



ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
& ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΕΡΓΑΣΤΗΡΙΟ ΗΛΕΚΤΡΑΚΟΥΣΤΙΚΗΣ & ΤΗΛΕΟΠΤΙΚΩΝ
ΣΥΣΤΗΜΑΤΩΝ



<http://avilab.ee.auth.gr/>

ΓΕ3206 - Τεχνολογία του Ήχου και της Εικόνας:

Αποθήκευση, Επεξεργασία, Μετάδοση

1. Διδάσκοντες

Χαράλαμπος Δημούλας – αναπληρωτής καθηγητής, 2310994245, babis@auth.gr,
babis@eng.auth.gr

Χρήστος Σεβαστιάδης, Δρ Μηχ. - ΕΔΙΠ, 2310996317, csevast@ece.auth.gr

Ιορδάνης Θωΐδης, Διπλ. Μηχ., Υπ. Δρ. ΑΠΘ, ithoidis@auth.gr

Γεώργιος Παπανικολάου, ομότιμος καθηγητής, 2310996284, pap@eng.auth.gr

2. Περιγραφή μαθήματος

Περιεχόμενο Μαθήματος (Syllabus):

Καταγραφή και ψηφιοποίηση οπτικοακουστικού σήματος, Αποθήκευση /αναπαραγωγή ηχητικού και οπτικοακουστικού περιεχομένου (αναλογικά και ψηφιακά μέσα, γραμμικά /μη-γραμμικά), Αλγόριθμοι και πρότυπα συμπίεσης ήχου, ακίνητης και κινούμενης εικόνας, Συστήματα επεξεργασίας ακουστικού και τηλεοπτικού σήματος, Μετάδοση ήχου /εικόνας, Διαδικτυακό ραδιόφωνο /τηλεόραση, Τηλεόραση Υψηλής Ευκρίνειας -HD, Ψηφιακός κινηματογράφος, Στερεοφωνικά (περιφερικά) συστήματα πολλών καναλιών, Στερεοσκοπικά συστήματα κινούμενης εικόνας, Τεχνικές διαχείρισης και σημασιολογικής επεξεργασίας οπτικοακουστικού περιεχομένου.

Στόχοι - Μαθησιακά αποτελέσματα:

Οι φοιτητές /-τριες λαμβάνουν θεωρητικές και πρακτικές γνώσεις για μέσα εγγραφής, μεθόδους επεξεργασίας, εργαλεία ανάλυσης και διαχείρισης δεδομένων ήχου και εικόνας. Εξοικειώνονται με τεχνολογίες συμπίεσης και κωδικοποίησης οπτικοακουστικής πληροφορίας (MPEG, JPEG, codecs κλπ.), με εφαρμογές και πρότυπα “μετάδοσης ροής” (streaming). Αποκτούν τεχνογνωσία στην εξαγωγή και χρήση παραμέτρων περιγραφής ήχου και εικόνας (features), σε αλγορίθμους μηχανικής μάθησης (machine learning), σε σύνθετες τεχνικές σημασιολογικής ανάλυσης, διαχείρισης και ανάκτησης πολυμεσικού περιεχομένου. Με την επιτυχή ολοκλήρωση του μαθήματος οι φοιτητές /-τριες θα είναι σε θέση να κατανοούν και εφαρμόζουν τις σύγχρονες τάσεις αποθήκευσης, επεξεργασίας και μετάδοσης ήχου – εικόνας, όπως διαμορφώνονται στη σημερινή εποχή των Μεγάλων Δεδομένων (Big Data) και του Σημασιολογικού Ιστού (Semantic Web).

3. Προτεινόμενη βιβλιογραφία

- Βρύζας Ν., Κατσαουνίδου Α., Κωτσάκης Ρ. & Δημούλας Χ. (2018). Διερεύνηση Τεχνικών Διπλής Συμπίεσης για την Ανίχνευση Παρουσίας Πολλαπλών Καταγραφών σε Αρχείο Ήχου. Στο 9^ο Εθνικό Συνέδριο "Ακουστική 2018". Πάτρα, 8-9 Οκτ. 2018.
- Βρύζας Ν. (2020). Αυτοματισμοί ανάλυσης και διαχείρισης οπτικοακουστικών ροών στα ψηφιακά μέσα και τη διαμεσική επικοινωνία. Διδακτορική διατριβή (Επιβλέπων: Δημούλας Χ.), Τμήμα Δημοσιογραφίας και Μέσων Μαζικής Επικοινωνίας ΑΠΘ, διαθέσιμο online (09/10/2021 τελευταία πρόσβαση): <http://ikee.lib.auth.gr/record/326357>
- Βρύσης Λ. (2019). Τεχνικές Οπτικοακουστικής Επεξεργασίας & Ανάλυσης της Διαμεσικής Επικοινωνίας. Διδακτορική διατριβή (Επιβλέπων: Παπανικολάου Γ.), Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών (ΤΗΜΜΥ) ΑΠΘ, διαθέσιμο online (04/10/2019 τελευταία πρόσβαση): <http://ikee.lib.auth.gr/record/305871>
- Δημούλας Χ. (2015). Τεχνολογίες συγγραφής και διαχείρισης πολυμέσων: Τεχνικές μη γραμμικής αφήγησης στα νέα ψηφιακά μέσα [Κεφάλαια 1 (1.2), 5 (5.2-5.4), 6, 8 (8.4), 9 (9.2)]. Πρόγραμμα συγγραφής ακαδημαϊκών συγγραμμάτων "kallipos" (www.kallipos.gr), ISBN: 978-960-603-221-9, <https://repository.kallipos.gr/handle/11419/4343>
- Καλλίρης Γ. & Δημούλας Χ. (2013). Ηλεκτρονικά ΜΜΕ. Ανοικτά Μαθήματα ΑΠΘ, Τμ. Δημοσιογραφίας και ΜΜΕ ΑΠΘ, online: <https://opencourses.auth.gr/courses/OCRS105/> (τελευταία πρόσβαση στις 2/10/2018).
- Κατσαουνίδου Α., Βρύζας Ν., Κωτσάκης Ρ. & Δημούλας Χ. (2016). Έλεγχος αυθεντικότητας ψηφιακού περιεχομένου με βάση τον ήχο. Πρακτικά 8^{ου} εθνικού συνεδρίου του Ελληνικού Ινστιτούτου Ακουστικής (ΕΛΙΝΑ) «Ακουστική 2016», Πειραιάς, 3-4 Οκτωβρίου 2016.
- Κωτσάκης Ρ. (2015). Εφαρμογή αλγορίθμων μηχανικής εκμάθησης για εξόρυξη και κατηγοριοποίηση πληροφοριών περιεχομένου στα οπτικοακουστικά μέσα. Διδακτορική διατριβή (Επιβλέπων: Καλλίρης Γ.). Τμήμα Δημοσιογραφίας και ΜΜΕ ΑΠΘ, διαθέσιμο online στο (27/10/2018 τελευταία πρόσβαση): <http://ikee.lib.auth.gr/record/270162>
- Τσίπας Ν. (2018). Συστήματα επεξεργασίας οπτικοακουστικών πόρων (Media Assets). Διδακτορική διατριβή (Επιβλέπων: Παπανικολάου Γ.). ΤΗΜΜΥ ΑΠΘ, διαθέσιμο online (04/10/2019 τελευταία πρόσβαση) στο: <http://ikee.lib.auth.gr/record/301464>
- Katsaounidou, A., Dimoulas, C., & Veglis, A. (2018). *Cross-Media Authentication and Verification: Emerging Research and Opportunities: Emerging Research and Opportunities*. IGI Global.
- Kotsakis, R., Kalliris, G., & Dimoulas, C. (2012). Investigation of broadcast-audio semantic analysis scenarios employing radio-programme-adaptive pattern classification. *Speech Communication*, 54(6), 743-762.
- Vryzas N., Katsaounidou A., Kotsakis R., Dimoulas C., and Kalliris G. (2018). Investigation of Audio Tampering in Broadcast Content. In *Proceedings of the 144th AES Convention, 2018 May 23–26, Milan*.
- Vryzas N., Katsaounidou A., Kotsakis R., Dimoulas C., and Kalliris G. (2019). Audio-Driven Multimedia Content Authentication as a Service. In *Proceedings of the 146th AES Convention, 2019 March 20–23, Dublin*.
- Vryzas N., Matsiola M., Kotsakis R., Dimoulas C. and Kalliris G. (2018). Subjective Evaluation of a Speech Emotion Recognition Interaction Framework. *Proceedings of 13th Audio Mostly (A Conference on Interaction with Sound)*. North Wales, 12-14 Sept. 2018.
- Vryzas N., Vrysis L., Kotsakis R. & Dimoulas C. (2018). Speech Emotion Recognition Adapted to Multimodal Semantic Repositories. *Proceedings of the 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP 18)*. Zaragoza, 6-7 Sept. 2018.
- Vagias C., Vrysis L., Dimoulas C. & Papanikolaou G. (2018). Speech enhancement on the cloud by means of adaptive Wavelet-domain Wiener filtering. *Proceedings of the 9th Greek National Conference "Acoustics 2018"*. Patra, 8-9 Oct. 2018.
- Lartillot O. (2013). *MIR toolbox guide*. Online (last access Oct. 3, 2018): www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox/MIRtoolbox1.5Guide

Loviscach, J., (2010). Do-it-yourself semantic audio. *Audio engineering society tutorial of the 128th AES Convention*, London (Tutorial T1), online (last access Oct. 3, 2018): <https://j3l7h.de/talks/2010-05-22 DIY Semantic Audio.pdf>

4. Αξιολόγηση Μαθήματος

10% Παρουσία - συμμετοχή στο μάθημα

15% Παρουσίαση θέματος εργασίας: Θεώρηση Προβλήματος – Ανάλυση (1^ο παραδοτέο)

35% Τελική παρουσίαση εργασίας (2^ο παραδοτέο)

40% Τελική γραπτή εργασία (3^ο παραδοτέο)

5. Θεματικές περιοχές εργασιών

Επεξεργασία σήματος ήχου – εικόνας (επεξεργασία πλάτους – φάσματος – δυναμικής περιοχής, ανίχνευση συμβάντων /δραστηριότητας, ανίχνευση κίνησης βίντεο, καταστολή θορύβου /enhancement κ.ά.).

Αναγνώριση προτύπων ήχου – εικόνας (αναγνώριση αντικειμένων, κατάτμηση συμβάντων, ανάκτηση περιεχομένου, συναισθηματική ανάλυση κ.ά.).

Αξιολόγηση τεχνικών συμπίεσης /μετάδοσης οπτικοακουστικών δεδομένων (τεχνικές υποκειμενικής και αντικειμενικής αξιολόγησης ποιότητας εμπειρίας θέασης /αναπαραγωγής, QoS/QoE).

Ανίχνευση και εντοπισμός παραποιήσεων σε αρχεία ήχου, κινούμενης και ακίνητης εικόνας (audiovisual content tampering detection, π.χ. ανίχνευση διπλής συμπίεσης, εντοπισμός μεταβάσεων διαγραφής /προσθήκης υλικού κ.ά.).

Σχετικοί σύνδεσμοι – διαγωνισμοί:

<https://dns-challenge.azurewebsites.net/rules/ICASSP2021>

<http://dcase.community/>

https://www.music-ir.org/mirex/wiki/MIREX_HOME

<https://www.kaggle.com/competitions>

<http://multimediaeval.org/>

<https://trecvid.nist.gov/>

6. Πρόγραμμα διαλέξεων

1ο μάθημα: Εισαγωγή – Χρονοδιάγραμμα – Αντικείμενα του μαθήματος
[11/10/2021]

Οδηγίες και σχετικό υλικό στο περιβάλλον elearning (<https://elearning.auth.gr/>)

2ο μάθημα: Σήματα Ήχου – Εικόνες: καταγραφή, ψηφιοποίηση, αποθήκευση
[18/10/2021]

Δημούλας Χ. (2015). Τεχνολογίες συγγραφής και διαχείρισης πολυμέσων: Τεχνικές μη γραμμικής αφήγησης στα νέα ψηφιακά μέσα [Κεφάλαια 1 (1.2), 5 (5.2-5.4)], Πρόγραμμα συγγραφής ακαδημαϊκών συγγραμμάτων "kallipos" (www.kallipos.gr), ISBN: 978-960-603-221-9, <https://repository.kallipos.gr/handle/11419/4343>

3ο μάθημα: Τεχνικές επεξεργασίας, συμπίεσης & μετάδοσης οπτικοακουστικού περιεχομένου [25/10/2021]

Βρύζας Ν. (2020). Αυτοματισμοί ανάλυσης και διαχείρισης οπτικοακουστικών ροών στα ψηφιακά μέσα και τη διαμεσική επικοινωνία. Διδακτορική διατριβή (Επιβλέπων: Δημούλας Χ.), Τμήμα Δημοσιογραφίας και ΜΜΕ ΑΠΘ, διαθέσιμο online (09/10/2021 τελευταία πρόσβαση): <http://ikee.lib.auth.gr/record/326357>

Δημούλας Χ. (2015). Τεχνολογίες συγγραφής και διαχείρισης πολυμέσων: Τεχνικές μη γραμμικής αφήγησης στα νέα ψηφιακά μέσα [Κεφάλαια 6, 8 (8.4), 9 (9.2)]. Πρόγραμμα συγγραφής ακαδημαϊκών συγγραμμάτων "kallipos" (www.kallipos.gr), ISBN: 978-960-603-221-9, <https://repository.kallipos.gr/handle/11419/4343>

4ο μάθημα: Ανάλυση – σημασιολογική επεξεργασία ήχου και εικόνες [01/11/2021]

Βρύσης Λάζαρος (2019). Τεχνικές Οπτικοακουστικής Επεξεργασίας και Ανάλυσης της Διαμεσικής Επικοινωνίας. Διδακτορική διατριβή (Επιβλέπων: Παπανικολάου Γ.), Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΑΠΘ, διαθέσιμο online στο (04/10/2019 τελευταία πρόσβαση): <http://ikee.lib.auth.gr/record/305871>

Κωτσάκης Ρήγας (2015). Εφαρμογή αλγορίθμων μηχανικής εκμάθησης για εξόρυξη και κατηγοριοποίηση πληροφοριών περιεχομένου στα οπτικοακουστικά μέσα. Διδακτορική διατριβή (Επιβλέπων: Καλλίρης Γ.), Τμήμα Δημοσιογραφίας και ΜΜΕ ΑΠΘ, διαθέσιμο online στο (27/10/2018 τελευταία πρόσβαση): <http://ikee.lib.auth.gr/record/270162>

Τσίπας Νίκος (2018). Συστήματα επεξεργασίας οπτικοακουστικών πόρων (Media Assets). Διδακτορική διατριβή (Επιβλέπων: Παπανικολάου Γ.), Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΑΠΘ, διαθέσιμο online (04/10/2019 τελευταία πρόσβαση) στο: <http://ikee.lib.auth.gr/record/301464>

5ο μάθημα: Παρουσίαση θέματος εργασίας (1ο παραδοτέο) [08/11/2021]

Θεώρηση Προβλήματος – Ανάλυση – Προτεινόμενη Αλγοριθμική Προσέγγιση
15λεπτη παρουσίαση κάθε ομάδα
Συζήτηση – αναφορά σχετικών εργαλείων

6ο μάθημα: Παρουσίαση θέματος εργασίας (1ο παραδοτέο) [15/11/2021]

««««»»»»

7ο μάθημα: Παρουσίαση θέματος εργασίας (1ο παραδοτέο) [22/11/2021]

««««»»»»

8ο μάθημα: Παρουσίαση σχετικών εργαλείων – συζήτηση [29/11/2021]

Εξαγωγής παραμέτρων (features) ήχου – εικόνας

Αξιολόγησης και ελέγχου αυθεντικότητας περιεχομένου

Μηχανικής και εμβαθύνουσας μάθησης (machine – deep learning)

Ενδεικτικά παραδείγματα:

Matlab, Sonic Visualizer, Weka, RapidMiner, Google Collab κ.ά.

<https://krisp.ai/>

<https://www.invid-project.eu/>

<http://reveal-mklab.iti.gr/reveal/>

<https://29a.ch/photo-forensics/#forensic-magnifier>

<https://www.sonicvisualiser.org/>

<https://colab.research.google.com/>

<https://www.cs.waikato.ac.nz/ml/weka/>

<https://rapidminer.com/>

9ο μάθημα: Παρουσίαση σχετικών εργαλείων – συζήτηση [6/12/2021]

Επεξεργασία σήματος στη γλώσσα Python (περιβάλλοντα, εργαλεία)

Βασικές βιβλιοθήκες επεξεργασίας & ανάλυσης ήχου - εικόνας

10ο μάθημα: Παρουσίαση σχετικών εργαλείων – συζήτηση [13/12/2021]

Περιβάλλοντα & Βιβλιοθήκες μηχανικής και βαθειάς μάθησης

11ο μάθημα: Παρουσίαση τελικής εργασίας (2ο παραδοτέο) [20/12/2021]

Παρουσίαση τελικής εργασίας - αποτελεσμάτων

15λεπτη παρουσίαση κάθε ομάδας

Συζήτηση

12ο μάθημα: Παρουσίαση τελικής εργασίας (2ο παραδοτέο) [10/01/2022]

««««»»»»

13ο μάθημα: Παρουσίαση τελικής εργασίας (2ο παραδοτέο) [17/01/2022]

««««»»»»