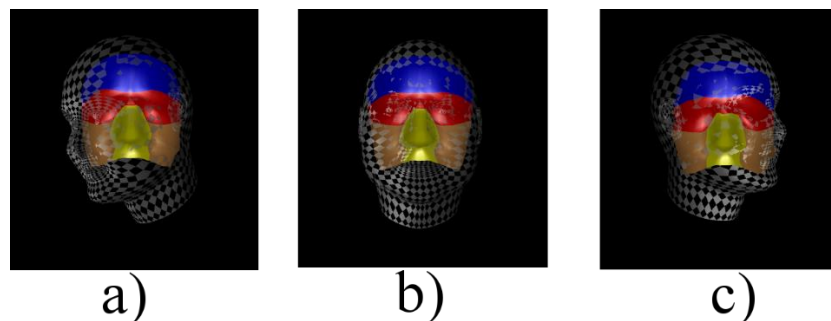


Figure 11: Incorrect registration phase of fimgen.



a)

b)

c)

Figure 12: Fimgen results on the same 3D face

a) produced by FaceGen without any transformations

b) manually rotated 30 degrees on the Z axis

c) manually rotated 30 degrees on the Z axis and -12 degrees on the X axis

6.1 Future Work

Below we present a few methods that might improve the recognition rate of URxD:

1. Take into account the texture map. The fwv distance along with the colour of the face can improve the recognition rate of the algorithm and help reevaluate failed identifications. Apart from the shape of the face, the fwv signature can contain additional information about the texture of the face, which means slightly larger signatures. If this technique

increases the running time significantly, it can only be used to clear up ambiguities regarding the identity of the face.

2. Higher resolution AFM. In this thesis, URxD was run in a 5 year old moderate laptop with 30-35 seconds for each face (in debug version; a release version is faster). A modern day computer can reduce the running time at least by half. Subsequently, the resolution of the AFM can be increased, so that more details can be captured. This method can be combined with the technique proposed in [9]. This technique takes advantage of the symmetry of the face and the half AFM mask can have a higher resolution.

There is another area where the current thesis can be applied to, one that saw a big rise in the last 10 years: Neural Networks and Machine Learning. The advancements in processing power gave the opportunity for face recognition algorithms to step into the area of Neural Networks. It would be interesting to witness how such a synthetic database could be employed, e.g. for training a machine learning face recogniser.

REFERENCES

- [1] Ioannis Kakadiaris, Georgios Passalis, George Toderici, Mohammed Murtuza, Yunliang Lu, Nikos Karampatziakis, and Theoharis Theoharis. Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. *IEEE transactions on pattern analysis and machine intelligence*, 29:640–9, 05 2007.
- [2] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3d faces. *SIGGRAPH'99 Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 09 2002.
- [3] V.Blanz and T.Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1063–1074, 2003.
- [4] EOS GitHub Repository. <https://github.com/patrikhuber/eos>

- [5] FaceGen Official Website. <https://facegen.com/>
- [6] FRGC 3D databse. <https://www.nist.gov/programs-projects/face-recognition-grand-challenge-frgc>
- [7] Bosphorus 3D databse. <http://bosphorus.ee.boun.edu.tr/Home.aspx>
- [8] BU-3DFE 3D database. https://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html
- [9] T. Theoharis, P.Perakis, G.Passalis, and I.A. Kakadiaris. Using facial symmetry to handle pose variations in real-world 3d face recognition. IEEE Transactions on Pattern Analysis & Machine Intelligence, 33(10):1938–1951, oct 2011.

Εισαγωγή στη Θεωρία Μητροειδών

Μερκούρης Παπαμιχαήλ

ΠΕΡΙΛΗΨΗ

Στην πτυχιακή εργασία παρουσιάσαμε μια εισαγωγή στη Θεωρία Μητροειδών μέσα από κάποια βασικά αποτελέσματα της βιβλιογραφίας. Μελετήσαμε την έννοια της ανεξαρτησίας, μέσω της διακριτής δομής του μητροειδούς. Είδαμε πως η ανεξαρτησία εμφανίζεται στη Θεωρία Γραφημάτων και την Γραμμική Άλγεβρα και θα αποδείξαμε κάποιες θεμελιώδεις ιδιότητες που εμφανίζουν τα ανεξάρτητα σύνολα. Στη συνέχεια, αξιωματικοποιήσαμε τις ιδιότητες διατυπώνοντας διαφορετικές δομές ανεξαρτησίας ή αναπαραστάσεις μητροειδών και αποδείξαμε την ισοδυναμία τους. Τέλος, εξετάσαμε τη σύνδεση των μητροειδών με προβλήματα βελτιστοποίησης, όπου ορίσαμε τον άπληστο αλγόριθμο. Εκεί αποδείξαμε ότι η βελτιστότητα του αλγορίθμου εξαρτάται από τον βαθμό, που η είσοδος προσεγγίζει την οργάνωση του μητροειδούς.

Θεματική Περιοχή: Διακριτά Μαθηματικά

Λέξεις Κλειδιά: μητροειδές, γραμμική ανεξαρτησία, άπληστος αλγόριθμος, διακριτές δομές, βελτιστοποίηση, γραμμική άλγεβρα, θεωρία γραφημάτων

ΕΠΙΒΛΕΠΩΝ

Σταύρος Κολλιόπουλος, Καθηγητής ΕΚΠΑ

1 ΕΙΣΑΓΩΓΗ

Η Θεωρία Μητροειδών εισήχθη από τον Hassler Whitney το 1935 [1],[2], ενώ είχε ανακαλυφθεί παράλληλα και ανεξάρτητα, από τον Ιάπωνα μαθηματικό Takeo Nakasawa [3]. Στο άρθρο του ο Whitney έδωσε τα αξιώματα ανεξαρτησίας και ονόμασε οποιαδήποτε δομή που ικανοποιεί αυτά τα αξιώματα, μητροειδές. Η κεντρική του παρατήρηση ήταν ότι αυτά τα αξιώματα παρέχουν μια αφαίρεση της έννοιας της ανεξαρτησίας που εμφανίζεται τόσο στα γραφήματα, όσο και στους πίνακες. Εξ αιτίας αυτού μεγάλο κομμάτι της ορολογίας, όπως θα δούμε και αργότερα, προέρχεται από την Θεωρία Γραφημάτων και την Γραμμική Άλγεβρα. Ένας άλλος σπουδαίος ερευνητής σε αυτή την περιοχή υπήρξε ο William Thomas Tutte, ο οποίος έγραψε ένα από τα πρώτα βιβλία στην Θεωρία Μητροειδών (1971), ως μια συλλογή διαλέξεων του συγγραφέα στην RAND Corporation το 1965. Τέλος, την σύνδεση μεταξύ Θεωρίας Μητροειδών και βελτιστοποίησης, την χρωστάμε στον Jack Edmonds [2], ο οποίος έδειξε ότι ένα μητροειδές μπορεί να οριστεί αλγοριθμικά, από την άπληστο αλγόριθμο και σε μια σειρά εργασιών διερεύνησε τις σχέσεις ανάμεσα στα μητροειδή και την πολυεδρική συνδυαστική [4].

1.1 Σχετική Βιβλιογραφία

Στην ελληνόφωνη βιβλιογραφία δεν υπάρχει κάποιο εισαγωγικό βιβλίο στην Θεωρία Μητροειδών, τουλάχιστον κατά την γνώση του συγγραφέα. Κάποιες αναφορές στην Θεωρία Μητροειδών μπορεί να βρει κανείς στα Διακριτές Μαθηματικές Δομές για την Επιστήμη Υπολογιστών [5] και στην ελληνική μετάφραση του Introduction to Algorithms [6],[7]. Παρ' όλα αυτά, έχουν γραφεί πλήθος εργασιών στην Ελληνική γλώσσα με θέμα την Θεωρία Μητροειδών. Αναφέρουμε σχετικά τις ακόλουθες διπλωματικές εργασίες 891011 και την διδακτορική διατριβή του Αθανάσιου Κουτσώνα [12]. Παρ' όλα αυτά, οι παραπάνω εργασίες εξετάζουν κάποιες πλευρές ή αποτελέσματα της Θεωρίας Μητροειδών, και δεν έχουν σαν σκοπό την εισαγωγή στο αντικείμενο.

Σε αυτό το πλαίσιο έρχεται η παρούσα εργασία ώστε να συμπληρώσει την υπάρχουσα βιβλιογραφία με μια Εισαγωγή στην Θεωρία Μητροειδών, για την Ελληνική γλώσσα. Η εργασία απευθύνεται σε προπτυχιακούς φοιτητές της

Επιστήμης Υπολογιστών και των Μαθηματικών, οι οποίοι έχουν παρακολουθήσει κάποιο μάθημα Διακριτών Μαθηματικών. Δεν απαιτείται προϋπάρχουσα γνώση Θεωρίας Γραφημάτων ή Γραμμικής Άλγεβρας, αφού όλες οι απαραίτητες έννοιες ορίζονται εκ νέου. Σκοπός της εργασίας είναι να παράσχει στους ενδιαφερόμενους θεμελιώδεις έννοιες και ορισμούς, ώστε να διευκολύνεται η ανάγνωση εξειδικευμένων εργασιών ή συγγραμμάτων.

Κύριες πηγές για την παρούσα εργασία είναι τα εισαγωγικά συγγράμματα στην Αγγλική γλώσσα, *Matroid Theory* [1] του Αυστραλοαμερικάνου James G. Oxley καθηγητή στο Louisiana State University και το *Topics in Matroid Theory* [2] του καθηγητή του Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης Λεωνίδα Σ. Πιτσούλη. Στην δομή της παρουσιάσής μας θα ακολουθήσουμε το [2]. Παρ' όλα αυτά, πολλές αποδείξεις, προήλθαν από το [1].

2 ΚΥΡΙΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΠΑΡΑΤΗΡΗΣΕΙΣ

Στην πτυχιακή εργασία κινηθήκαμε επαγωγικά προς τον ορισμό της έννοιας του μητροειδούς. Ξεκινήσαμε με την αναφορά σε κάποιους στοιχειώδεις ορισμούς και παρατηρήσεις. Είδαμε την εμφάνιση της ανεξαρτησίας στα δάση γραφημάτων και στα γραμμικώς ανεξάρτητα σύνολα διανυσμάτων. Στην συνέχεια, συγκεντρώσαμε τις παρατηρήσεις μας σε συστήματα αξιωμάτων. Έτσι, ορίσαμε διαφορετικές αναπαραστάσεις μητροειδών, όπου αποδεικνύονται ισοδύναμες. Τέλος, παρατηρούμε την σύνδεση των μητροειδών με προβλήματα βελτιστοποίησης. Εκεί ορίζουμε τον άπληστο αλγόριθμο και δείχνουμε έναν αλγοριθμικό χαρακτηρισμό για τα μητροειδή.

2.1 Διακριτές Συναρτήσεις και Υπομετρικότητα

Η κεντρικότερη, ίσως, έννοια που εισήχθη στο πρώτο κεφάλαιο της πτυχιακής εργασίας είναι αυτή της υπομετρικής συνάρτησης¹. Όπως θα δούμε στην

¹ Στην πτυχιακή εργασία ακολουθήσαμε την μετάφραση του Γεωργίου Σταθόπουλου για τον όρο submodularity 9.

συνέχεια, η συνάρτηση που αντιστοιχεί σε ένα σύνολο το μέγεθος τους μεγαλύτερου ανεξάρτητου υποσυνόλου του είναι υπομετρική.

Ορισμός 2.1.1 *Υπομετρικότητα (Submodularity).* Έστω Ω ένα πεπερασμένο σύνολο. Έστω, επίσης, μια συνάρτηση από το δυναμοσύνολο του Ω στους πραγματικούς αριθμούς, $f : 2^\Omega \rightarrow \mathbb{R}$. Θα λέμε ότι η f είναι υπομετρική (submodular), αν για κάθε $X \subseteq Y$ και $x \in \Omega \setminus Y$, ισχύει:

$$f(X \cup x) - f(X) \geq f(Y \cup x) - f(Y), \quad \forall X \subseteq Y \subseteq \Omega, x \in \Omega \setminus Y \quad (1)$$

Δηλαδή, η αξία που παίρνουμε προσθέτοντας το x στο πιο «άδειο» σύνολο X είναι μεγαλύτερη, από το να προσθέσουμε το ίδιο x στο πιο «κορεσμένο» σύνολο Y . Αποδεικνύεται πως ισοδύναμα μπορούμε να χρησιμοποιήσουμε αντί της (1), την ακόλουθη σχέση:

$$f(S) + f(T) \geq f(S \cup T) + f(S \cap T), \quad \forall X, Y \subseteq \Omega \quad (2)$$

2.2 Παρατηρήσεις στην Θεωρία Γραφημάτων και την Γραμμική Άλγεβρα

Στο δεύτερο κεφάλαιο της πτυχιακής ακολουθήσαμε την παρουσίαση του [2]. Καταπιαστήκαμε με δύο, φαινομενικά ξένα, πεδία των μαθηματικών, την Θεωρία Γραφημάτων και τους Διανυσματικούς Χώρους. Όπως θα δούμε στην συνέχεια, σε αυτά τα δύο πεδία σημαντικό ρόλο παίζει η έννοια της ανεξαρτησίας. Θα την συναντήσουμε στα δέντρα και τα δάση στην Θεωρία Γραφημάτων, όπως και στα σύνολα γραμμικώς ανεξάρτητων διανυσμάτων του Διανυσματικού Χώρου. Καθώς προχωράμε, θα μας αναπτυχθεί η διαίσθηση πως παρά τις διαφορές τους ορισμένα θεωρήματα διατυπώνουν την ίδια μαθηματική αλήθεια, απλά σε διαφορετικές γλώσσες. Έτσι θα έχουμε προετοιμάσει το έδαφος ώστε στην συνέχεια να αφήσουμε πίσω μας τις ειδικές περιπτώσεις και θα επιχειρήσουμε να διατυπώσουμε έναν γενικότερο ορισμό της ανεξαρτησίας.

Μέσα από τις παρατηρήσεις μας, δημιουργείται η εντύπωση πως η έννοια της γραμμικής ανεξαρτησίας σε διανυσματικούς χώρους «μοιάζει» με την ιδιότητα της ακυκλικότητας σε γραφήματα. Είναι γνωστό πως ένα άκυκλο υπογράφημα ονομάζεται και δάσος (forest). Όπως θα δούμε παρακάτω, τα γραμμικώς

ανεξάρτητα σύνολα μοιράζονται με τα δάση ενός γραφήματος την Ιδιότητα Επαύξεσης.

Θεώρημα 2.2.1 *Ιδιότητα Επαύξεσης (Augmentation Property).* Έστω ένα γράφημα G . Θεωρούμε δύο δάση F_1, F_2 , με $|E(F_1)| > |E(F_2)|$. Τότε θα υπάρχει κάποια ακμή e στο $E(F_1) \setminus E(F_2)$ τέτοια ώστε το $F_2 \cup e$ να είναι δάσος.

Ομοίως μπορούμε να σκεφτούμε και για σύνολα γραμμικώς ανεξάρτητων διανυσμάτων ενός διανυσματικού χώρου. Πράγματι, θεωρήστε έναν διανυσματικό χώρο V και δύο σύνολα γραμμικώς ανεξάρτητων διανυσμάτων X_1, X_2 , του V , με $|X_1| > |X_2|$, τότε θα υπάρχει πάντα κάποιο διάνυσμα x στο $X_1 \setminus X_2$ τέτοιο ώστε, αν προσθέσουμε το x στο X_2 , θα έχουμε πάλι ένα γραμμικώς ανεξάρτητο σύνολο.

Ακολουθώντας την σύνδεση μεταξύ γραμμικής ανεξαρτησίας και ακυκλικότητας, γνωρίζουμε πως ένα μέγιστο άκυκλο γράφημα ονομάζεται δάσος *επικάλυψης* (spanning forest). Υπενθυμίζουμε, επίσης, πως οι βάσεις ενός διανυσματικού χώρου είναι ακριβώς τα μεγιστικά γραμμικώς ανεξάρτητα σύνολά του. Αποδεικνύεται πως τόσο τα δάση επικάλυψης, όσο και οι βάσεις ενός διανυσματικού χώρου μοιράζονται την Ιδιότητα της ανταλλαγής.

Θεώρημα 2.2.2 *Ιδιότητα Ανταλλαγής (Exchange Property).* Έστω ένα γράφημα G . Έστω, επίσης, δύο δάση επικάλυψης F_1, F_2 . Τότε θα υπάρχει κάποια ακμή f_1 του F_1 και κάποια ακμή f_2 του F_2 τέτοια ώστε το υπογράφημα $(F_1 \setminus f_1) \cup f_2$ να είναι δάσος επικάλυψης.

Εντελώς, συμμετρικά και αντικαθιστώντας τα δάση επικάλυψης με βάσεις ενός διανυσματικού χώρου, έχουμε την αντίστοιχη ιδιότητα στην Γραμμική Άλγεβρα.

Μένοντας στην ίδια αντιστοιχία, εφόσον έχουμε θεωρήσει τα δάση ως ανεξάρτητα (παρ' όλο που δεν έχουμε προχωρήσει σε έναν αυστηρό ορισμό), τότε τα ελαχιστικά εξαρτημένα γραφήματα θα είναι *κύκλοι*. Έτσι, αποδεικνύεται πως τόσο οι κύκλοι, όσο και τα ελαχιστικά γραμμικώς εξαρτημένα σύνολα διανυσμάτων διαθέτουν την ιδιότητα της απαλοιφής.

Θεώρημα 2.2.3 *Ιδιότητα Απαλοιφής (Elimination Property).* Έστω ένα γράφημα G . Έστω, επίσης, δύο κύκλοι C_1, C_2 , με C_1 διαφορετικό του C_2 . Θεωρούμε, επίσης να

υπάρχει κάποια κοινή ακμή e των C_1, C_2 . Τότε το γράφημα $(C_1 \cup C_2) \setminus e$ είναι ένας κύκλος του G .

Ομοίως, σκεφτόμαστε και για ελαχιστικά γραμμικώς εξαρτημένα σύνολα διανυσμάτων, στην Γραμμική Άλγεβρα.

Στην συνέχεια εστιάζουμε στον *βαθμό* (rank) ενός γραφήματος και ενός συνόλου γραμμικώς ανεξάρτητων διανυσμάτων. Ειδικότερα, θα λέμε βαθμό $r(G)$ ενός γραφήματος G , το μέγεθος ενός δάσους επικάλυψής του,

$$r(G) = |E(F)|, \quad \text{όπου } F \text{ ένα δάσος επικάλυψης του } G$$

Επιπλέον, σε ένα πεπερασμένο σύνολο γραμμικώς ανεξάρτητων διανυσμάτων S , ονομάζουμε βαθμό $r(S)$ το μέγεθος του μεγαλύτερο γραμμικώς ανεξάρτητου υποσυνόλου του.

$$r(S) = |T|, \quad \text{όπου } T \subseteq S, T \text{ γραμμικώς ανεξάρτητο}$$

Τόσο ο βαθμός ενός γραφήματος, όσο και ενός συνόλου διανυσμάτων μοιράζονται κάποιες πολύ συγκεκριμένες ιδιότητες. Ειδικότερα, η συνάρτηση βαθμού είναι υπομετρική (submodular). Δίνουμε παρακάτω μόνο τις ιδιότητες της συνάρτησης βαθμού σε γραφήματα.

Θεώρημα 2.2.4 *Ιδιότητες Συνάρτησης Βαθμού.* Για ένα γράφημα G η συνάρτηση βαθμού ικανοποιεί τα παρακάτω:

Εύρος τιμών της συνάρτησης, $0 \leq r(G) \leq |E(G)|$.

Αύξουσα συνάρτηση. Για $H \subseteq_{\text{υπ}} G$, $r(H) \leq r(G)$.

Υπομετρικότητα. Έστω δύο υπογραφήματα $H, K \subseteq_{\text{υπ}} G$, τότε ισχύει:

$$r(H \cup K) \leq r(H) + r(K) - r(H \cap K)$$

2.3 Ισοδύναμοι Ορισμοί Μητροειδούς

Οι παρατηρήσεις που κάναμε προηγουμένως μας επιτρέπουν να ανεβούμε ένα επίπεδο αφαίρεσης. Θα ονομάσουμε τα αντικείμενα με τα οποία θα ασχοληθούμε μητροειδή (matroids), τα οποία θα ορίσουμε μόνο μέσω των

αξιωμάτων που θα υπακούουν. Ένα μητροειδές μπορεί να κωδικοποιεί είτε ένα σύνολο διανυσμάτων, είτε το σύνολο των δέντρων ενός γραφήματος, είτε οποιοδήποτε άλλο μαθηματικό αντικείμενο στο οποίο παρουσιάζεται η έννοια της ανεξαρτησίας. Θα δούμε ότι σε ένα μητροειδές εμφανίζονται οι έννοιες των βάσεων, των κυκλωμάτων (βλ. κύκλοι σε ένα γράφημα), αλλά και ο βαθμός.

Μάλιστα, θα ανακαλύψουμε πως αυτές οι έννοιες είναι αλληλένδετες. Αποδεικνύοντας ορισμένες ιδιότητες για κάθε μια τους, και αναγάγωντάς τις σε αξιώματα θα οδηγηθούμε σε διαφορετικούς, αλλά ισοδύναμους, ορισμούς ενός μητροειδούς. Στο παρόν κεφάλαιο θα δούμε τέσσερεις τέτοιους ορισμούς. Τα *1-μητροειδή*, τα οποία βασίζονται στην έννοια της ανεξαρτησίας. Τα *β-μητροειδή*, τα οποία βασίζονται στην έννοια της βάσης. Τα *γ-μητροειδή*, τα οποία βασίζονται στην έννοια του κυκλώματος. Και, τέλος, τα *τ-μητροειδή*, τα οποία βασίζονται στη συνάρτηση βαθμού. Οι περισσότερες αποδείξεις που παρουσιάστηκαν στην πτυχιακή εργασία προέρχονται από το [1].

2.3.1 Ανεξάρτητα Σύνολα

Ξεκινάμε με έναν αφηρημένο ορισμό της ανεξαρτησίας. Αφού δεν κάνουμε καμία παραδοχή για τα στοιχεία που χαρακτηρίζουμε ανεξάρτητα, θα θεωρήσουμε ένα αφηρημένο πεπερασμένο σύνολο E . Την ανεξαρτησία, επί του E , εκφράζει το σύνολο I , το οποίο ορίζει ποια στοιχεία από το E είναι ανεξάρτητα μεταξύ τους. Παρ' όλα αυτά για να είναι συνεπές το I με την διαίσθηση που μας έχει δημιουργηθεί για την ανεξαρτησία απαιτούμε να υπακούει σε ορισμένα αξιώματα. Το σύνολο E , μαζί με την οργάνωση του I , θα αποτελούν ένα σύστημα ανεξαρτησίας. Ένα μητροειδές θα είναι ένα σύστημα ανεξαρτησίας, στο οποίο απαιτούμε να ισχύει η ιδιότητα επαύξησης.

Ειδικότερα, δίνουμε τους ακόλουθους ορισμούς.

Ορισμός 2.3.1 Σύστημα Ανεξαρτησίας. Έστω ένα πεπερασμένο σύνολο E . Έστω, επίσης, μια συλλογή συνόλων από το E , $I \subseteq 2^E$. Θα λέμε ότι η δυάδα $(E, I) = S$ είναι ένα σύστημα ανεξαρτησίας (independence system), αν υπακούει στα εξής αξιώματα:

$\emptyset \in \mathcal{I}$

Αν $I \in \mathcal{I}$ και $I' \subseteq I$, τότε $I' \in \mathcal{I}$.

Όπως είπαμε, το μητροειδές είναι μια ειδική περίπτωση συστήματος ανεξαρτησίας, όπου αξιωματοποιούμε την ιδιότητα επαύξησης.

Ορισμός 2.3.2 *I-Μητροειδές*. Έστω ένα σύστημα ανεξαρτησίας $M = (E, \mathcal{I})$. Θα λέμε ότι το M είναι ένα I -μητροειδές ή απλά μητροειδές (matroid), αν ικανοποιεί επίσης το αξίωμα επαύξησης:

Ιδιότητα Επαύξησης. Αν τα I_1 και I_2 ανήκουν στον \mathcal{I} , και $|I_1| > |I_2|$, τότε υπάρχει ένα στοιχείο $e \in I_1 \setminus I_2$ τέτοιο ώστε $I_2 \cup e \in \mathcal{I}$.

Στον παραπάνω ορισμό, αναφερθήκαμε σε I -μητροειδές, επειδή σε αυτόν τον ορισμό τα αξιώματα αφορούν την οικογένεια ανεξάρτητων συνόλων \mathcal{I} . Στις επόμενες ενότητες θα δούμε εναλλακτικούς ορισμούς για τα μητροειδή, οι οποίοι βασίζονται σε διαφορετικά σύνολα αξιωμάτων².

Από την παραπάνω συζήτηση προκύπτει άμεσα πως δεδομένου ενός γραφήματος G , μπορούμε να κατασκευάσουμε ένα μητροειδές $M_G = (E, \mathcal{F})$, όπου $E = E(G)$ και το \mathcal{F} να περιέχει τα δάση του G . Όπως είδαμε παραπάνω η δομή αυτή θα ικανοποιεί τα αξιώματα $I1$ - $I3$. Θα λέμε το μητροειδές που προκύπτει από γράφημα **γραφικό μητροειδές** (graphic matroid). Ομοίως μπορούμε να θεωρήσουμε ένα πεπερασμένο σύνολο διανυσμάτων S και να θεωρήσουμε το μητροειδές $M_S = (S, \mathcal{I})$, όπου το \mathcal{I} να περιέχει όλα τα γραμμικώς ανεξάρτητα υποσύνολα του S . Θα ονομάζουμε το μητροειδές που προκύπτει από ένα πεπερασμένο σύνολο διανυσμάτων **διανυσματικό μητροειδές** (vector matroid).

2.3.2 Βάσεις

Γενικεύουμε, τώρα, την έννοια της βάσης, ενός διανυσματικού χώρου, και του δάσους επικάλυψης, ενός γραφήματος. Έτσι, όταν μιλάμε για *βάση* (basis) σε ένα

2 Οι αναγνώστες που είναι εξοικειωμένοι με την υπολογιστική γεωμετρία μπορούν να κάνουν την αναλογία με την αναπαράσταση ενός κυρτού πολυτόπου. Εκεί κάνουμε τον διαχωρισμό μεταξύ V -αναπαράστασης, όπου μας δίνονται οι κορυφές του πολυτόπου, και H -αναπαράστασης, όπου μας δίνονται τα υπερεπίπεδα των πλευρών του.

σύστημα ανεξαρτησίας, θα εννοούμε ένα μεγιστικό ανεξάρτητο σύνολο. Αποδεικνύεται πως οι βάσεις σε ένα μητροειδές έχουν τις εξής ιδιότητες.

Θεώρημα 2.3.1 *Ιδιότητες Βάσης.* Έστω ένα μητροειδές M και B το σύνολο βάσεων του. Θα ισχύει:

$$B1. \quad B \neq \emptyset$$

$B2.$ *Ιδιότητα Ανταλλαγής.* Αν $B1, B2$ δύο στοιχεία του B , τότε για κάθε $x \in B1 \setminus B2$, υπάρχει $y \in B2 \setminus B1$, τέτοιο ώστε το $(B1 \setminus x) \cup y \in B$.

Μάλιστα, μπορούμε να κινηθούμε και προς την αντίθετη πλευρά. Δηλαδή, να ορίσουμε μία δομή $M = (E, B)$, όπου $B \subseteq 2^E$ και να ικανοποιεί τα αξιώματα $B1 - B2$. Θα ονομάζουμε αυτή την δομή *B-μητροειδές*. Βεβαίως, αποδεικνύεται πως τα B-μητροειδή είναι *ισοδύναμα* με τα I-μητροειδή. Ειδικότερα, δεδομένου του συνόλου βάσεων B , μπορούμε να κατασκευάσουμε την οικογένεια ανεξάρτητων συνόλων I , από την σχέση,

$$I = \{ I \subseteq E \mid \text{υπάρχει } B \in B, \text{ τέτοιο ώστε } I \subseteq B \} \quad (3)$$

Αποδεικνύεται πως η δομή $M' = (E, I)$, όπως έχει οριστεί παραπάνω ικανοποιεί τα αξιώματα $I1 - I3$.

2.3.3 Κυκλώματα

Ως τώρα έχουμε μιλήσει για τα ανεξάρτητα σύνολα ενός μητροειδούς. Μάλιστα, στην προηγούμενη ενότητα, κάναμε λόγο για τα μεγιστικά ανεξάρτητα σύνολα, τις βάσεις. Σε αυτή την ενότητα θα ασχοληθούμε με τα εξαρτημένα σύνολα, ενός μητροειδούς M . Τα ελαχιστικά εξαρτημένα σύνολα θα τα λέμε *κυκλώματα* (circuits) και είναι η γενίκευση των κύκλων γραφημάτων σε μητροειδή. Αποδεικνύεται πως το σύνολο κυκλωμάτων ενός μητροειδούς έχει τις ακόλουθες ιδιότητες.

Θεώρημα 2.3.2 *Ιδιότητες Κυκλωμάτων.* Έστω ένα μητροειδές M και C το σύνολο κυκλωμάτων του. Θα ισχύει:

$$C1. \quad \emptyset \text{ δεν ανήκει στο } C$$

C2. Αν C_1, C_2 στοιχεία του C , με $C_1 \subseteq C_2$, τότε $C_1 = C_2$

C3. *Ιδιότητα Απαλοιφής.* Αν C_1, C_2 δύο διακεκριμένα στοιχεία του C και $e \in C_1 \cap C_2$, τότε υπάρχει ένα στοιχείο $C_3 \in C$ τέτοιο ώστε $C_3 \subseteq (C_1 \cup C_2) \setminus e$.

Αν αξιωματοποιήσουμε τις παραπάνω ιδιότητες, τότε μία δομή $M_C = (E, C)$, που ικανοποιεί τα C1 – C3, θα λέγεται *C-μητροειδής*. Ομοίως με πριν, αποδεικνύεται πως τα C-μητροειδή είναι ισοδύναμα με τα I-μητροειδή. Ειδικότερα, δεδομένου του συνόλου κυκλωμάτων C , μπορούμε να κατασκευάσουμε την οικογένεια ανεξάρτητων συνόλων I , από την σχέση,

$$I = \{ I \subseteq E \mid \text{δεν υπάρχει } C \in C, \text{ τέτοιο ώστε } C \subseteq I \} \quad (4)$$

Αποδεικνύεται πως η δομή $M' = (E, I)$, όπως έχει οριστεί παραπάνω ικανοποιεί τα αξιώματα I1 – I3.

2.3.4 Βαθμός

Όπως, προηγουμένως, ορίζουμε ως βαθμό $r(X)$ ενός συνόλου $X \subseteq E$, ενός μητροειδούς $M = (E, I)$, ως το μέγεθος του μεγαλύτερου ανεξάρτητου υποσυνόλου του³. Αποδεικνύονται οι εξής ιδιότητες.

Θεώρημα 2.3.3 *Ιδιότητες Συνάρτησης Βαθμού.* Έστω ένα μητροειδής $M = (E, I)$. Έστω, επίσης, δύο σύνολα $X, Y \subseteq E$. Τότε, ισχύει:

$$R1. \text{ Αν } X \subseteq E, \text{ τότε } 0 \leq r(X) \leq |X|$$

$$R2. \text{ Αν } X \subseteq Y \subseteq E, \text{ τότε } r(X) \leq r(Y)$$

$$R3. \text{ Υπομετρικότητα. } r(X \cup Y) + r(X \cap Y) \leq r(X) + r(Y)$$

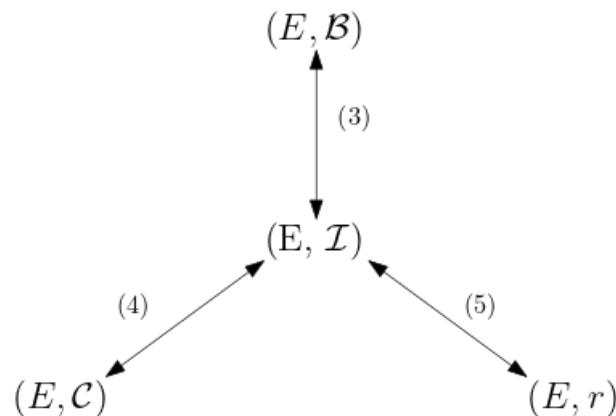
Αξιωματικοποιώντας τις παραπάνω ιδιότητες ορίζουμε το r -μητροειδής $M = (E, r)$. Η ισοδυναμία του r -μητροειδούς με το I -μητροειδής προκύπτει από την σχέση,

$$I = \{ I \subseteq E \mid r(I) = |I| \} \quad (5)$$

3 Η συνάρτηση βαθμού είναι καλώς ορισμένη, αφού αποδεικνύεται πως όλα τα μεγιστικά ανεξάρτητα υποσύνολα ενός συνόλου $X \subseteq E$, είναι ισοπληθικά.

2.3.5 Συμπεράσματα

Στην συζήτηση που προηγήθηκε στις προηγούμενες ενότητες είδαμε πως μπορούμε να αναπαραστήσουμε το ίδιο μαθηματικό αντικείμενο με διαφορετικά σύνολα αξιωμάτων. Ξεκινήσαμε από μια δυάδα (E, I) , για την οποία διατυπώσαμε ένα σύνολο αξιωμάτων που πρέπει να ικανοποιεί το σύνολο I , στα $I1 - I3$. Ονομάσαμε το (E, I) μητροειδές. Στην συνέχεια ονομάσαμε το σύνολο μεγιστικών συνόλων από το I , σύνολο βάσεων B του μητροειδούς. Έπειτα, αποδείξαμε ορισμένες βασικές ιδιότητες για το B . Στην συνέχεια αξιωματοποιήσαμε αυτές τις ιδιότητες στα αξιώματα $B1 - B2$, ενώ περιγράψαμε την δυάδα (E, B) ως B -μητροειδές. Είδαμε ότι οι δυάδες (E, I) και (E, B) είναι το ίδιο αντικείμενο και η (3) είναι η σχέση που ορίζει το I , δοθέντος του B . Έτσι, είδαμε πως ο κατάλογος αξιωμάτων $I1 - I3$ οδηγεί στον $B1 - B2$ και το αντίστροφο. Ανάλογες συνδέσεις κάναμε και για τα σύνολα αξιωμάτων που βασίζονται στις δυάδες (E, C) και (E, r) . Τον καρπό των κόπων μας, παρουσιάζουμε στο ακόλουθο σχήμα.



Σχήμα 1: Σχηματική αναπαράσταση των διαφορετικών ορισμών για μητροειδή. Μέσω της σχέσης (3) πηγαίνουμε από τα B -μητροειδή στα I -μητροειδή. Μέσω της σχέσης (4) πηγαίνουμε από τα C -μητροειδή στα I -μητροειδή. Τέλος, μέσω της σχέσης (5) πηγαίνουμε από τα r στα I -μητροειδή.

2.4 Άπληστος Αλγόριθμος

Προηγουμένως, εξετάσαμε πολλούς διαφορετικούς ορισμούς για την ίδια μαθηματική έννοια. Εδώ θα δώσουμε ακόμη έναν ορισμό για τα μητροειδή, χωρίς όμως να παρουσιάσουμε άλλον ένα κατάλογο αξιωμάτων. Τα μητροειδή

αποτελούν σπάνια περίπτωση μαθηματικού αντικειμένου, το οποίο μπορεί να χαρακτηριστεί από έναν αλγόριθμο. Επιπλέον, η φύση του αλγοριθμικού ορισμού αναδεικνύει τη σημασία των μητροειδών στην θεωρία βελτιστοποίησης. Και σε αυτό το κεφάλαιο της πτυχιακής εργασίας ακολουθήσαμε την παρουσίαση του [2].

Θεωρήστε το ακόλουθο πρόβλημα βελτιστοποίησης σε συστήματα ανεξαρτησίας. Θυμηθείτε πως στα συστήματα ανεξαρτησίας απαιτούμε να ισχύουν μόνο τα αξιώματα I1 – I2 και όχι το Αξίωμα Επαύξεσης.

Πρόβλημα 2.4.1 *Μεγιστοβαρούς (Ελαχιστοβαρούς) Βάσης.* Έστω ένα σύστημα ανεξαρτησίας $S = (E, I)$. Έστω, επίσης, κάποια συνάρτηση $w : E \rightarrow \mathbb{R}$, η οποία να είναι γραμμική, δηλαδή για κάθε $X \subseteq E$ να ισχύει

$$w(X) = \sum_{e \in X} w(e)$$

και $w(\emptyset) = 0$, κατόπιν συμβάσεως. Θεωρούμε το ακόλουθο πρόβλημα διακριτής βελτιστοποίησης σε συστήματα ανεξαρτησίας,

$$\begin{array}{ll} \max(\min) & w(X) \\ \text{τ. ω.} & X \in B \end{array}$$

όπου B το σύνολο βάσεων του συστήματος ανεξαρτησίας S .

Θα λέμε κάθε σύνολο $X \subseteq E$ *εφικτή λύση* (feasible solution), αν $X \in B$, ενώ αν επιπλέον είναι και *μεγιστοβαρές* θα κάνουμε λόγο για *βέλτιστη λύση* (optimum solution). Τα περισσότερα προβλήματα συνδυαστικής βελτιστοποίησης μπορούν να εκφραστούν σαν προβλήματα βελτιστοποίησης πάνω σε συστήματα ανεξαρτησίας, όπου ο στόχος είναι να βρεθεί μια μεγιστοβαρής ή ελαχιστοβαρής βάση.

Το Πρόβλημα Μεγιστοβαρούς Βάσης, που θέλουμε να αντιμετωπίσουμε είναι μια γενίκευση του Προβλήματος Ελάχιστου Δάσους Επικάλυψης. Έτσι, θα γενικεύσουμε τον γνωστό Αλγόριθμο του Kruskal, που επιλύει το τελευταίο. Ειδικότερα, δίνουμε τον εξής αλγόριθμο, το οποίο ονομάζουμε *άπληστο*.

Αλγόριθμος 4.2.1: Άπληστος Αλγόριθμος

Είσοδος: Ένα βεβαρυμένο σύστημα ανεξαρτησίας $S = (E, \mathcal{I})$, με συνάρτηση βάρους $w: E \rightarrow \mathbb{R}$.

Έξοδος: Μία εφικτή λύση στο **Πρόβλημα Μεγιστοβαρούς Βάσης**.

```

1  $B \leftarrow \emptyset$ 
2 while  $E \neq \emptyset$  do
3    $e \leftarrow \arg \max\{w(e) \mid e \in E\}$ 
4    $E \leftarrow E \setminus e$ 
5   if  $B \cup e \in \mathcal{I}$  then
6      $B \leftarrow B \cup e$ 
7   end
8 end
```

Σχήμα 2: Ο Άπληστος Αλγόριθμος.

Αποδεικνύεται πως η βελτιστότητα του άπληστου αλγόριθμου σε ένα σύστημα ανεξαρτησίας ισοδυναμεί με την Ιδιότητα Επαύξεσης. Ειδικότερα, δίνουμε τον ακόλουθο ορισμό.

Ορισμός 2.4.1 *Άπληστο Μητροειδές.* Έστω κάποιο σύστημα ανεξαρτησίας $S = (E, \mathcal{I})$. Θα λέμε ότι το S είναι ένα άπληστο μητροειδές, ή απλά μητροειδές (matroid)· αν για οποιαδήποτε συνάρτηση βάρους $w: E \rightarrow \mathbb{R}$ το Πρόβλημα Μεγιστοβαρούς Βάσης, που προκύπτει, επιλύεται βέλτιστα από τον Άπληστο Αλγόριθμο.

Φυσικά, τα άπληστα μητροειδή είναι ισοδύναμα με τα \mathcal{I} -μητροειδή, αφού αν ο Άπληστος Αλγόριθμος δίνει την βέλτιστη λύση για το $S = (E, \mathcal{I})$, τότε το \mathcal{I} έχει την Ιδιότητα Επαύξεσης.

Ο τελευταίος ορισμός διαφέρει από εκείνους που εξετάσαμε προηγουμένως. Εδώ δεν χρησιμοποιήσαμε έναν κατάλογο αξιωμάτων για να χαρακτηρίσουμε κάποια σύνολα. Αντίθετα, παρατηρήσαμε πως αν η είσοδος σε έναν αλγόριθμο, διαθέτει κάποια ισχυρή οργάνωση, τότε και μόνον τότε, ο αλγόριθμος θα δίνει την βέλτιστη λύση. Έτσι, είδαμε πως μπορούμε να χρησιμοποιήσουμε μια αφηρημένη διακριτή δομή για να επιχειρηματολογήσουμε για προβλήματα βελτιστοποίησης.

3 ΠΕΡΑΙΤΕΡΩ ΘΕΜΑΤΑ

Η θεωρία μητροειδών έχει προχωρήσει πολύ από το άρθρο του Whitney το 1935 και έχει επεκταθεί σε ένα πλήρη κλάδο των μαθηματικών με πληθώρα εφαρμογών στην επιστήμη της πληροφορικής. Στην παρούσα εργασία παρουσιάσαμε μόνο τις βασικές έννοιες αυτής της πλούσιας θεωρίας. Για περισσότερα σας παραπέμπουμε στα [1],[2], τα οποία υπήρξαν και οι κύριες πηγές για την πτυχιακή εργασία.

Επιπλέον, έχουν προταθεί διαφορές ενδιαφέρουσες παραλλαγές την έννοιας του μητροειδούς. Το 1980 σε μια προσπάθεια για περαιτέρω χαρακτηρισμό των προβλημάτων που επιδέχονται τον άπληστο αλγόριθμο, προτάθηκαν από τον Γερμανό μαθηματικό Bernhard H. Korte και τον Ούγγρο μαθηματικό László Lovász η έννοια του απληστοειδούς (greedoid) [13]. Εκεί το αξίωμα (I2) έχει αντικατασταθεί από την Ιδιότητα Προσβασιμότητας (Accessibility Property). Αντί να απαιτούμε για κάποιο ανεξάρτητο σύνολο I κάθε υποσύνολό του να είναι επίσης ανεξάρτητο, ζητάμε να υπάρχει κάποιο $e \in I$ τέτοιο ώστε το $I \setminus e$ να είναι επίσης ανεξάρτητο [13].

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] James G. Oxley, Matroid Theory. Oxford University Press, 1992.
- [2] Leonidas S. Pitsoulis, Topics in Matroid Theory. Springer, 2014
- [3] Hirokazu Nishimura and Susumu Kuroda, A Lost Mathematician, Takeo Nakasawa: The forgotten father of matroid theory. Birkhauser Verlag, 2009
- [4] Alexander Schrijver, Combinatorial Optimization: Polyhedra and Efficiency, 2003.
- [5] Δημήτριος Γεωργίου and Ευστάθιος Αντωνίου and Ανέστης Χατζιμιχαηλίδης, Διακριτές Μαθηματικές Δομές για την Επιστήμη Υπολογιστών. Σύνδεσμος Ελλήνων Ακαδημαϊκών Βιβλιοθηκών, 2015.
- [6] Thomas H. Cormen and Charles E. Leiserson and Ronald L. Rivest and Clifford Stein, Introduction to Algorithms. MIT Press, 2001.

- [7] Thomas H. Cormen and Charles E. Leiserson and Ronald L. Rivest and Clifford Stein, Εισαγωγή στους Αλγορίθμους. Πανεπιστημιακές Εκδόσεις Κρήτης, 2003.
- [8] Μαρία Πετροπαναγιωτάκη, Παραμετρικοί Αλγόριθμοι και Μητροειδή: η χρήση των συνόλων αντιπροσώπευσης, <https://pergamos.lib.uoa.gr/uoa/dl/object/1325895>, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, 2016.
- [9] Γεώργιος Σταθόπουλος, Πολύεδρα, Μητροειδή και Υπομετρικά Συστήματα, Θεωρία και Εφαρμογές, <https://pergamos.lib.uoa.gr/uoa/dl/object/1319497>, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, 2012.
- [10] Θεόδωρος Δημητρακόπουλος, Προσεγγιστικοί Αλγόριθμοι για το Generalized Min-Sum Set Cover, <https://pergamos.lib.uoa.gr/uoa/dl/object/1319647>, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, 2014.
- [11] Κρυσταλλένια Τσιρικίδου, Αλγόριθμος Αναγνώρισης Γραφικών Μητροειδών, <http://ikee.lib.auth.gr/record/291881/>, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 2015.
- [12] Αθανάσιος Κουτσώνας, Διδιαστατότητα: Θεωρία και Αλγοριθμικές Εφαρμογές, <https://pergamos.lib.uoa.gr/uoa/dl/object/1308818>, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, 2014.
- [13] Anders Björner and Günter M. Ziegler, Introduction to greedoids, 1992.

Σχήμα Οδηγός-Ακόλουθος Κινούμενων Ρομπότ με τη χρήση RGB-D Κάμερας και MobileNets

Παναγιώτης Α. Πετροπουλάκης

ΠΕΡΙΛΗΨΗ

Τα Ρομποτικά Συστήματα που ενεργούν σε αδόμητα περιβάλλοντα, όπως σε αυτά που και εμείς δρούμε, βασίζονται κυρίως στην όραση και στις αισθήσεις τους μέσω της χρήσης διαφόρων αισθητήρων. Για να μπορέσουν τα συστήματα αυτά να αποτυπώσουν και να καταλάβουν τον περιβάλλοντα χώρο τους αποδοτικά, η Τεχνητή Νοημοσύνη προσφέρει καινοτόμες, αλλά και ανθεκτικές μεθόδους, όπως ο εντοπισμός των αντικειμένων (Object Detection). Στην παρούσα εργασία, εξετάζουμε την απόδοση ενός τέτοιου συστήματος εντοπισμού που ονομάζεται MobileNets [1] σε ένα σχήμα Οδηγός-Ακόλουθος Κινούμενων Ρομπότ με την χρήση RGB-D Κάμερας. Όπως θα αναλύσουμε, η επιλογή αυτού του συστήματος εντοπισμού έγινε αποκλειστικά με βάση τις τεχνικές προδιαγραφές του Ρομποτικού Συστήματός μας και των επιμέρους στοιχείων του. Ακόμα, τα δεδομένα που εκπαιδεύουμε το μοντέλο αυτό είναι πραγματικά, τα οποία τα διαθέτουμε προς ελεύθερη χρήση. Πολλαπλά πειράματα στα Pioneer 3-AT και Summit-XL Ρομπότ έδειξαν ότι τα MobileNets προσφέρουν εξαιρετική ακρίβεια. Το σύστημα Οδηγός-Ακόλουθος που προτείνουμε λειτουργεί στα 19Hz στην GeForce GTX1060 GPU με το πολύ 5% λανθασμένες προβλέψεις.

Ο πηγαίος κώδικας του συστήματος μας, τα βίντεο από τα πειράματα που πραγματοποιήσαμε και τα δεδομένα εκπαίδευσης βρίσκονται στον παρακάτω σύνδεσμο: https://github.com/PetropoulakisPanagiotis/BSc_thesis.

Θεματική Περιοχή: Ρομποτική

Λέξεις-Κλειδιά: Ρομποτική, Οπτικός Έλεγχος, Σχήμα Οδηγός-Ακόλουθος, Βαθιά Μάθηση, RGB-D Κάμερα

ΕΠΙΒΛΕΠΟΝΤΕΣ ΚΑΘΗΓΗΤΕΣ

Κωσταντίνος Κυριακόπουλος, Καθηγητής, Εθνικό Μετσόβιο Πολυτεχνείο

Ιωάννης Εμίρης, Καθηγητής, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

1 ΕΙΣΑΓΩΓΗ

Τα Αυτόνομα Ρομποτικά Συστήματα σταδιακά υπεισέρχονται σε διάφορες πτυχές της ζωής μας. Από τους βιομηχανικούς βραχίονες που βρίσκονται ήδη σε εργοστάσια που εκτελούν βαριές εργασίες, θα μεταβούμε σε μια διαφορετική πραγματικότητα. Drones θα αρχίσουν να παραδίδουν δέματα στα σπίτια μας, καθώς και η συνεργασία Κινούμενων Ρομπότ σε περιπτώσεις αναζήτησης και διάσωσης θα γίνει πραγματοποιήσιμη (Εικόνα 1).



Εικόνα 1: Παραδείγματα Ρομποτικών Συστημάτων

Τα περισσότερα από αυτά τα Αυτόνομα Συστήματα βασίζονται κυρίως σε αισθητήρες όρασης ή αλλιώς κάμερες για κατανοήσουν τον περιβάλλοντα χώρο τους και να πάρουν τις κατάλληλες αποφάσεις. Ως αποτέλεσμα, στην Ρομποτική, ο τομέας του ελέγχου μέσω όρασης, Vision-Based Control[2-3], εξελίσσεται ραγδαία.

Πιο ακριβέστερα, ο έλεγχος μέσω όρασης αναφέρετε στην χρήση οπτικών δεδομένων για την εξαγωγή ταχυτήτων ενός Ρομπότ. Τα δεδομένα όρασης τις περισσότερες φορές προέρχονται από μια κάμερα η οποία είναι τοποθετημένη πάνω σε ένα Ρομπότ, ή αλλιώς eye-in-hand διάταξη. Αυτή η ερευνητική περιοχή βασίζεται ως επί το πλείστον σε τεχνικές μηχανικής όρασης. Ιδιαίτερα, καινοτόμες λύσεις πάνω σε προβλήματα όρασης από τις επιστημονικές περιοχές Object Detection, Image Segmentation, Image Reconstruction και Object Tracking, έχουν προσφέρει λύσεις σε σημαντικά προβλήματα όρασης που απασχολούν την Ρομποτική.

Με την σειρά μας, εξετάζουμε ένα σχήμα Οδηγός-Ακόλουθος με την χρήση RGB-D κάμερας επικεντρώνοντας κυρίως την προσοχή μας στο πρώτο σκέλος του Vision-Based Control, δηλαδή την εξαγωγή χρήσιμων οπτικών δεδομένων τα οποία είναι ιδιαίτερα σημαντικά για τον έλεγχο του Ρομπότ Ακόλουθου. Το σχήμα αυτό αποτελείτε καταβάση από δύο Ρομπότ και μια κάμερα και ο τελικός στόχος του είναι ο Οδηγός Ρομπότ να βρίσκεται μονίμως στο πεδίο όρασης του Ακόλουθου Ρομπότ.

Η πρώτη δυσκολία σε αυτά τα σχήματα έγκειται στον εντοπισμό του Οδηγού Ρομπότ, συνήθως μέσω της χρήσης της RGB πληροφορίας της κάμερας. Η πρόσφατη βιβλιογραφία της μηχανικής όρασης πλέον δείχνει, ότι οι κλασικές μέθοδοι εντοπισμού, για παράδειγμα η εξαγωγή hard-crafted features με μεθόδους, όπως η SURF[4] και η ORB[5], έχουν πλέον αντικατασταθεί με τεχνικές βαθιάς μηχανικής μάθησης, οι οποίες φαίνεται να αποτελούν και τον δρόμο για να οδηγηθούμε στην πλήρη αυτονομία.

Επειδή το σύστημα μας πρέπει να λαμβάνει σε ελάχιστο χρόνο αποφάσεις, είναι επιτακτική η ανάγκη εύρεσης και επιλογής μιας τεχνικής εντοπισμού αντικειμένων που θα εξάγει αποτελέσματα σε πραγματικό χρόνο με βάση τα τεχνικά χαρακτηριστικά του συστήματός μας και ιδιαίτερα της κάρτας γραφικών (GeForce GTX1060) όπου λαμβάνουν χώρα όλοι οι υπολογισμοί των δεδομένων της κάμερας.

Συνεπώς, στην παρούσα εργασία, διεξάγουμε μια αναλυτική ανασκόπηση της βιβλιογραφίας των τεχνικών εντοπισμού αντικειμένων, κατά την οποία καταλήγουμε στην υιοθέτηση των MobileNets[1] για το σύστημά μας. Η πλήρης ανάλυση βρίσκεται στην πρωτότυπη έκδοση, όπου ξεκινάμε από τα R-CNN variants[6-8], συνεχίζουμε με τεχνικές εξαγωγής αποτελεσμάτων σε πραγματικό χρόνο, YOLO[9], SSD[10], YOLO9000[11], αναλύουμε ένα βασικό συστατικό των MobileNets, τα μοντέλα ResNet[12], και τέλος αναφέρουμε τα MobileNets[13,1]. Στην παρούσα περίληψη, αναφέρουμε αναλυτικά την πρώτη και βασική δημοσίευση των MobileNets[13].

Ακολούθως, αναλύουμε τον τρόπο που εκπαιδεύουμε το μοντέλο μας και τις διάφορες μετατροπές που πραγματοποιούμε για να πετύχουμε τα επιθυμητά αποτελέσματα. Στην συνέχεια, εξετάζουμε την συμπεριφορά του συστήματος

εντοπισμού σε πραγματικά πειράματα πριν συνεχίσουμε με το κομμάτι του ελέγχου.

Μόλις λύσουμε το πρόβλημα εντοπισμού του Οδηγού Ρομπότ, πρέπει να χρησιμοποιήσουμε την πληροφορία που εξάγει το σύστημα αυτό, δηλαδή τις συντεταγμένες εικονοστοιχείων ενός κουτιού που περικλείει τον Οδηγό Ρομπότ, με σκοπό τον υπολογισμό κατάλληλων ταχυτήτων για το Ακόλουθο Ρομπότ ώστε να οδηγηθεί κατάλληλα προς τον Οδηγό του. Σε αυτό το σημείο, χρησιμοποιούμε την πληροφορία βάθους της κάμερας και βασιζόμενοι στο κεντρικό θεώρημα προοπτικής της εικόνας, εκτιμούμε προσεγγιστικά την τρισδιάστατη θέση του Οδηγού Ρομπότ. Τέλος, χρησιμοποιώντας έναν ελεγκτή, υπολογίζουμε κατάλληλες ταχύτητες και ελέγχουμε την συμπεριφορά του ολικού συστήματος σε πειράματα με πραγματικά Ρομπότ.

Κατά το τελευταίο κεφάλαιο, αναφέρουμε τα συμπεράσματά μας, αλλά και τις αδυναμίες του συστήματός μας, δίνουμε κατευθύνσεις για το πως μπορεί να βελτιωθεί αυτό το σύστημα και κάνουμε μια συζήτησή πάνω σε διάφορες δημοσιεύσεις που σχετίζονται άμεσα με το πρόβλημα που αντιμετωπίσαμε και για το πως μπορούν να εφαρμοστούν σε τέτοια σχήματα.

2 ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

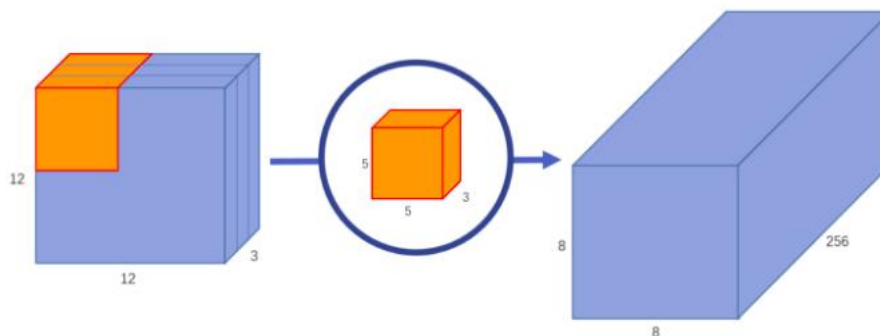
2.1 MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications

Ο Andrew G. Howard και η ομάδα του [13], πρότειναν μια κλάση αποτελεσματικών μοντέλων τα οποία ονομάζονται MobileNets. Στην δημοσίευσή τους, κατάφεραν να δημιουργήσουν μοντέλα τα οποία είναι αποτελεσματικά ως προς το μέγεθος και την ταχύτητά τους. Αυτό το πέτυχαν αντικαθιστώντας τα τυπικά convolutional layers με τα αποδοτικά depthwise separable layers που πρότειναν. Επίσης, δημιούργησαν και δύο καθολικές υπερ-παραμέτρους με σκοπό την περαιτέρω μείωση των εσωτερικών παραμέτρων των μοντέλων αυτών με ομοιόμορφο τρόπο. Τα νέα αυτά layers πρόκειται για

ένας είδος παραγοντοποίησης των τυπικών convolutional layers, κατά την οποία, η τυπική πράξη αυτή διασπάτε σε δύο βήματα, σε depthwise και σε 1x1 pointwise convolutions.

Μια τυπική πράξη convolution ($s=1, p=1$) με N φίλτρα μεγέθους $D_K \times D_K \times M$, πάνω σε ένα feature map εισόδου με διαστάσεις $D_F \times D_F \times M$, εξάγει ένα τελικό feature map διαστάσεων $D_F \times D_F \times N$, με συνολικό κόστος $N * D_K * D_K * M * D_F * D_F$ πολλαπλασιασμών.

Σαν παράδειγμα (Εικόνα 2), τα 256 φίλτρα ($s=1, p=0$) διαστάσεων $5 \times 5 \times 3$, εξάγουν δεδομένα διαστάσεων $8 \times 8 \times 256$, με συνολικό κόστος $256 * (5 * 5 * 3 * 8 * 8) = 1.228.800$ πολλαπλασιασμών.



Εικόνα 2: Τυπική πράξη convolution

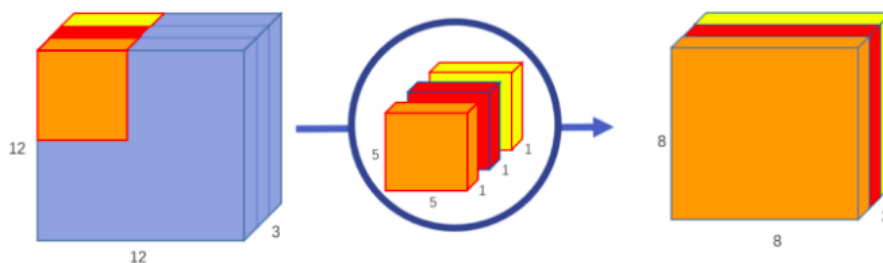
Στα MobileNets όμως, ο όρος $N (=256)$ στο κόστος πολλαπλασιασμών κλαδεύεται με τη χρήση depthwise separable convolutions τα οποία έχουν την ιδιότητα να “διασπούν τα στάδια φιλτραρίσματος και συνδυασμού” της πράξης convolution.

Αρχικά, κατά την πράξη depthwise convolution, εφαρμόζετε ένα φίλτρο ($s=1, p=1$) σε ένα feature map ανά κανάλι εισόδου. Ένα αρχικό feature map διάστασης $D_F \times D_F \times M$ μετατρέπεται σε μια έξοδο διαστάσεων $D_F \times D_F \times M$ όταν εφαρμοστούν N φίλτρα διαστάσεων $D_K \times D_K \times 1$, με συνολικό κόστος $D_K * D_K * M * D_F * D_F$ πολλαπλασιασμών. Με βάση το προηγούμενο παράδειγμά μας, 3 φίλτρα διαστάσεων $5 \times 5 \times 1$, παράγουν έξοδο διαστάσεων $8 \times 8 \times 3$ (Εικόνα 3) με συνολικό κόστος $5 * 5 * 3 * 8 * 8 = 4.800$ πολλαπλασιασμών.

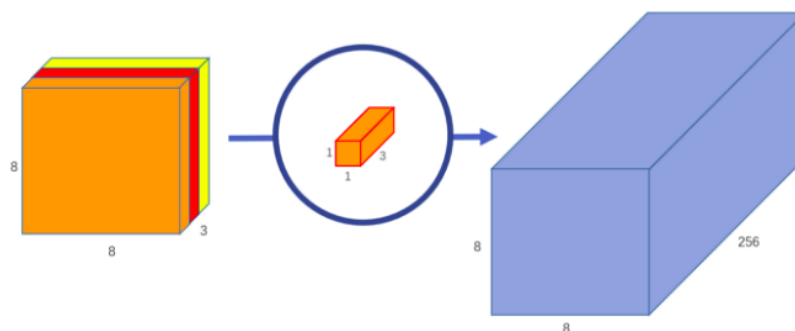
Για να συνδυάσουμε τα feature maps που έχουν προκύψει (Εικόνα 4) κατά το προηγούμενο βήμα, εκτελούμε πλέον-μόνο-μια φορά, αντί για $N(=256)$, την πράξη 1×1 pointwise convolution. Το συνολικό κόστος στο προηγούμενο παράδειγμά μας από 1.228.800 πολλαπλασιασμούς, μειώνετε σε κόστος $53.952 = 49.152 + 4.800$ πολλαπλασιασμών.

Με βάση των ορισμό της δημοσίευσης, τα depthwise separable convolutions χρειάζονται να εκτελεστούν $D_K * D_K * M * D_F * D_F + N * D_F * D_F * M$ πολλαπλασιασμοί. Η μείωση του κόστους που παίρνουμε είναι της τάξεως:

$$\frac{D_K * D_K * M * D_F * D_F + N * D_F * D_F * M}{N * D_K * D_K * M * D_F * D_F} = \frac{1}{N} + \frac{1}{D_K^2}$$



Εικόνα 3: Depthwise convolution πράξη



Εικόνα 4: Pointwise convolution πράξη

Το βασικό μοντέλο που προτείνουν οι συγγραφείς, αποτελείτε από 28 layers και φίλτρα depthwise διαστάσεων 3×3 . Μετά από κάθε depthwise και pointwise πράξη, ακολουθεί ένα batchnormalization layer και ως activation function έχουν

επιλέξει την ReLU. Πιο συγκεκριμένα, μόνο το πρώτο layer είναι μια τυπική convolution πράξη. Τα υπόλοιπα convolutions είναι depthwise separable και στο τέλος, έχουν προσθέσει ένα average pooling και ένα fc layer, το οποίο τροφοδοτεί τα δεδομένα σε ένα τελικό softmax layer για την εξαγωγή ετικετών. Ακόμα, οι συγγραφείς προτείνουν μικρό regularization και ελάχιστο ή και καθόλου weight decay, επειδή τα μοντέλα αυτά είναι μικρά, καθώς διαθέτουν λίγες εσωτερικές παραμέτρους, και επομένως, είναι και λιγότερο επιρρεπή σε over-fitting.

Τέλος, οι συγγραφείς ακόμα πρότειναν δύο καθολικές υπερ-παραμέτρους, οι οποίες μειώνουν ακόμα περισσότερο το μέγεθος των μοντέλων αυτών, σε περίπτωση που πρέπει να χρησιμοποιηθούν σε συστήματα που διαθέτουν μικρή υπολογιστική ισχύ (π.χ. κινητά). Η πρώτη από αυτές, ονομάζεται width multiplier α , και αυτή μπορεί να “λεπτύνει” τα μοντέλα αυτά. Πιο συγκεκριμένα, κάθε feature map με διάσταση εισόδου M και διάσταση εξόδου N περικόπτετε σε feature map με διάσταση εισόδου αM και εξόδου αN και με αυτό τον τρόπο πετυχαίνουμε μια τετραγωνική μείωση του κόστους α^2 . Η τελευταία παράμετρος ονομάζεται resolution multiplier ρ , η οποία προσαρμόζει το μέγεθος την εικόνας εισόδου. Με την χρήση των παραμέτρων αυτών, το συνολικό κόστος πολλαπλασιασμών της κλάσης αυτών των μοντέλων γίνεται πλέον: $D_K * D_K * \alpha M * \rho D_F * \rho D_F + \alpha N * \rho D_F * \rho D_F * M$.

3 ΣΧΗΜΑ ΟΔΗΓΟΣ-ΑΚΟΛΟΥΘΟΣ

Σε αυτό το κεφάλαιο θα ορίσουμε τυπικά το σχήμα μας Οδηγός-Ακόλουθος με βάση τον ορισμό από την δημοσίευση [14] χωρίς να λάβουμε υπόψη μας περιορισμούς εισόδου και θα αναφέρουμε επίσης τα προβλήματα που προσπαθούμε να επιλύσουμε στην περίπτωση μας.

3.1 Περιγραφή Προβλήματος

Ορισμός 1. Θέτουμε $d > 0$ και αναπαριστούμε δύο Ρομπότ ως $R_0 = P_0^T$ και $R_1 = P_1^T$. Λέμε ότι τα R_0 και R_1 ανήκουν ασυμπτωτικά σε ένα d-σχήμα με Οδηγό Ρομπότ το R_0 , αν $\lim_{t \rightarrow \infty} (P_0(t) - (P_1(t) + d)) = 0$.

Πρόβλημα 1. Θέτουμε $d > 0$ και αναπαριστούμε δύο Ρομπότ ως $R_0 = P_0^T$ και $R_1 = P_1^T$. Βρείτε τις κατάλληλες ταχύτητες v_1 και ω_1 του R_1 , ώστε τα R_0 και R_1 να ανήκουν ασυμπτωτικά σε ένα d-σχήμα με Οδηγό Ρομπότ το R_0 .

Στην περίπτωση μας θέτουμε όπου $d = 1.7m$, ώστε να αποφύγουμε την μεταξύ επαφή των δύο Ρομπότ. Ο αριθμός αυτός εξαρτάτε κατά βάση από το μέγεθός τους.

Πρόβλημα 2. Το ποσοστό των λανθασμένων οπτικών προβλέψεων του συστήματος εντοπισμού μέσω της χρήσης της εικόνας της κάμερας στο 1.7-σχήμα δεν θα πρέπει να είναι παραπάνω από 5%.

4 ΜΕΘΟΔΟΛΟΓΙΑ

4.1 Εντοπισμός Οδηγού



Εικόνα 5: Παραδείγματα δεδομένων εκπαίδευσης

Από την στιγμή που δεν υπάρχουν στον παγκόσμιο ιστό παραδείγματα εκπαίδευσης για τύπου Pioneer 3-AT Ρομπότ (Οδηγός), συλλέξαμε συνολικά

2641 φωτογραφίες για τις οποίες αποδώσαμε ετικέτες χειροκίνητα. Η συλλογή τους πραγματοποιήθηκε με την χρήση joysticks και ρυθμό 3fps, κυρίως στον χώρο που θα πραγματοποιούσαμε τα τελικά μας πειράματα (Αριστερή εικόνα 5), αλλά και εμπλουτίσαμε τα δεδομένα μας ακόμα και με παραδείγματα από μερικούς εσωτερικούς χώρους (Δεξιά εικόνα 5) για να αποφύγουμε τις false positive προβλέψεις.

Επειδή το μέγεθος των δεδομένων μας δεν είναι εξαιρετικά μεγάλο για την εκπαίδευση ενός νευρωνικού από την αρχή, χρησιμοποιήσαμε το pre-trained νευρωνικό από τη βιβλιοθήκη [Tensorflow Models API](#), το [ssd_mobilenet_v2_coco](#). Επίσης, επικαλεστήκαμε και τεχνικές Data Augmentation, με την βοήθεια του python πακέτου [imgaug](#) για να δημιουργήσουμε τεχνητά δεδομένα με βάση τα παραδείγματά μας. Από την μια, με τη χρήση Transfer Learning, παρόλο που το dataset COCO[15] δεν περιέχει Ρομπότ, χρησιμοποιώντας ένα pre-trained μοντέλο του οποίου τα βάρη των αρχικών layers έχουν εκπαιδευτεί να εντοπίζουν γωνίες αποδοτικά, επωφεληθήκαμε από αυτή την ιδιότητα, παρακάμπτοντας την ανάγκη εύρεσης μεγάλου όγκου δεδομένων για την εκπαίδευση αυτών των layers. Από την άλλη, δημιουργήσαμε συνολικά 56242 τεχνητά δεδομένα με Data Augmentation με Affine μετασχηματισμούς ως προς τον x και y άξονα στα κουτιά πρόβλεψης και στις ίδιες τις εικόνες με ποσοστά 20% και 2% αντίστοιχα, ώστε να μην γίνει μεγάλη αλλοίωση της προοπτικής. Για περισσότερες πληροφορίες αναφερθείτε και στο αρχείο [scripts/aug.py](#).

Ακόμα, δημιουργήσαμε και scripts για την διαχείριση των παραπάνω αρχείων, όπως και για την πραγματοποίηση ελέγχου σχετικά με την εγκυρότητα των ορίων των κουτιών σε περιπτώσεις όπου βρίσκονται εκτός της εικόνας, λόγω των παραπάνω μετασχηματισμών.

Ακολούθως, ξεκινήσαμε την συστηματική εκπαίδευση του μοντέλου. Αρχικά, όπως προτείνετε και στην δημοσίευση YOLO9000[11], με k-means clustering, υπολογίσαμε τα κατάλληλα anchors για τα δεδομένα μας. Το clustering αυτό βασίζεται στην μετρική $d(box, centroid) = 1 - IOU(box, centroid)$ η οποία, σε αντίθεση με την Ευκλείδεια απόσταση που προσδίδει μεγαλύτερο error στα μεγαλύτερα κουτιά, είναι κατάλληλη για την εκτίμηση anchors με το καλύτερο δυνατό IOU score, το οποίο είναι ανεξάρτητο από το μέγεθος των κουτιών.

Τρέξαμε τον αλγόριθμο clustering στα πραγματικά μας δεδομένα πολλές φορές για να δημιουργήσουμε σταθερά αποτελέσματα, λόγω της τυχαιότητας του αλγορίθμου που υπάρχει στην αρχική ανάθεση των clusters και συνεπώς στην ευαισθησία αυτών. Η υλοποίηση βρίσκεται στο [scripts/kmeans.py](https://github.com/panagiotis-petropoulakis/scripts/kmeans.py).

Η εκπαίδευση του μοντέλου έγινε στο cloud μέσω της εφαρμογής [Google Colab](https://colab.research.google.com/) (αρχείο [train_collab.ipynb](#)) με μέγεθος εισόδου 300×300 (η επιλογή αυτή έγινε σύμφωνα με ανάγκες μας). Με βάση τα αποτελέσματα στο validation dataset, παρατηρήθηκε ότι το ιδανικό batch size είναι 16 με την καλύτερη σχέση ταχύτητας εκπαίδευσης, overfitting και απόδοσης. Για λόγους εξοικονόμησης ταχύτητας εκπαίδευσης, υιοθετήθηκε ο Adam[16] optimizer, παρόλο που ο SGD ήταν ελάχιστα καλύτερος, για τον οποίο συνιστούμε την χρήση του στο πρόβλημα μας, μιας και συνήθως αποφεύγει τοπικά μέγιστα και πετυχαίνει καλύτερο generalization. Επίσης, εξακριβώσαμε ότι σε pre-trained μοντέλα αρχικά μικρά learning rates και η σταδιακή τους αύξηση έχει τα επιθυμητά αποτελέσματα. Στο δικό μας πειραματικό σενάριο, στο οποίο ο περιβάλλοντας χώρος δεν είναι περιέχει πολλά αντικείμενα, βρήκαμε ότι τα MobileNets με depth multiplier 0.75 έχουν καλύτερη απόδοση από ότι τα μεγαλύτερα και βαθύτερα variants. Ο πηγαίος κώδικας του μοντέλου εντοπισμού μπορεί να βρεθεί στον φάκελο [objectDetector](#).

Το μοντέλο εντοπισμού προτού συνεχίσουμε το δοκιμάσαμε και πειραματικά. Τα αποτελέσματα φαίνονται στον παρακάτω πίνακα και περισσότερες πληροφορίες, όπως και βίντεο, μπορούν να βρεθούν στην πρωτότυπη έκδοση.

Πίνακας 1: Πειραματικά αποτελέσματα συστήματος εντοπισμού

Experiment	Total Frames	False Detections
media/modelEvaluation/1	629	00
media/modelEvaluation/2	2398	09
media/modelEvaluation/3	529	00
media/modelEvaluation/4	1301	00
media/modelEvaluation/5	1604	00
media/modelEvaluation/6	59	01
media/modelEvaluation/7	233	03

4.2 Εκτίμηση της Θέσης του Οδηγού Ρομπότ

Για να εξάγουμε την τρισδιάστατη θέση του Οδηγού Ρομπότ με βάση τις συντεταγμένες εικονοστοιχείων ($x_{min}, y_{min}, x_{max}, y_{max}$) των κουτιών που εκτιμούμε, χρησιμοποιήσαμε τις εξισώσεις από το central perspective imaging model, οι οποίες λαμβάνουν υπόψη τις εσωτερικές παραμέτρους της κάμερας (o_x, o_y, f_x, f_y) σε συνδυασμό με την χρήση της πληροφορίας του βάθους της κάμερας.

Ορίσαμε ως θέση του Οδηγού Ρομπότ τις τρισδιάστατες συντεταγμένες του κέντρου του κουτιού που εκτιμούμε και χρησιμοποιώντας το παραπάνω θεώρημα παίρνουμε τις παρακάτω εξισώσεις:

$${}^cX = \frac{1}{f_x}(x_{center} - o_x)Z \text{ και } {}^cY = \frac{1}{f_y}(y_{center} - o_y)Z.$$

Αξίζει να σημειωθεί ότι δημιουργήσαμε, λόγω της μη μεταβλητότητας των εσωτερικών παραμέτρων, αρχικά ένα map που αποθηκεύσαμε τους σταθερούς υπολογισμούς στην παραπάνω εξίσωση, κατά την οποία ο μόνος άγνωστος είναι το Z , ώστε να κερδίσουμε σε ταχύτητα. Περισσότερα και στα αρχεία [visualServo/helpers.py](#) και [visualServo/visualServo.py](#)

4.3 Εξαγωγή των Ταχυτήτων του Ακόλουθου Ρομπότ

Με βάση τις τρισδιάστατες συντεταγμένες του Οδηγού Ρομπότ για να υπολογίσουμε την γραμμική ταχύτητα που πρέπει να έχει το Ακόλουθο Ρομπότ εφαρμόσαμε τον παρακάτω proportional ελεγκτή:

$$u^* = K_u \sqrt{\Delta_x^2 + \Delta_y^2} - d^*, K_u > 0,$$

με (Δ_x, Δ_y) την σχετική θέση του Οδηγού Ρομπότ ως προς τον Ακόλουθο και $d^* = 1.7$.

Για να στρίψουμε κατάλληλα το Ακόλουθο Ρομπότ προς το κέντρο του Οδηγού Ρομπότ και να υπολογίσουμε την γωνιακή του ταχύτητα χρησιμοποιήσαμε τον παρακάτω proportional ελεγκτή:

$$\omega^* = K_\omega \Delta_{theta}, K_\omega > 0,$$

$$\text{με } \Delta_{\theta} = \arctan\left(\frac{\Delta_y}{\Delta_x}\right).$$

Για την εύρεση των κατάλληλων κερδών (K_v, K_ω) και τον έλεγχο της συμπεριφοράς του ελεγκτή χρησιμοποιήσαμε μέσω του [Gazebo](#), την προσομοίωση [summit_xl_sim](#) που προσφέρει η Robotnik Automation για το Summit-XL Ρομπότ. Στον φάκελο [catkin_ws/.../controller/](#) βρίσκεται η υλοποίηση του ελεγκτή μας καθώς και μιας ρουτίνας που εξαγει “τυχαίες” ταχύτητες για έναν Οδηγό Ρομπότ κατά την προσομοίωση. Σημειώνουμε ότι η κινηματική ανάλυση (τύπου skid-steering) του Summit-XL αναγράφεται στην πρωτότυπη έκδοση και βασίζεται στις παρακάτω δημοσιεύσεις [17-18].

5 ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Αναφέρουμε ότι κατά τη διάρκεια των πειραμάτων χρησιμοποιήθηκαν επιπλέον δύο υπολογιστές. Ο ένας υπολογιστής ήταν τοποθετημένος στο Ακόλουθο Ρομπότ και διέθετε την GPU κάρτα. Σε αυτόν πραγματοποιούνταν όλες οι υπολογιστικές πράξεις, όπου το αποτέλεσμα αυτών, δηλαδή οι ταχύτητες, αποστέλλονταν μέσω Ethernet στο Ακόλουθο Ρομπότ. Ο άλλος υπολογιστής, ήταν απομακρυσμένα συνδεδεμένος, μέσω Wi-Fi, με όλα τα συστήματα με σκοπό την εκκίνησή τους και τον έλεγχο των πειραμάτων. Η παραπάνω επικοινωνία έγινε δυνατή με την χρήση και την κατάλληλη προσαρμογή του [multimaster_fkie](#) ROS πακέτου. Τέλος, στο σύστημα μας δεν λάβαμε υπόψη την ύπαρξη και αποφυγή τυχόν εμποδίων και η κίνηση του Οδηγού πραγματοποιήθηκε με την βοήθεια ενός joystick.



Εικόνα 6: Σχήμα Οδηγός-Ακόλουθος

Ο ρυθμός λειτουργίας του συστήματός μας είναι στα 14-15 Hz με παράλληλη αποθήκευση της πληροφορίας της κάμερας, ενώ χωρίς την διαδικασία της καταγραφής αυτής λειτουργεί στα 19 Hz.

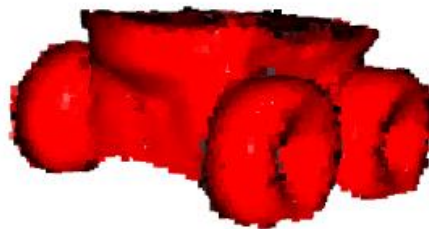
Όπως μπορούμε να παρατηρήσουμε και στον Πίνακα 2, το σύστημα μας έχει τα επιθυμητά αποτελέσματα και διαθέτει μικρό ποσοστό λανθασμένων προβλέψεων. Οι δύο τελευταίες στήλες αναφέρονται στις περιπτώσεις για τις οποίες το σύστημα μας δεν μπόρεσε να εντοπίσει τον Οδηγό Ρομπότ με βάση την RGB πληροφορία και στις περιπτώσεις που ο αισθητήρας βάθους αδυνατούσε να επιτρέψει πληροφορία για το κέντρο του κουτιού, αντιστοίχως. Στον φάκελο experiments μπορούν να βρεθούν και τα βίντεο των πειραμάτων αυτών (Βλ. εικόνα 6).

Πίνακας 2: Πειραματικά αποτελέσματα του σχήματος Οδηγός-Ακόλουθος

Experiment	Total Frames	False Detections	Kinect Failures
media/experiments/1	377	03	00
media/experiments/2	362	13	03
media/experiments/3	247	03	00
media/experiments/4	209	01	00
media/experiments/5	316	09	11
media/experiments/6	205	13	00
media/experiments/7	307	00	15
media/experiments/8	193	04	02

6 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΣΥΖΗΤΗΣΗ

Το σχήμα Οδηγός-Ακόλουθος που προτείνουμε είναι πραγματικού χρόνου με μικρό ποσοστό λανθασμένων οπτικών προβλέψεων. Σχετικά με τις παρατηρήσεις μας, μια από τις αδυναμίες του έγκειται στον ορισμό μας σχετικά με την τρισδιάστατη θέση του Οδηγού Ρομπότ (το κέντρο του κουτιού), επειδή σε ορισμένες περιπτώσεις ο αισθητήρας βάθους αδυνατούσε να επιστρέψει πληροφορία, λόγω θορύβου και αντανάκλασεων που συνέβαιναν όταν το κέντρο αυτό ταυτιζόταν με κάποιο σημείο της οροφής του Οδηγού. Για αυτό τον λόγο, εξέτασα το συγκεκριμένο κομμάτι περαιτέρω για να βρω πως μπορεί το σύστημά μας να βελτιωθεί. Διαπίστωσα ότι μια αποδοτική λύση φαντάζει να είναι η επεξεργασία ολόκληρου το point cloud του Οδηγού Ρομπότ (Εικόνα 7). Επομένως, εξοικειώθηκα από την μεριά μου με την επεξεργασία 3D δεδομένων με την βοήθεια του πακέτου Open3D[19], καθώς και με διάφορες τεχνικές point cloud registration[20-21]. Με αυτές τις μεθοδολογίες είναι εφικτό να βρούμε την πραγματική θέση του Οδηγού καθώς και να υπολογίσουμε και την σχετική περιστροφή των δύο Ρομπότ και να χρησιμοποιήσουμε έναν ελεγκτή διαφορετικής φύσεως. Σημειώνεται ότι για την εξαγωγή ορθών αποτελεσμάτων, συνιστάται η υιοθέτηση τεχνικών Image Segmentation, όπως Mask R-CNN[22], για την καλύτερη δυνατή απομόνωση του point cloud του Οδηγού από τον περιβάλλοντα χώρο και ιδιαίτερα, του πατώματος.



Εικόνα 7: Point cloud Οδηγού Ρομπότ

Ως τελικά σχόλια, μπορούμε να αναφέρουμε ότι πέρα από point cloud registration τεχνικές, πλέον οι 6D pose estimation μεθοδολογίες πετυχαίνουν εξαιρετικά αποτελέσματα για την εύρεση της θέσης αντικειμένων από μια φωτογραφία. Ιδιαίτερα, στην δημοσίευση [23], για την συλλογή δεδομένων και απόδοση ετικετών, αναφέρεται ακόμα ότι σε δύσκολες περιπτώσεις συλλογής

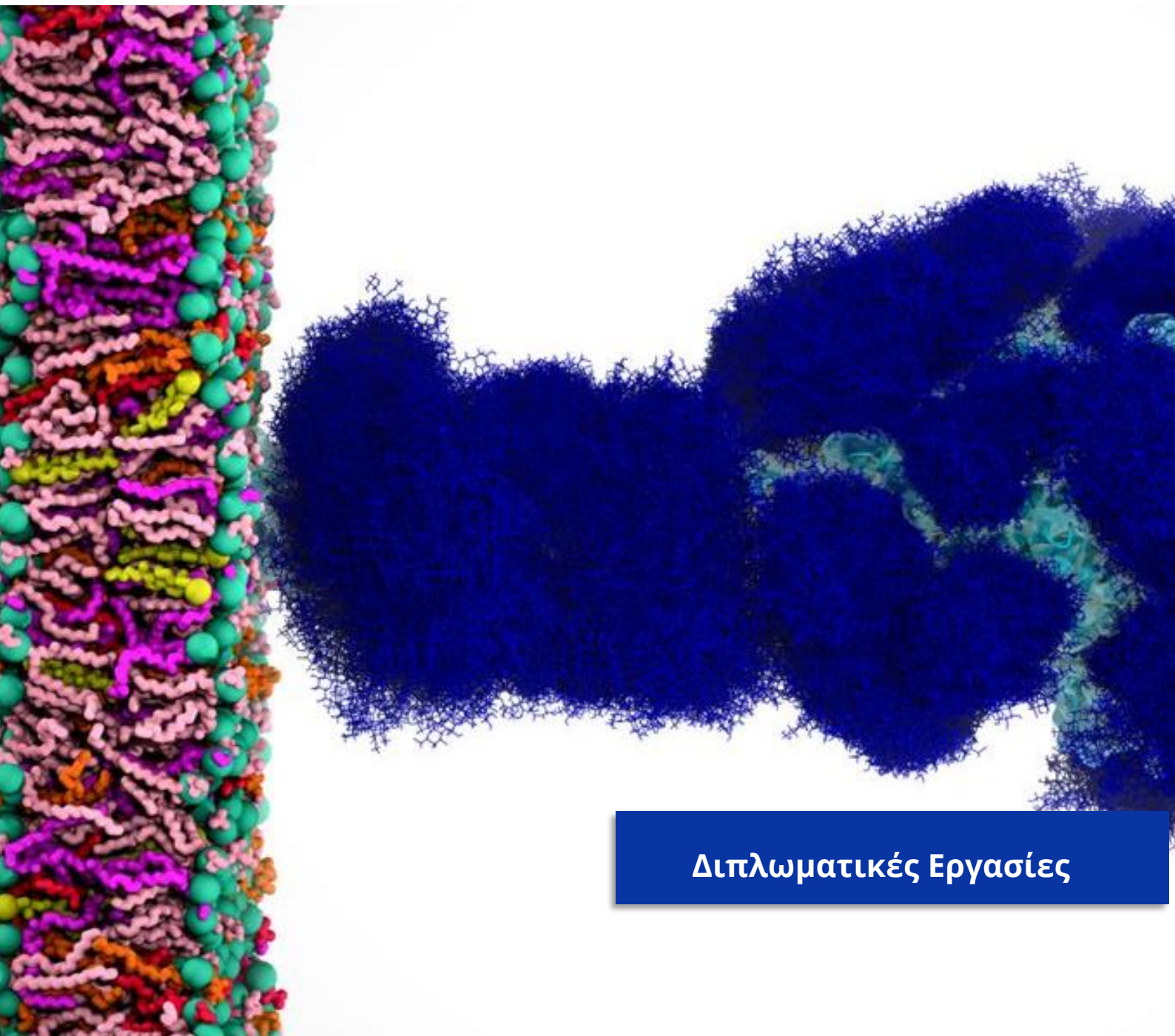
τους, για παράδειγμα σε υποβρύχια Ρομπότ, προτείνεται αρχικά η δημιουργία τεχνητών δεδομένων με ετικέτες μέσω προσομοίωσης και στην συνέχεια η πραγματοποίηση image translation, με την βοήθεια του μοντέλου CycleGAN[24], στα τεχνητά αυτά παραδείγματα με σκοπό την μετατροπή τους σε μορφή όσο πιο ρεαλιστική γίνεται. Με διαφορετικό τρόπο, στην [25] εργασία, εκπαιδεύσανε ένα CNN μοντέλο για να εκτιμήσουν την σχετική θέση της κάμερας χρησιμοποιώντας σαν είσοδο δύο εικόνες, και στην συνέχεια, με την προσθήκη position-based visual servo ελεγκτή, πραγματοποίησαν πειράματα πάνω σε Drone. Από την άλλη, αν βρεθούμε σε ένα δωμάτιο με πολλά πανομοιότυπα Ρομπότ έγκειται το ερώτημα το πως τα ξεχωρίζουμε. Πάνω σε αυτό το ερώτημα, έρχονται να απαντήσουν τεχνικές Object Tracking, όπως η [26]. Τέλος, μια πολλά υποσχόμενη δημοσίευση, η [27], συνδυάζει Deep Reinforcement Learning μεθοδολογίες για την επίλυση του image based visual servoing προβλήματος, κατά την οποία, μαθαίνονται τα dynamics των οπτικών στοιχείων (features) και στην συνέχεια, με Q-Learning, εκτιμώνται οι ταχύτητες που πρέπει να έχει ένα Drone για να ακολουθήσει ένα Κινούμενο Ρομπότ.

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

- [1] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in CVPR, 2018.
- [2] F. Chaumette and S. Hutchinson, "Visual servo control, Part I: Basic approaches," in IEEE Robotics and Automation Magazine, 2007.
- [3] F. Chaumette and S. Hutchinson, "Visual servo control, Part II: Advanced approaches," in AISTATS, 2007.
- [4] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded up robust features," in European Conference on Computer Vision, 2006.
- [5] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An Efficient Alternative to SIFT or SURF," in ICCV, 2011.
- [6] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in CVPR, 2014.
- [7] R. Girshick, "Fast R-CNN," in ICCV, 2015.

- [8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in ICCV, 2015.
- [9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in CVPR, 2016.
- [10] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "SSD: Single shot multibox detector," in European Conference on Computer Vision, 2016.
- [11] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in CVPR, 2017.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in CVPR, 2016.
- [13] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenk, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," in CVPR, 2017.
- [14] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques, "Leader-Follower Formation Control of Non-holonomic Mobile Robots with Input Constraints," in Automatica, 2008.
- [15] T. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick, "Microsoft COCO: Common Objects in Context," in European Conference on Computer Vision, 2014.
- [16] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization," in International Conference on Learning Representations, 2014.
- [17] K. Kozłowski and D. Pazderski, "MODELING AND CONTROL OF A 4-WHEEL SKID-STEERING MOBILE ROBOT," in International Journal of Applied Mathematics and Computer Science, 2004.
- [18] L. Caracciolo, A. D. Luca, and S. Iannitti, "Trajectory tracking control of a four-wheel differentially driven mobile robot," in IEEE Int. Conf. Robotics and Automation, Detroit, 1999.49
- [19] Q. Zhou, J. Park, and V. Koltun, "Open3D: A Modern Library for 3D Data Processing," arXiv:1801.09847, 2018.
- [20] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in Digital Imaging and Modeling, 2001.
- [21] Q. Zhou, J. Park, and V. Koltun, "ORB: An Efficient Alternative to SIFT or SURF," in ECCV, 2016.
- [22] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in ICCV, 2017.

- [23] B. Joshi, M. Modasshir, T. Manderson, H. Damron, M. Xanthidis, A. Q. Li, I. Rekleitis, and G. Dudek, "DeepURL: Deep Pose Estimation Framework for Underwater Relative Localization," arXiv:2003.05523, 2020.
- [24] J. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," in ICCV, 2017.
- [25] A. Saxena, H. Pandya, G. Kumar, A. Gaud, and K. M. Krishna, "Exploring Convolutional Networks for End-to-End Visual Servoing," in ICRA, 2017.
- [26] Q. Wang, L. Zhang, L. Bertinetto, W. Hu, and P. H. Torr, "Fast Online Object Tracking and Segmentation: A Unifying Approach," in CVPR, 2019.
- [27] A. X. Lee, S. Levine, and P. Abbeel, "Learning Visual Servoing with Deep Features and Fitted Q-Iteration," in ICLR, 2017.



Διπλωματικές Εργασίες

A Mixed Reality Dashboard Based on a Distributed Data Streaming Architecture

Deligiannakis N. Nektarios

ABSTRACT

As computer science grows and evolves, new technologies come to surface. One of these is called Mixed Reality. However, the creation of applications using MR technologies is very difficult, due to the high cost of the MR devices and the small developer community that insists on innovating in Mixed Reality applications.

We present a Mixed Reality Dashboard that receives live data streams from a distributed platform, Apache Kafka and Zookeeper, and demonstrates functionalities that can be used to monitor and visualize live data from any sensor or device. Data visualization is very important making large data sets easier to understand.

The application consists of a complex User Interface. That makes use of Application Programming Interfaces (API) and the Kafka Rest Proxy, for receiving and visualizing data to users. resulting in a greatly extensible application.

Keywords: augmented reality, mixed reality, data visualization, HoloLens, data graph, REST, Apache Kafka, REST Proxy, User Interface, Application Programming Interface, Kafka and Zookeeper.

SUPERVISOR

Stathes P. Hadjiefthymyades, Professor NKUA

1 INTRODUCTION

In this thesis we have created a Mixed Reality Dashboard deployed on the HoloLens 1 device. This dashboard presents various and basic functionalities that are able to handle live data in various formats from a network. The data types that are currently supported are data in the form of JSON data type and video feed from the internet.

The data stream may be measurements from sensors or Unmanned Vehicle (UxV) camera devices. In order for the device to get the information from a network, we used Apache Kafka, a distributed streaming platform. To demonstrate the whole architecture, we have designed a User Interface that is needed to project the information in a mixed reality context. Also, we used an already set up server, used from our research lab, that runs Apache Kafka and Zookeeper, in order to demonstrate the data exchange and communication with the application on the HoloLens 1.

Various difficulties and problems came up, during the development of this application. Many trials and fails had to happen in order to develop and create a stable and reliable version of the above application. We have to consider that the application development for HoloLens 1 is not fully supported due to the fact that the new HoloLens 2 device is now available.

The tools used to design, create and develop our dashboard are the Unity Editor 2018 LTS[1], the Mixed Reality Toolkit (MRTK)[2][6], Microsoft Visual Studio, .NET Framework 4.5 - 4.6 in C and some Python for creating some testing scripts. The data stream may be measurements from sensors or UxV camera devices. In order for the device to get the information from a network, we used Apache Kafka, a distributed streaming platform. To demonstrate the whole architecture, we have designed a User Interface that is needed to project the information in a mixed reality context. Also, we used an already set up server, used from our research lab, that runs Apache Kafka and Zookeeper, in order to demonstrate the data exchange and communication with the application on the HoloLens 1 in a distributed manner [11].

Finally, we have to mention that at the current time of the publication of this thesis, no related work has been done, in designing and developing such a dashboard. Therefore, the challenges that aroused where numerous.

2 MIXED REALITY FEATURES

In our application, we do project holograms to the user. We use 2D objects for our UI and it is possible to add any hologram to our application. On HoloLens however, because it is a Mixed Reality device, holograms blend with your physical environment to look and sound like they're part of our world. Even when holograms are all around you, you can still see your surroundings, move freely, and interact with the physical environment. We call this experience “mixed reality” [3].

HoloLens 1 is equipped with 4 environment understanding cameras, an RGB camera and a depth camera. Spatial Mapping is a process that allows HoloLens to understand its environment and surroundings connecting an object’s position and orientation from the virtual to the real world [5]. This can be achieved with detecting common characteristics of the real world like the position and orientation of surfaces and the real-world lighting conditions.

Getting around HoloLens is a lot like using your smart phone. You can use your hands to manipulate holographic windows, menus, and buttons. Instead of pointing, clicking, or tapping, you'll use your gaze, your voice, and hand gestures to select apps and holograms and to get around HoloLens.

Our device, HoloLens 1 have a gesture frame which is an area the device can detect the movements of the hand in front of it. In order to enable any hand gesture capabilities in our applications we had to enable them through the Mixed Reality Toolkit. This toolkit offers the option to adopt any default hand gestures but also create new ones. It is preferred to alter the existing profiles as we are not certain yet that integrating new hand gestures works the optimum way. Therefore, keeping up with the updates from the toolkit is essential [2].

The gesture shown below, is the main interaction hand movement that our application understands and tracks highly efficiently. The user has to air tap on the component he desires to zoom by tapping on it. Each object has its own resize script as the operations to the unity objects transform that have to be performed to each component for are position depended. Also, we needed to add some functionality to make the UI follow the users view. Without it, the user would lose sight of the UI if he committed to any type of movement. This would happen because the anchors of the UI would not change when the user moves. This is important in order to make the UI visible to the user even if the user has moved his head around.

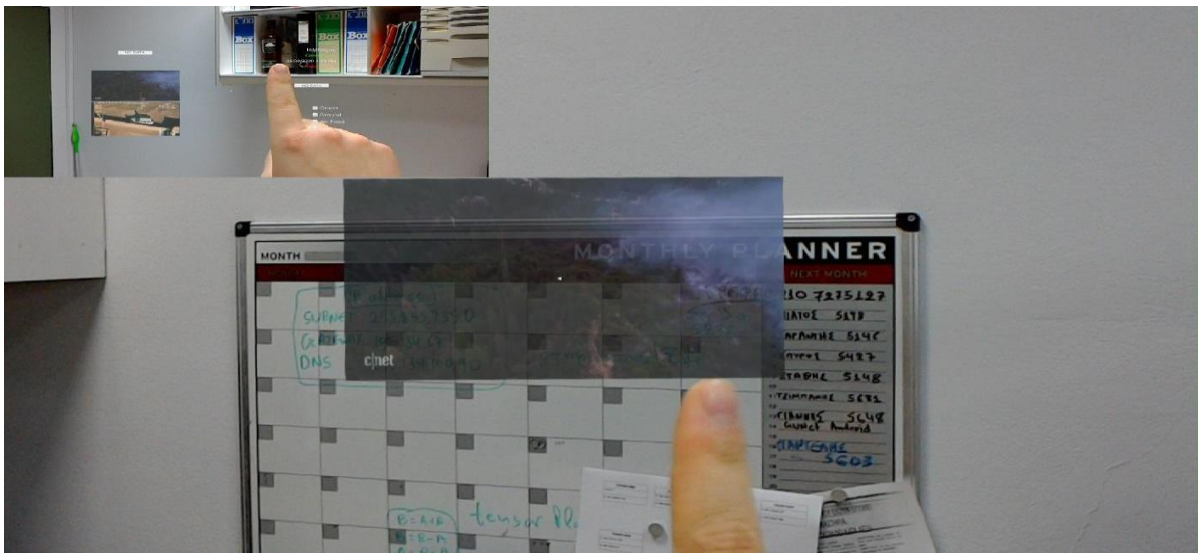


Figure 1: By air-tapping on one of the components the component comes to the center of the users view and gets larger

3 GENERAL OVERVIEW

In the above image we give a general overview of the applications architecture. All the various components are coordinated by the application deployed to the HoloLens device. The User Interface is a Panel that consists of nine parts (or panels) equally drawn on the main Panel. Each panel is an independent panel with specific functionality. The main Panel consists of the listed sub Panels: Data Graphs, Video Graphs and Options Pane.

The User Interface has some features enabled that offer a better visualization for the user. Tooltips over the data are also provided, giving an even better understanding of the data to the user. Also, the UI has specific scripts attached to it, in order to follow the user's view through the device. This is very necessary because the devices spatial awareness is enabled and helps the user to keep track of the UI and its position in space.

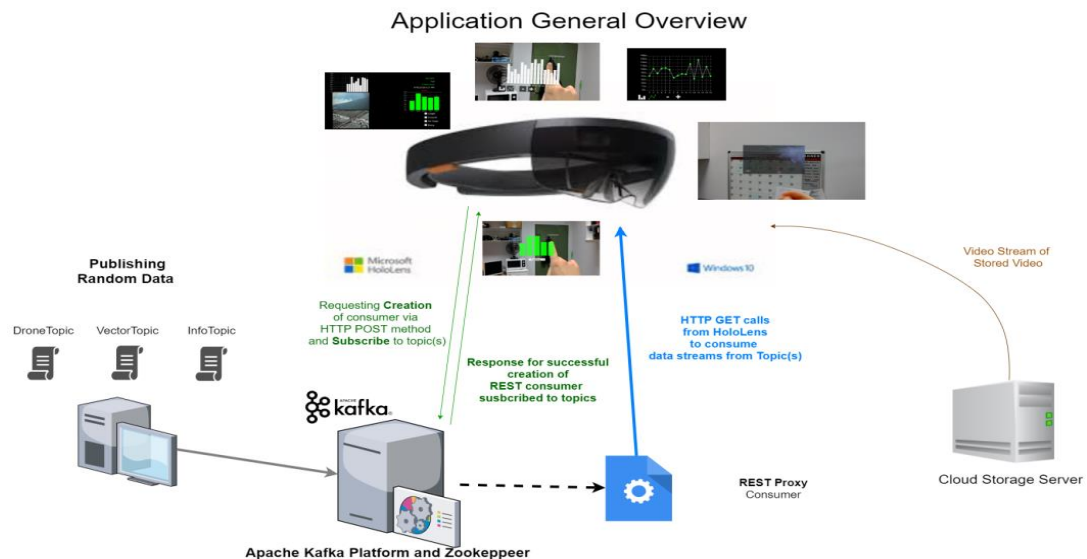


Figure 2: General Overview of the application

4 DATA GRAPHS

We will start with the Data Graph panel and the graphs it consists of which are the Simple Graph and the Dynamic Graph. This panel, a numeric graph is created and designed to visualize numeric values for a specific phenomenon. **The Simple Graph** in this example holds values of altitude simulating the altitude of a drone. The data that is projected on the graph is very easy to be configured or altered in the Unity editor, in order to visualize any phenomenon, the user needs. The co-routine receives data if available and translates it to a form usable for the graph components. Upon successful reception of messages, the graph gets updated and redesigned (if needed), visualizing the new incoming live data.

The graph component also offers some options in order to change the way the graph visualizes the data. The user can choose between a bar chart graph and a

line graph. Also, the user may wish to monitor a subset of the data in the graph. As we will also discuss in a next section this component receives data from the altitude Kafka topic. When receiving new value, a list of the values is updated and re-loaded to the graph component. Also, we have a sliding window as we can call it, which allows us to monitor only the last x values we received. The value x is a value that we can set from the editor and temporarily change from the application itself. The graph obviously gets updated with the same rate the values are received if a new value is available. In case the values are constantly changing, the graph gets rebuild.

This sliding window is very important and could also be used for other purposes like applying some data algorithm on this selected window. For instance, we could apply sliding window algorithms at the already created window of values. In many cases, it is possible to have values with a specific type of measurement, like meters or feet for distance. This feature is also available from the editor and from the Graph menu to select. However, any change might have a temporary effect on the visualization of the data, meaning that the graph opts for the preset options.

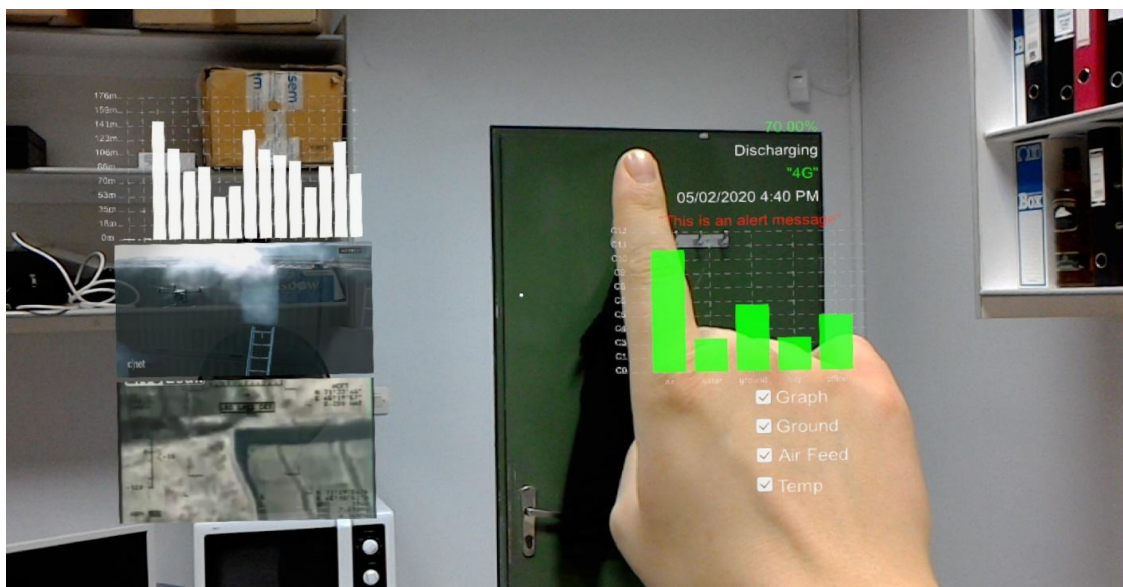


Figure 3: Application view from the HoloLens device

The Dynamic Graph component is similar to the above Graph component. However, it has features that make it more dynamic to the received data. First of all, this dynamic graph is a bar chart. Each bar represents the last value of a set of values for different phenomena e.g. we visualize temperatures for different subjects, like temperature in the air, water, ground etc. In case new sensors or sources of data start to send data to our application, the graphs gets dynamically redesigned and reformed in order to project the new data.

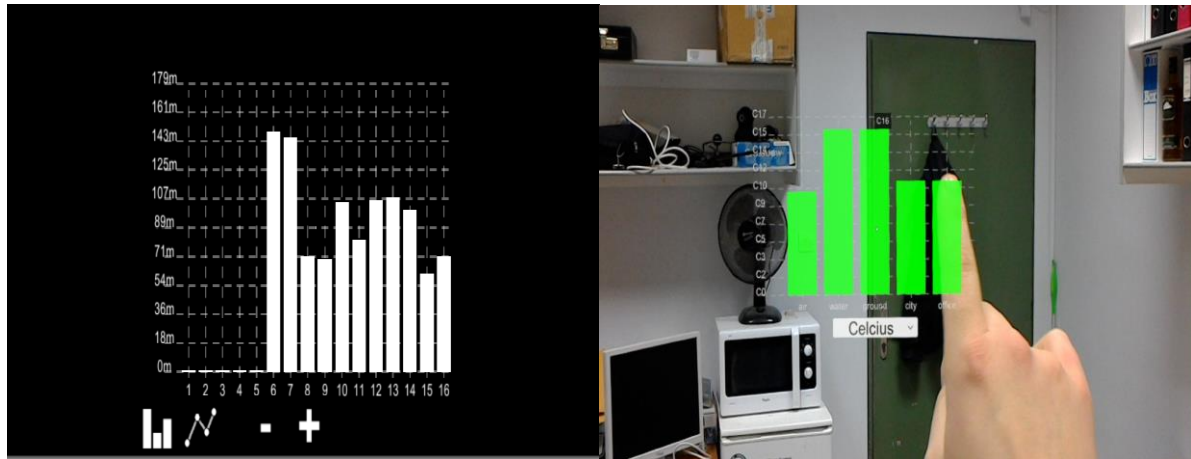


Figure 4 and 5: Zoom in and Sliding Window, and the settings available

5 VIDEO GRAPHS

Video graphs are created to offer live video feed to the user. Currently, we consume video feed from data servers, in order to measure any performance issues. It is possible to consume video frames from a Kafka topic and construct the video in the device but this aspect is on test mode at this time. Each video component has the zoom in or zoom out function with the use of air tap on the video. The two components are the Air Feed and the Ground Feed components.

The **Air Feed** component is intended to grab a live video feed from a topic and project it to the user. We named Air Feed to emphasize the use case scenario of a drone flying above an area and feeding the Kafka topic with its on board camera feed. However, we stumbled upon many and various difficulties and limitations from the tools we used. For security reasons, it is not allowed to consume any

data type from the internet. Also, not any type of web operations with data streams is allowed because of the just mentioned reason.

We achieved to stream video from a server and optimize it in order to avoid conflicts between the other internet transactions with internet (from the graphs etc.). We have stored a video in a cloud service, and used the share link that the platform provided and used it as source for our video component. From there, the video player script attached to the component, takes the URL and starts downloading frames. If the frames that are already streamed suffice for a smooth video play, the video starts to play and the component continues to download frames. Similar to the Air Feed component, the **Ground Feed** video component is also intended to consume live video feed from the network. The user again is able to enlarge the video by air tapping on it.

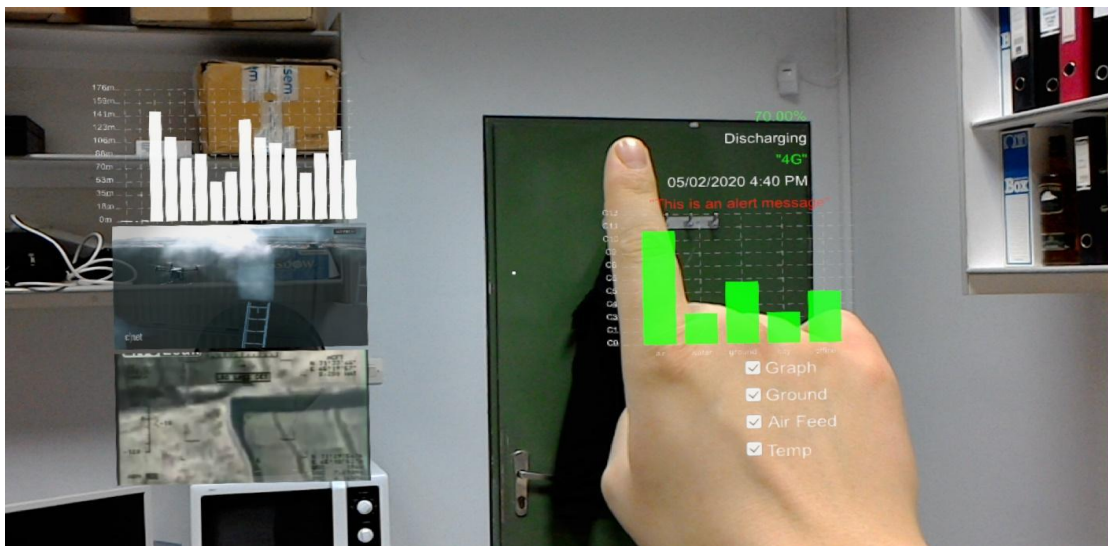


Figure 6: View of the application with use of the received data streams. Both video feed are visible at the left side of the dashboard.

It is necessary to point out that only with the above three graphical components, we already strain a lot the device with great workload on the CPU, the GPU that has to render all changes in the frames and to the device's memory. Multiple operations execute behind the scenes and therefore is not possible to completely control the workload of some processes. With many tests we found that the consumption of frames needs some effort rendering our UI components as expected.

5.1 Options Section

The Options Pane component holds more options regarding the entire layout and the components. Currently, the user is able to enable/disable a component. This is made in order to help the user to monitor the data he desires or currently needs. Also, we have to consider that the users view is limited towards his real surrounding space. These options intent to help the user free up his view through the device's view and move safely in his real surroundings.

6 SCRIPTS AND SCRIPTING APIS

In order to create our application and also implement some logic we had to write some code. For the scripts we have created we used the C# programming language and some scripting APIs. We used the Unity scripting API which offers the documentation and some basic examples for the use of functions and classes Unity offers. This scripting API is the core of our application and there were no alternatives to use. We used the .NET framework from Microsoft that was used to build and deploy our application to the HoloLens [8].

Co-routines are a type of functions but with some more and very important characteristics. In general, functions are called and executed inside the lifecycle of a frame. This means that a function has to run to its completion and this had to happen in a single frame. Also, many times it is needed to fulfill a task that is going to take multiple frames to complete. We could spread this work across multiple frames with the use of the Update function but this solution will certainly backfire on us with performance issues. Co-routines are like functions but have the ability to pause their execution and return the control to the main thread of the application, waiting for synchronous input, etc.

7 DATA CONSUMPTION

The above section in the way and the methodology our application consumes data from the network. We have to point out, that the HoloLens device and most Augmented Reality devices, have not yet offered robust solutions for consuming live data streams. In order to visualize data in the Mixed Reality Dashboard application, we have to generate data in a specific format and then create the application's routines for consuming the data from the Kafka REST Proxy.

First of all, the data format has to be JSON. For demo purposes we have created our own producers written in Python, which create JSON Objects and send them to a specific Kafka Topic. The application's routines consume the data from each Topic (one Topic per UI component), strip it from any metadata that Kafka adds to the message (for configuration purposes) and parse it in order to extract the specific values we need for our graph.

The above procedure is simple but has to be done as dynamic as possible in order to make the UI support as generic data types as possible. Our application achieves the above assumption by parsing the data in a general manner without having hardcoded information on our graph components. For instance, the presentation of the values in the graph uses the key values inside the received JSON objects for the tooltips and properties of the graphs. Additionally, the dynamic graph has no pre-configured values or properties and is able to dynamically visualize data in relation to the received data parsed from the JSON object. It automatically adjusts properties of the graph, like names of the datasets visualized.

The video feed components require a video stream to project to the user. These components have mechanisms enabled in order to prepare the video to be played, by gathering some frames before playing, and also to handle possible audio sources and other quality properties. However, the application limits the type of video stream that can be received and projected. This is a great challenge due to the fact that performance may degrade significantly on the device, in the process of not only rendering holograms but also decoding and rendering a video stream. Furthermore, our goal is to consume video feed from a Kafka topic,

which is currently on a testing phase and has great developing difficulties. We have proved that it is possible and may offer a smooth and reliable experience to the user.

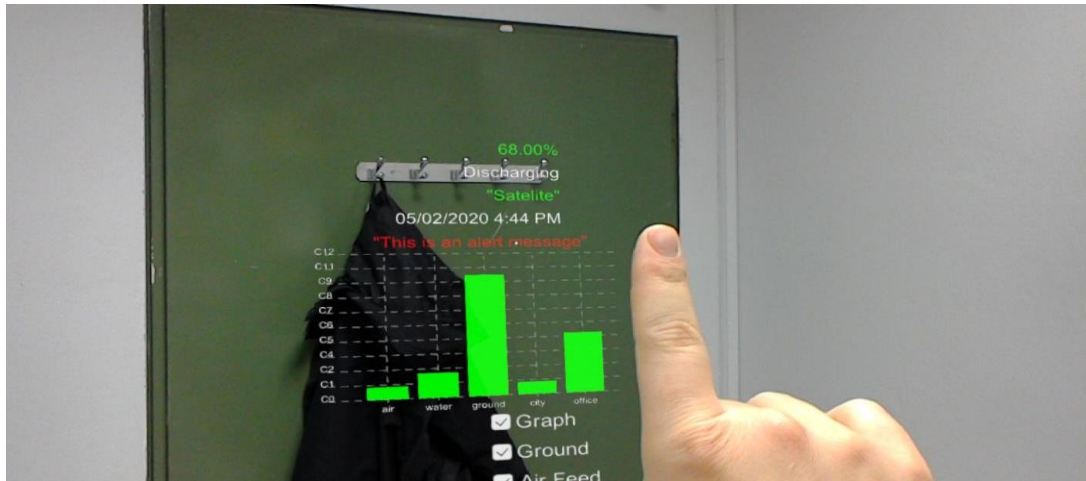


Figure 7: The message section

Also, we have to mention that this section of the UI has the potential to make use of the reception of any kind of events with use of a panel dedicated for status information and received messages or alerts. The user may be in a position to contact someone and receive or publish messages to the Kafka services that will be able to be consumed from another application.

8 APACHE KAFKA PLATFORM

A very challenging aspect of the application we present, is the communication with a network using the Apache Kafka [10] distributed streaming platform with help of the Apache Zookeeper. Kafka implements queues that are able to handle events very efficiently and fast. In our application we used the Confluent Platform, which implements at its core Apache Kafka and is the only enterprise event streaming platform as of now. The Confluent Platform is almost fully documented and offers many different **connectors** that offer more application specific automations for handling events. Apache Zookeeper is an open-source server which enables highly reliable distributed coordination.

An application has to subscribe to a **Topic** in order to publish or consume messages in real time to any other application. Kafka guarantees some very important assumptions with the most important that messages will be consumed in the order they were published to the topic. Topics are unique named buses, used for data transmission between nodes. Nodes can publish or subscribe to a topic, but data production and consumption are decoupled meaning that there is an indirect connection between the nodes. We have to point out that, implementing the above assumptions is not a feature that the device was designed for. Therefore, our application is a prototype at this kind of event tracking via the Kafka Confluent Platform.

In order to produce or consume data from a server that executes the Apache Kafka streaming platform, some configuration has to happen first. Our application is limited by the .NET framework, on how to connect and get data from a network. First of all, we have to use the UnityWebRequest object provided from the Unity Scripting API. This object is used to implement many types of connections to web servers and is the only one that worked on our HoloLens device. Despite this constraint we kept pushing for a solution, as Apache Kafka is a robust distributed platform for big data processing, significantly expanding our application's potential use cases [11].

The UnityWebRequest object, limited us to use the Apache Kafka REST Proxy connectors, as we could not use the Confluent Kafka libraries offered on many other programming languages, challenging us even more. In simple words, our application does not load any Kafka libraries or dll files but uses UnityWebRequest objects to connect to a specific URI. The above URI requires some steps to be created as it has to refer to a specific topic(s), consuming the data in a way to handle them as a simple JSON format. Each message received from Kafka, has our required values encapsulated (the sensor data is bolded and the rest are metadata useful for configuration purposes).

*Message produced: {"altitude": **104**}*

Message received:

```
[{"topic":"DroneTopic","key":
null,"value":{"altitude":104},"partition":0,"offset":80752}]
```

As we can see, the initial message is encapsulated in JSON data message consumed from the Kafka topic "DroneTopic". The above message holds some very basic values like the partition and the message offset. These values can be used in order to recover from a network failures or any other cause that forced a delay or halt on consuming data from the application. In our application we opted for the current and latest messages (or values) in order to keep track of a phenomenon in real time, ignoring any previous data that might be stored. The procedure for this assumption is described in the previous section.

We would also like to mention the other two topics we are using which are the Vector topic and the InfoTopic. The Vector topic is used to present the dynamic loading of content from the graphs. The producer of the message may alter the dataset he wants to visualize to the receiver. If this is happening the application will adapt to these changes and the graph will take the form of the new received content. The InfoTopic is used to demonstrate the reception of messages to the application user.

9 CONCLUSIONS

The presented application offers an integrated solution for visualizing live data streams in an Augmented environment. The technologies used consist of tools and APIs that are the state of the art. We used tools and libraries that are constantly evolving. Although this came with many challenges, we succeeded to overcome the numerous difficulties occurred by the innovative nature of our application. We have demonstrated a use case of an Augmented Reality application, using the AR device HoloLens and visualizing data consumed from a robust and concrete distributed live data stream platform.

REFERENCES

- [1] Unity Editor, <https://unity.com/products/core-platform>
- [2] Mixed Reality Toolkit, <https://github.com/microsoft/MixedRealityToolkit-Unity>
- [3] How Augmented Reality Works
<https://computer.howstuffworks.com/augmented-reality.htm>
- [4] VR, AR, MR, and What Does Immersion Actually Mean?,
https://www.thinkwithgoogle.com/intl/en-gee/insights-trends/industry-perspectives/vr-ar-mr-and-what-does-immersion-actually-mean/?gclid=CjwKCAiAvonyBRB7EiwAadauqTgTO-BWnLavH7O0Zpg7a6Wfd9VDpsD6WbCa7e2xMI1bQIycxCK2rxoCdkgQAvD_BwE
- [5] HoloLens Spatial Mapping – A developers guide
<https://lightbuzz.com/hololens-spatial-mapping/>
- [6] How to develop for Microsoft HoloLens
<https://www.hypergridbusiness.com/2017/01/how-to-develop-for-microsoft-hololens/>
- [7] HoloLens for Research, <https://www.microsoft.com/en-us/research/academic-program/hololens-for-research/>
- [8] MR Basics 100: Getting started with Unity, <https://docs.microsoft.com/en-us/windows/mixed-reality/holograms-100>
- [9] UnityWebRequest,
<https://docs.unity3d.com/ScriptReference/Networking.UnityWebRequest.html>
- [10] What is Confluent Platform?
<https://docs.confluent.io/current/platform.html>
- [11] Cloud Computing: Theory and Practice 1st Edition by Dan C. Marinescu
<https://eclass.uoa.gr/modules/document/file.php/D416/CloudComputingTheoryAndPractice.pdf>

Αξιοποίηση της πλατφόρμας Arduino στην εκπαίδευση: Σχεδιασμός μαθησιακών δραστηριοτήτων με βάση το πλαίσιο ECLiP

Ιορδάνης Δ. Κούσης

ΠΕΡΙΛΗΨΗ

Η εκπαιδευτική ρομποτική εισάγεται ολοένα και περισσότερο σε όλες τις βαθμίδες της εκπαίδευσης και ιδιαίτερα σε εκπαιδευτικά συστήματα που ακολουθούν την εκπαίδευση STEM. Η εργασία προτείνει την αξιοποίηση της πλατφόρμας Arduino για την εισαγωγή της εκπαιδευτικής ρομποτικής στην πρωτοβάθμια και δευτεροβάθμια εκπαίδευση. Συγκεκριμένα, προτείνεται ο σχεδιασμός μαθησιακών δραστηριοτήτων βάσει του πλαισίου σχεδιασμού μαθησιακών δραστηριοτήτων ECLiP. Η πρόταση περιλαμβάνει τον εμπλουτισμό του ECLiP με εκπαιδευτικά στοιχεία ώστε να μπορεί να εφαρμοστεί σε δραστηριότητες εκπαιδευτικής ρομποτικής και την αξιοποίηση εναλλακτικών προγραμματιστικών περιβαλλόντων για τον προγραμματισμό του Arduino ώστε να καλύπτει διαφορετικές μαθησιακές ανάγκες.

Λέξεις κλειδιά: Εκπαιδευτική Ρομποτική, Πληροφορική, Arduino, ECLiP

ΕΠΙΒΛΕΠΟΝΤΕΣ

Αγορίτσα Γόγουλου, μέλος ΕΔΙΠ

1 ΕΙΣΑΓΩΓΗ

Η ραγδαία τεχνολογική εξέλιξη που έχει συντελεστεί τα τελευταία χρόνια σε συνδυασμό με το γεγονός ότι αυτή συνεχίζεται με αμείωτο ρυθμό έχει φέρει στο προσκήνιο τη συζήτηση που αφορά το μέλλον των επαγγελματιών (Nübler, 2016), των δεξιοτήτων και φυσικά της εκπαίδευσης που θα πρέπει προηγείται. Δεξιότητες, όπως κριτική σκέψη, επίλυση προβλημάτων, υπολογιστική σκέψη, συνεργασία, αποτελούν το επίκεντρο της εκπαιδευτικής μεταρρύθμισης σε πολλές χώρες. Προς αυτή την κατεύθυνση η αξιοποίηση των ΤΠΕ και της εκπαιδευτικής ρομποτικής μπορούν να συμβάλλουν σημαντικά (Eguchi, 2014).

Πρόσφατες ερευνητικές εργασίες αναδεικνύουν ότι η εκπαιδευτική ρομποτική έχει προσελκύσει το ενδιαφέρον και την προσοχή των ερευνητών ως ένα χρήσιμο εργαλείο για την ανάπτυξη γνωστικών και κοινωνικών δεξιοτήτων στους μαθητές, ειδικά στα γνωστικά αντικείμενα STEM - Science, Technology, Engineering, Mathematics (Alimisis, 2013; He & Liang, 2019).

Η εκπαιδευτική ρομποτική, σχετίζεται τόσο με την παρατήρηση ή τον απλό χειρισμό ηλεκτρονικών μοντέλων (ρομπότ σε κλίμακα) όσο και με την εμπλοκή του μαθητή με τον προγραμματισμό, την επίλυση προβλήματος και τη λήψη αποφάσεων σε σχέση με τη συμπεριφορά του μοντέλου με στόχο την ανάπτυξη της δημιουργικής σκέψης. Η ένταξη της ρομποτικής στην εκπαίδευση μπορεί να θεμελιωθεί στο επιχείρημα της χρήσης της τεχνολογίας ως νοητικό εργαλείο (Mikropoulos & Bellou, 2013). Η σχεδίαση μαθησιακών δραστηριοτήτων υπό τη μορφή των συνθετικών εργασιών διευκολύνει την προσαρμογή στα ενδιαφέροντα και στις δεξιότητες των μαθητών (Φράγκου κ.α., 2010).

Υπό κατάλληλες προϋποθέσεις, η εκπαιδευτική ρομποτική μπορεί να χρησιμοποιηθεί από τον μαθητή των πρώτων τάξεων του Δημοτικού, έως και τον φοιτητή ή τον επιστήμονα που εμπλέκεται με σύνθετες καταστάσεις και γλώσσες προγραμματισμού για τον καθορισμό της συμπεριφοράς του ρομπότ (He & Liang, 2019). Καθώς στην αγορά υπάρχουν ποικίλα συστήματα ρομποτικής, τα οποία διαφέρουν ως προς το κόστος, τα κατασκευαστικά χαρακτηριστικά, τα δομικά υλικά, τον βαθμό προσυναρμολόγησης τμημάτων και τον προγραμματισμό τους, είναι προφανές ότι χρειάζεται προσοχή ώστε να

επιλέγεται εκείνο το περιβάλλον υλικού και λογισμικού που ικανοποιεί τις απαιτήσεις της βαθμίδας για την οποία προορίζεται καθώς και τα επιδιωκόμενα μαθησιακά αποτελέσματα. Οι σημαντικότερες παράμετροι που συνεκτιμώνται για την κατάλληλη επιλογή και ταυτόχρονα οι προκλήσεις για την αξιοποίηση της ρομποτικής είναι (Schön, Ebner & Kumar, 2014; Vandavelde et al., 2013)

- Το κόστος
- Ο βαθμός προσυναρμολόγησης τμημάτων
- Το είδος προγραμματιστικού περιβάλλοντος (γλώσσες προγραμματισμού, λειτουργικά, διαθέσιμα συστήματα)
- Ο διαθέσιμος χρόνος διδασκαλίας
- Ο βαθμός κατάρτισης του εκπαιδευτικού

Τα συστήματα εκπαιδευτικής ρομποτικής χωρίζονται σε δύο κατηγορίες: αυτή των «κλειστών» συστημάτων και των εμπορικών πακέτων όπου η ελευθερία δημιουργίας περιορίζεται από τις δυνατότητες και τα τυχόν εξαρτήματα που παρέχει ο κατασκευαστής (π.χ. Lego Mindstorms, Fischer Technik, Edison, Vex Robotics) και αυτή των «ανοικτών» συστημάτων όπου ο καθένας μπορεί να δημιουργήσει ό,τι θέλει χωρίς περιορισμούς (π.χ. Arduino, Raspberry Pi). Τα «κλειστά» πακέτα έχουν υψηλό κόστος αγοράς ενώ τα «ανοικτά» συστήματα απαιτούν μεγαλύτερο κόπο (και ενδεχομένως χρόνο) δημιουργίας.

Το θεωρητικό υπόβαθρο και οι βασικές γνώσεις που πρέπει να έχουν οι μαθητές προκειμένου να ανταπεξέλθουν σε ένα εκπαιδευτικό πρόγραμμα ρομποτικής ποικίλλουν καθώς εξαρτώνται από παράγοντες όπως το σύστημα που θα χρησιμοποιηθεί και η ηλικία των μαθητών. Συγκεκριμένα για το Arduino απαιτούνται βασικές γνώσεις υπολογιστή, γνώσεις περιήγησης στο Διαδίκτυο, βασικές γνώσεις σε αρχές προγραμματισμού και γνώσεις ηλεκτρολογικών /ηλεκτρονικών κυκλωμάτων. Για κατηγορίες μαθητών που δε διαθέτουν το υπόβαθρο σε προγραμματισμό C και σε θέματα βασικής ηλεκτρονικής /ηλεκτρολογίας (όπως σε μαθητές δημοτικού) μπορεί να ακολουθηθεί η προσέγγιση με την αξιοποίηση οπτικών γλωσσών (Scratch) για τον

προγραμματισμό και χρήση εισαγωγικών ενοτήτων για τα ηλεκτρονικά εξαρτήματα που θα χρησιμοποιούνται.

Στο πλαίσιο της παρούσας εργασίας προτείνεται η αξιοποίηση και ο εμπλουτισμός του πλαισίου ECLiP (Exploratory + Collaborative Learning in Programming) (Gogoulou, Gouli & Grigoriadou, 2009) στον σχεδιασμό μαθησιακών δραστηριοτήτων για την εισαγωγή μαθητών της πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης στην εκπαιδευτική ρομποτική αξιοποιώντας την πλατφόρμα Arduino. Επιπλέον, παρουσιάζεται ένα σύνολο μαθησιακών δραστηριοτήτων που καλύπτει βασικές έννοιες της εκπαιδευτικής ρομποτικής καθώς και τα αποτελέσματα από την πιλοτική πειραματική εφαρμογή.

2 ECLIP ΚΑΙ ARDUINO

Το ECLiP είναι ένα πλαίσιο σχεδιασμού δραστηριοτήτων που αξιοποιεί χαρακτηριστικά από τη διερευνητική και συνεργατική μάθηση και προτείνει μια διαδικασία η οποία αποτελείται από τρία βήματα: (i) Δημιουργία κινήτρου για μάθηση, (ii) Οικοδόμηση της γνώσης μέσω της Διερεύνησης+Συνεργασίας (Constructing knowledge through Exploration+Collaboration) και (iii) Εφαρμογή – Εκλέπτυνση της γνώσης (Applying-Refining knowledge). Οι δραστηριότητες που σχεδιάζονται με βάση το ECLiP πρέπει να αποτελούν ένα ενιαίο σύνολο, να έχουν συνοχή και συνέχεια και να στοχεύουν τόσο στην οικοδόμηση της γνώσης (2ο Βήμα) όσο και στην εφαρμογή της (3ο Βήμα). Απαραίτητη προϋπόθεση για την επίτευξη αυτών των στόχων αποτελεί η ενεργοποίηση των μαθητών και η πρόκληση του ενδιαφέροντος τους για μάθηση (1ο Βήμα) (Gogoulou et al., 2009).

Η επιλογή του Arduino έγινε καθώς είναι μία τεχνολογία φτηνή, αξιόπιστη, ευέλικτη και πολυδιάστατη (Jamieson & Herdtner, 2015). Επιπλέον, είναι εύκολο στην εγκατάσταση, διαθέτει καλά διαμορφωμένο προγραμματιστικό περιβάλλον, υποστηρίζεται από μεγάλη κοινότητα στο Διαδίκτυο, μπορεί να εγκατασταθεί σε όλα τα γνωστά λειτουργικά συστήματα, έχει χαμηλό κόστος όχι μόνο στο υλικό αλλά και στο λογισμικό, και μπορεί να προγραμματιστεί με πληθώρα γλωσσών προγραμματισμού.

Για τον σχεδιασμό των μαθησιακών δραστηριοτήτων εκπαιδευτικής ρομποτικής διαμορφώθηκαν και εμπλουτίστηκαν οι τρεις φάσεις ως εξής:

ΦΑΣΗ 1 : Δημιουργία κινήτρου και πρόκληση ενδιαφέροντος

Στη φάση αυτή οι δραστηριότητες έχουν σκοπό, εκτός από την πρόκληση του ενδιαφέροντος, να αναδειχθούν οι ελλείψεις των γνώσεων και η ανάγκη αναζήτησης νέας γνώσης. Για την πρόκληση του ενδιαφέροντος είναι σημαντικό να γίνει χρήση απλών, γνωστών προβλημάτων από την καθημερινότητα. Αυτό θα αναδείξει τη χρησιμότητα της ρομποτικής, της ηλεκτρονικής και του προγραμματισμού ενώ θα προκληθεί το ενδιαφέρον για την αποκωδικοποίηση και την εξερεύνηση της τεχνολογικής πραγματικότητας που συναντά ο μαθητής. Οι δραστηριότητες μπορούν να περιλαμβάνουν:

- Παρουσίαση ενός μέσου (video, παρουσίαση, εικόνες) και πρόσκληση των μαθητών για σχολιασμό.
- Έτοιμη κατασκευή (μαύρο κουτί) και πρόσκληση για σχολιασμό επί της λειτουργίας της ή της προγραμματιστικής υλοποίησης που «κρύβει».
- Συζήτηση και ανάδειξη των εμπειριών των μαθητών.
- Αναφορά σε δραστηριότητας και έργα που έχουν εκπονηθεί και διαφαινόμενες προοπτικές επέκτασης με νέα δεδομένα.

ΦΑΣΗ 2 : Οικοδόμηση νέας γνώσης μέσω διερεύνησης - συνεργασίας

Ο σκοπός σε αυτή τη φάση είναι οι μαθητές να διερευνήσουν μόνοι τους τα διάφορα χαρακτηριστικά των υλικών που θα χρησιμοποιηθούν στις δραστηριότητες και των προγραμματιστικών τεχνικών που θα κληθούν να ακολουθήσουν. Πειραματίζονται πάνω σε αυτά με σκοπό την κατανόηση εφαρμόζοντας τροποποιήσεις σε υλικό που τους δίνεται και δοκιμάζουν να επιλύσουν παρόμοια μικρά προβλήματα. Επίσης, αξιοποιούνται μέθοδοι συνεργασίας (όπως το μοντέλο οδηγού- παρατηρητή για την καλλιέργεια δεξιοτήτων στην επιχειρηματολογία, στον κριτικό σχολιασμό, στη σύνθεση απόψεων, κ.α.). Στη φάση αυτή μπορούν να σχεδιαστούν δραστηριότητες οι οποίες:

- Με δεδομένο ένα έτοιμο κύκλωμα και πρόγραμμα, οι μαθητές καλούνται να μαντέψουν (ή να ανακαλύψουν) παρατηρώντας το, ποια είναι η λειτουργία του και στη συνέχεια να συγκρίνουν το πραγματικό με το προσδοκώμενο αποτέλεσμα.
- Πάνω σε ένα κύκλωμα που δίδεται έτοιμο, οι μαθητές καλούνται να πειραματιστούν με τροποποιήσεις για την επίτευξη μιας παραλλαγής (που μπορεί να σκεφτούν και οι ίδιοι).

ΦΑΣΗ 3 : Εφαρμογή και εκλέπτυνση της γνώσης

Στην τρίτη φάση οι μαθητές καλούνται να δημιουργήσουν, να αξιολογήσουν και να αξιολογηθούν και να αναπτύξουν περαιτέρω τις γνώσεις τους. Οι δραστηριότητες στη φάση αυτή περιλαμβάνουν:

- Τον αναστοχασμό: γίνεται κριτική αποτίμηση των ιδεών και απόψεων
- Ανάπτυξη μιας νέας εφαρμογής: υλοποιείται ένα project
- Αξιολόγηση: αυτοαξιολόγηση ή αξιολόγηση από ομότιμους

Η υλοποίηση ενός μικρού project, στο πλαίσιο της φάσης 3, το οποίο θα στηρίζεται στις γνώσεις που μέχρι τώρα έχουν αποκτηθεί, είναι ουσιαστική για την εφαρμογή της νεοαποκτηθείσας γνώσης και την αποσαφήνιση τυχόν παρανοήσεων. Οι μαθητές αφήνονται ελεύθεροι στο πλαίσιο εκπόνησης ενός έργου να συνεργαστούν, να πειραματιστούν, να προβληματιστούν και να καταλήξουν σε ένα μοναδικό για αυτούς αποτέλεσμα. Ως αποτέλεσμα δε νοείται μόνο η τελική έκβαση της κατασκευής (που μπορεί να είναι ένα απλό αναβόσβημα ενός LED) αλλά και τα επιμέρους στάδια τα οποία για κάθε ομάδα μπορεί είναι διαφορετικά. Για την ολοκλήρωση της φάσης 3, προτείνεται οι μαθητές να εμπλέκονται σε διαδικασίες αναστοχασμού και αξιολόγησης (αυτο-αξιολόγηση και ομότιμη αξιολόγηση) (Sluijsmans, 1999; Topping, 2017).

Τέλος, η ανάπτυξη συνεργατικού κλίματος με τον σχηματισμό μικρών ή μεγαλύτερων ομάδων είναι άλλος ένας σημαντικός πυλώνας στο πλαίσιο ECLiP. Ειδικά για την περίπτωση της ρομποτικής οι συνεργασίες μπορούν να είναι

πολυδιάστατες (π.χ. προγραμματιστής, βοηθός του, ομάδα σχεδιασμού, ομάδα υλοποίησης κ.λπ.).

3 ΜΑΘΗΣΙΑΚΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ

Με βάση το εμπλουτισμένο πλαίσιο ECLiP, σχεδιάστηκαν και προτείνονται πέντε (5) Φύλλα Εργασίας (ΦΕ) για εισαγωγή στην εκπαιδευτική ρομποτική με το Arduino και βασικά περιφερειακά του εξαρτήματα (Πίνακας 1).

Για τον προγραμματισμό του Arduino και την υλοποίηση των ζητούμενων στα φύλλα εργασίας οι μαθητές καλούνται να χρησιμοποιήσουν τη γλώσσα Wiring C μέσα από το περιβάλλον Arduino IDE αλλά και γλώσσες οπτικού προγραμματισμού όπως η S4A και η Ardublock ώστε ανάλογα με την πρότερη γνώση και τους επιδιωκόμενους γνωστικούς στόχους όσον αφορά στον προγραμματισμό, να επιλέγουν το επιθυμητό προγραμματιστικό περιβάλλον. Επιπρόσθετα χρησιμοποιείται το περιβάλλον Fritzing μέσα από το οποίο μπορεί κάποιος όχι μόνο να προγραμματίσει το Arduino αλλά και να σχεδιάσει το κύκλωμα που καλείται να υλοποιήσει. Έτσι παράλληλα οι μαθητές μαθαίνουν για τον σχεδιασμό CAD στα ψηφιακά κυκλώματα.

Στο πλαίσιο ενός ολοκληρωμένου προγράμματος ρομποτικής-πληροφορικής προτείνεται να συμπεριληφθεί ένα κεντρικό μακροπρόθεσμο project. Για παράδειγμα, στο πλαίσιο ενός project, οι μαθητές μπορεί να χωριστούν σε δύο ομάδες, οι οποίες θα έχουν ως στόχο να κατασκευάσουν ένα ίδιο αντικείμενο (π.χ. κινούμενο όχημα) το οποίο θα εκτελεί κοινές και μη κοινές ενέργειες. Κατά τη διάρκεια της υλοποίησης οι ομάδες θα πρέπει να συνεργαστούν ώστε να πάρουν αποφάσεις σχετικά με θέματα όπως η αλληλεπίδραση των δύο οχημάτων.

Πίνακας 1. Προτεινόμενα Φύλλα Εργασίας

Ενότητα	Τίτλος	Γνωστικά πεδία - λέξεις κλειδιά
Εξαρτήματα γενικού σκοπού	Φύλλο Εργασίας 1 Το Arduino ανάβει το Φως Η Δίοδος Εκπομπής Φωτός (Led)	Οπτική, ηλεκτρισμός, προγραμματισμός
	Φύλλο Εργασίας 2 Το Arduino ρυθμίζει το Ηλεκτρικό Ρεύμα Η Μεταβλητή Ωμική Αντίσταση (Ποτενσιόμετρο)	Ηλεκτρισμός, προγραμματισμός, έλεγχος
Αισθητήρες	Φύλλο Εργασίας 3 Το Arduino ελέγχει τη Θερμοκρασία Ο Αισθητήρας Θερμοκρασίας	Θερμοκρασία, προγραμματισμός
	Φύλλο Εργασίας 4 Το Arduino αντιλαμβάνεται τον Χώρο Ο Αισθητήρας Μέτρησης Απόστασης	Υπέρηχος, προγραμματισμός, μηχανική
Μηχανισμοί	Φύλλο Εργασίας 5 Το Arduino σε Κίνηση Ο Dc Κινητήρας και ο Σερβοκινητήρας	Μηχανική, ηλεκτρισμός, προγραμματισμός

Τα Φύλλα Εργασίας 1 και 2 εφαρμόστηκαν πιλοτικά. Ενδεικτικά περιγράφεται συνοπτικά το ΦΕ 1, «Το Arduino ανάβει το Φως - Η Δίοδος Εκπομπής Φωτός (Led)» όπου γίνεται γνωριμία με το Arduino και τον βασικό τρόπο προγραμματισμού του εν λόγω μικροελεγκτή με τις γλώσσες Wiring C και S4A (Εικόνα 1α). Επίσης, εξετάζονται τα βασικά δομικά υλικά της ηλεκτρονικής και με αυτά υλοποιείται ένα κύκλωμα. Στη συνέχεια γίνεται σύνδεση με τον μικροελεγκτή και με χρήση της πλακέτας και της γλώσσας προγραμματισμού, το κύκλωμα λειτουργεί εκτελώντας τις ενέργειες που σχεδιάστηκαν. Το ΦΕ αποτελείται από πέντε δραστηριότητες. Στο τέλος των ενοτήτων δίνεται προς υλοποίηση ένα μικρό project ώστε οι μαθητές να εφαρμόσουν τις γνώσεις που αποκόμισαν στις προηγούμενες δραστηριότητες.

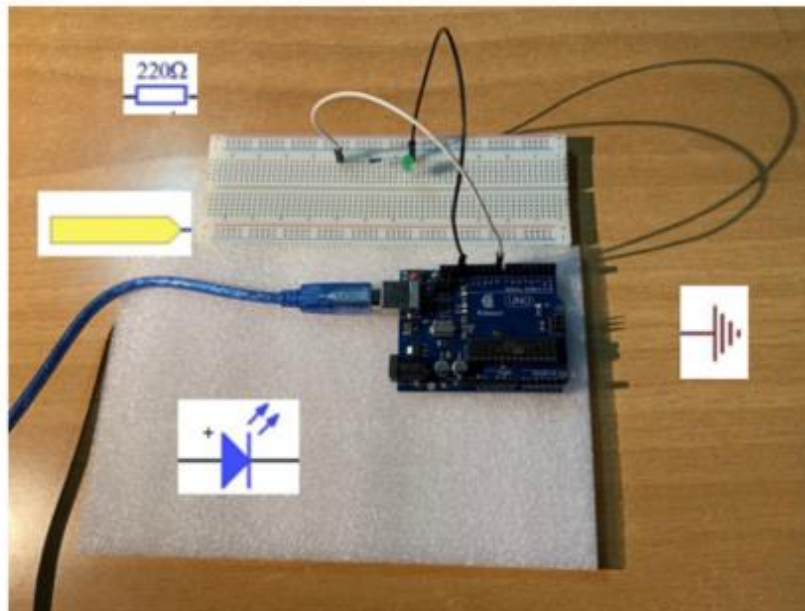
Στην 1^η δραστηριότητα οι μαθητές παρακολουθούν αρχικά ένα σύντομο video clip που παρουσιάζονται έργα (projects) και ιδέες με επίκεντρο τον μικροελεγκτή Arduino με στόχο να προκληθεί το ενδιαφέρον των μαθητών. Στη 2^η Δραστηριότητα οι μαθητές προσπαθούν να αναγνωρίσουν τα εξαρτήματα που τους δίνονται και απαντούν σε σχετικές ερωτήσεις σχετικά με τη χρήση

τους. Επιδιώκεται η δημιουργία νέας γνώσης μέσω της παρατήρησης και της διερεύνησης. Συνεργάζονται σε ομάδες και ανταλλάσσουν γνώσεις οπότε παράλληλα αντιλαμβάνονται την αξία της ομαδικής εργασίας. Ακολουθώντας (δραστηριότητα 3) οι μαθητές διερευνούν περαιτέρω το ηλεκτρονικό κύκλωμα. Παρατηρούν το σχέδιο και παίζουν ένα μικρό παιχνίδι αντιστοιχίσεων.



Ερωτήματα :

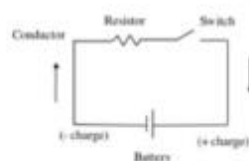
1. Αναγνωρίστε στο διάγραμμα και στη φωτογραφία τα εξαρτήματα που έχετε μπροστά σας. Συγκρίνοντας το διάγραμμα και τη φωτογραφία μπορείτε να αντιστοιχήσετε στην παρακάτω εικόνα τα εξαρτήματα στα pins του διαγράμματος;



2. Με χρήση των υλικών που έχετε στη διάθεσή σας σχεδιάστε με λεπτομέρεια και υλοποιήστε το κύκλωμα που βλέπετε στη φωτογραφία σύμφωνα με το παρακάτω παράδειγμα. Χρησιμοποιήστε τα σύμβολα που βλέπετε στην φωτογραφία. Στη θέση του Arduino βάλτε ένα σχήμα δικής σας έμπνευσης.

Παράδειγμα σχεδίασης ενός κυκλώματος:

Basic Electrical Circuit (Diagram)



Εικόνα 1α: Απόσπασμα Φύλλου Εργασίας 1 (Δρ1)

Στη συνέχεια ζητείται η εφαρμογή των όσων έχουν μάθει μέσα από την απλή τροποποίηση ενός βασικού ηλεκτρολογικού διαγράμματος. Στην 4^η δραστηριότητα δίνονται έτοιμα δύο τμήματα κώδικα σε διαφορετικές γλώσσες προγραμματισμού και ζητούνται απαντήσεις σε ερωτήσεις σχετικά με εντολές που περιέχουν (Εικόνα 1β). Σε αυτή τη φάση ζητείται οι μαθητές «διαισθητικά» να κατανοήσουν το πρόγραμμα και να απαντήσουν. Εκτελώντας τα προγράμματα και παρατηρώντας το αποτέλεσμα επιδιώκεται η κατανόηση της λειτουργίας των εντολών και της σχέσης τους με το υλικό μέρος. Στη συνέχεια ζητείται η τροποποίηση του προγράμματος προκειμένου να επιτευχθεί κάποιο ζητούμενο αποτέλεσμα. Στην 5^η δραστηριότητα δίνεται ένα project ώστε οι μαθητές να εφαρμόσουν τις γνώσεις που αποκόμισαν στις προηγούμενες δραστηριότητες. Το project αφορά γνωστό αντικείμενο της καθημερινής ζωής, τον φάρο, ώστε να παρουσιάζει αυξημένο ενδιαφέρον. Αρχικά δίνονται κατανοητές και σύντομες οδηγίες ώστε με τη βοήθεια του Διαδικτύου να εντοπιστούν οι πληροφορίες και οι απαιτήσεις της δραστηριότητας. Συγκεκριμένα καλούνται να εντοπίσουν πληροφορίες γενικά για τη σημασία και την τεχνολογία των φάρων (αντικείμενο της τεχνικής/τεχνολογίας γνωστό σε όλους) και κατόπιν να βρουν πληροφορίες για έναν συγκεκριμένο (υπαρκτό) φάρο ώστε να καταφέρουν να τον προσομοιάσουν στη συνέχεια με τη βοήθεια του Arduino και του προγραμματισμού. Οι μαθητές, κάνοντας μικρές τροποποιήσεις στον κώδικα του προγράμματος που ήδη έχουν χρησιμοποιήσει από τις προηγούμενες δραστηριότητες και χρησιμοποιώντας το ίδιο κύκλωμα, προσπαθούν να προσομοιάσουν τη λειτουργία του φάρου εργαζόμενοι σε ομάδες και αναλαμβάνοντας συγκεκριμένους ρόλους.

4 ΕΦΑΡΜΟΓΗ

Τα δύο πρώτα φύλλα εργασίας εφαρμόστηκαν κατά τη διάρκεια δύο διδακτικών ωρών σε φοιτητές του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του ΕΚΠΑ στο πλαίσιο του μαθήματος «Διδακτική της Πληροφορικής». Πρέπει να αναφερθεί ότι κανείς από τους συμμετέχοντες δεν είχε πρότερη επαφή με Arduino και εκπαιδευτική ρομποτική γενικότερα. Κατά τη διάρκεια της

εφαρμογής συμπληρώθηκαν από τους φοιτητές οι απαντήσεις στα ερωτήματα των φύλλων εργασίας και χρησιμοποιήθηκαν οι πλακέτες ώστε να γίνει κατάλληλος προγραμματισμός του Arduino σύμφωνα με τα ζητούμενα.



Τρέξτε τα προγράμματα Arduino IDE και S4A. Στη συνέχεια φορτώστε τα προγράμματα led.ino και led.sb αντίστοιχα που θα βρείτε στον φάκελο του Φύλλου Εργασίας 1. Παρατηρήστε τον κώδικα στις δυο γλώσσες προγραμματισμού και απαντήστε στα ερωτήματα που ακολουθούν.

Arduino IDE	S4A
<pre>int ledPin=12; void setup() { pinMode(ledPin,OUTPUT); } void loop() { digitalWrite(ledPin,HIGH); delay(1000); digitalWrite(ledPin,LOW); delay(1000); }</pre>	

Μελετήστε το πρόγραμμα στη γλώσσα Wiring C. Ποιος ο ρόλος της μεταβλητής ledPin; Τι δηλώνει η τιμή που δίνεται;

Ποιο θεωρείτε ότι είναι το αποτέλεσμα της εντολής digitalWrite(ledPin,HIGH) και ποιο της digitalWrite(ledPin,LOW);

Ποιο θεωρείτε ότι είναι το αποτέλεσμα της εκτέλεσης του κώδικα;

Εικόνα 1β: Απόσπασμα Φύλλου Εργασίας 1 (Δρ4)

Για τη διερεύνηση της καταλληλότητας των ΦΕ, θεμάτων μαθησιακών σχεδιασμού και διδακτικής προσέγγισης και προτιμήσεων όσον αφορά το προγραμματιστικό περιβάλλον, οι φοιτητές συμπλήρωσαν ηλεκτρονικά ένα ερωτηματολόγιο (8 από 9 φοιτητές).

Από τις απαντήσεις προκύπτει ότι οι αξιοποιούμενες διδακτικές προσεγγίσεις και η ακολουθία των δραστηριοτήτων ικανοποίησε τους φοιτητές και διευκόλυνε τη μάθηση σε ικανοποιητικό βαθμό. Οι φοιτητές λειτούργησαν μόνοι τους, η παρέμβαση ήταν ιδιαίτερα περιορισμένη κυρίως λόγω των πρότερων γνώσεων των φοιτητών στον τομέα του hardware. Τα παραρτήματα υπηρέτησαν τον στόχο τους καθώς παρείχαν τεχνικά θέματα εγκατάστασης του υλικού και του λογισμικού που χρησιμοποιούνται στα ΦΕ. Οι συμμετέχοντες στην πράξη έδειξαν μεγάλο ενδιαφέρον για την πλατφόρμα το οποίο εκφράστηκε με την ενεργή συμμετοχή τους καθ' όλη τη διάρκεια της εφαρμογής. Το ενδιαφέρον αυτό εκφράστηκε με την επιθυμία της πλειοψηφίας να εμβαθύνει μελλοντικά στην αξιοποίηση του Arduino στην εκπαίδευση. Όσον αφορά την άποψη των συμμετεχόντων για την καταλληλότητα της πλατφόρμας Arduino στις διάφορες βαθμίδες της εκπαίδευσης, από τις απαντήσεις προκύπτει ότι η πλατφόρμα μπορεί να αξιοποιηθεί σε όλες τις εκπαιδευτικές βαθμίδες με το μεγαλύτερο ποσοστό να προτείνει το Λύκειο ως την καταλληλότερη βαθμίδα.

Από τις απαντήσεις που δόθηκαν στις ερωτήσεις των δραστηριοτήτων εξάγονται τα ακόλουθα συμπεράσματα:

- Το ΦΕ1 συμπληρώθηκε από όλες τις ομάδες πλήρως (πλην μιας ομάδας η οποία δεν εκτέλεσε το project) ενώ το ΦΕ2 ελάχιστα. Λόγω έλλειψης εργαστηριακού περιβάλλοντος και λόγω μη έγκαιρης εγκατάστασης του απαιτούμενου λογισμικού (όπως ζητήθηκε), παρουσιάστηκε και αντιμετωπίστηκε αριθμός προβλημάτων αρχικής εγκατάστασης των απαραίτητων προγραμμάτων τα οποία ξεπεράστηκαν στην πορεία πλην όμως αναλώθηκε χρόνος. Στην πράξη, 2 ομάδες ασχολήθηκαν σε ικανοποιητικό βαθμό με τα προγράμματα του ΦΕ2 και τελικά επιχείρησαν και κατάφεραν να περατώσουν το project του ΦΕ2 χωρίς όμως να συμπληρώσουν τα ζητούμενα στις ερωτήσεις.
- Το γνωστικό επίπεδο διαπιστώθηκε ότι ήταν υψηλό και σχεδόν κοινό στον τομέα του προγραμματισμού – αναμενόμενο καθώς πρόκειται για φοιτητές πληροφορικής. Ωστόσο στον τομέα του μικροελεγκτή οι γνώσεις διαπιστώθηκε ότι ποικίλουν από ελάχιστες (1 ομάδα) μέχρι αρκετά προχωρημένες (1 ομάδα) και πολύ προχωρημένες (2 ομάδες).

- Οι απαντήσεις που δόθηκαν στο σύνολο των ερωτήσεων των δραστηριοτήτων του ΦΕ1 ήταν περιεκτικές και στη σωστή κατεύθυνση.
- Η δραστηριότητα 5 (Project) δεν εκπονήθηκε σε μια περίπτωση, γεγονός το οποίο εκτιμάται ότι οφείλεται στη δυσκολία που προέκυψε λόγω της μη ενασχόλησης του συνόλου των ατόμων της ομάδας με μικροελεγκτές ή εν γένει ηλεκτρονική (όπως απαντήθηκε και στη σχετική ερώτηση).

Συμπερασματικά για τη βελτιστοποίηση της εφαρμογής διαφαίνεται ότι απαιτείται μια εισαγωγική ενότητα (εισαγωγή στο Arduino) και οπωσδήποτε, κατά την εφαρμογή των ΦΕ, περιβάλλον εργαστηρίου με προεγκατεστημένα και δοκιμασμένα τα απαραίτητα προγράμματα που χρησιμοποιούνται.

5 ΣΥΜΠΕΡΑΣΜΑΤΑ

Οι ανοικτές τεχνολογίες όπως είναι το Arduino αποτελούν μια εναλλακτική πλατφόρμα εκπαιδευτικής ρομποτικής, η οποία όμως δεν έχει αξιοποιηθεί ευρέως στην πρωτοβάθμια και δευτεροβάθμια εκπαίδευση. Εξαίρεση αποτελούν τα επαγγελματικά Λύκεια στα οποία καθώς υπάρχει ήδη το κατάλληλο περιβάλλον, αξιοποιούνται συναφείς τεχνολογίες σε αρκετά μεγάλο βαθμό.

Η τεχνολογία Arduino παρουσιάζει πλεονεκτήματα όπως κόστος και διαθεσιμότητα υλικού ενώ το αντίστοιχο λογισμικό προσφέρεται δωρεάν. Υποστηρίζεται δε από πολύ μεγάλη κοινότητα και χρησιμοποιείται σε πανεπιστημιακά μαθήματα.

Ο σχεδιασμός κατάλληλων μαθησιακών δραστηριοτήτων και η επιλογή κατάλληλων περιβαλλόντων προγραμματισμού, μπορεί να διευκολύνει την εισαγωγή της εκπαιδευτικής ρομποτικής με την τεχνολογία Arduino στην πρωτοβάθμια και γενική δευτεροβάθμια εκπαίδευση.

Η προτεινόμενη προσέγγιση μαθησιακών δραστηριοτήτων διαφοροποιείται από τις υπάρχουσες τύπου «tutorial» γιατί έχει ως πρωταρχικό στόχο την ανάπτυξη της ενεργητικής μάθησης από τους μαθητές στηριζόμενη από

επικοινωνητικές - κonstrouκτιβιστικές προσεγγίσεις και τη συνεργασία σε ομάδες.

Στα άμεσα σχέδια περιλαμβάνεται η εφαρμογή των δραστηριοτήτων στην πρωτοβάθμια και δευτεροβάθμια εκπαίδευση σε πραγματικές συνθήκες και ο αναλυτικός σχεδιασμός του τελικού project.

ΑΝΑΦΟΡΕΣ

- [1] Alimisis, D. (2013). Educational robotics: Open questions and new challenges. *Themes in Science and Technology Education*, 6(1), 63-71.
- [2] Breiner J., Johnson C., Harkness S., & C. Koehler (2012). What is STEM? A discussion about Conceptions of STEM in education and partnerships. *School Science and Mathematics*, 112(1), 3-11.
- [3] Alimisis, D. (2013). Educational robotics: Open questions and new challenges. *Themes in Science and Technology Education*, 6(1), 63-71.
- [4] Eguchi A. (2014). Educational Robotics for Promoting 21st Century Skills. *Journal of Automation, Mobile Robotics & Intelligent Systems*, 8(1), 5-11.
- [5] Gogoulou, A., Gouli, E., Grigoriadou, M. (2009). Teaching Programming with ECLiP Didactical Approach. *Proceedings of the IADIS International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2009)* (pp. 204-211), Rome, 20-22 November.
- [6] He, Y. & Liang, L. (2019). Application of Robotics in Higher Education in Industry 4.0 Era. *Universal Journal of Educational Research*, 7(7), 1612-1622.
- [7] Jamieson, P. & Herdtner, J. (2015). More missing the Boat — Arduino, Raspberry Pi, and small prototyping boards and engineering education needs them. *IEEE Frontiers in Education Conference (FIE)* (pp. 1-6), El Paso, TX, 2015.
- [8] Mikropoulos, T. A., & Bellou, I. (2013). Educational Robotics as Mindtools. *Themes in Science and Technology Education*, 6(1), 5-14.
- [9] Nübler I. (2016). *New technologies: A jobless future or a golden age of job creation?* Working paper, International Labour Organisation. Τελευταία

πρόσβαση, Φεβρουάριος 2020

https://www.ilo.org/global/research/publications/working-papers/WCMS_544189/lang--en/index.htm

- [10] Sluijsmans, D., Dochy, F., & Moerkerke, G. (1999). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research*, 1, 293-319.
- [11] Schön, S., Ebner, M., & Kumar, S. (2014). The Maker Movement. Implications of new digital gadgets, fabrication tools and spaces for creative learning and teaching. *eLearning Papers*, 39, 14-25.
- [12] Topping, K. (2017). Peer Assessment: Learning by Judging and Discussing the Work of Other Learners. *Interdisciplinary Education and Psychology*, 1(1), 1-17. [https://doi.org/10.31532/](https://doi.org/10.31532/InterdiscipEducPsychol.1.1.007) InterdiscipEducPsychol.1.1.007
- [13] Vandeveld, C., Saldien, J., Ciocci, M. C., & Vanderborght, B. (2013). Overview of technologies for building robots in the classroom. In *International conference on robotics in education* (pp. 122-130).
- [14] Φράγκου Σ., Γρηγοριάδου Μ., Παπανικολάου Σ. (2010). Σχεδιάζοντας δραστηριότητες ρομποτικής για μαθητές Γυμνασίου. Στο Γρηγοριάδου Μ. (επιμ) Πρακτικά του 5ου Πανελλήνιου Συνεδρίου «Διδακτική της Πληροφορικής» (σ. 216-223). Τμήμα Πληροφορικής και Τηλεπικοινωνιών, ΕΚΠΑ, Αθήνα, 9-11 Απριλίου, 2010.