|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **UČNI NAČRT PREDMETA / COURSE SYLLABUS** | | | | | | | | | | | | | | | | | |
| **Predmet:** | | | Govorne in slikovne tehnologije | | | | | | | | | | | | | | |
| **Course title:** | | | Speech and Image Technology | | | | | | | | | | | | | | |
|  | | | | |  | | | | | | | |  | |  | | |
| **Študijski program in stopnja**  **Study programme and level** | | | | | **Študijska smer**  **Study field** | | | | | | | | **Letnik**  **Academic year** | | **Semester**  **Semester** | | |
| Univerzitetni študijski program prve stopnje Elektrotehnika | | | | | **Elektronika** | | | | | | | | 3. | | letni | | |
| 1st cycle academic study programme Electrical Engineering | | | | | **Electronics** | | | | | | | | **3.** | | **summer** | | |
|  | | | | | | | | | | | | | | | | | |
| **Vrsta predmeta / Course type** | | | | | | | | | | | | Obvezni - strokovni/ compulsory professional | | | | | |
|  | | | | | | | | | | | |  | | | | | |
| **Univerzitetna koda predmeta / University course code:** | | | | | | | | | | | | 64154 | | | | | |
|  | | | | | | | | | | | | | | | | | |
| **Predavanja**  **Lectures** | **Seminar**  **Seminar** | | | **Vaje**  **Tutorial** | | | **Klinične vaje**  **work** | | | | **Druge oblike študija** | | | **Samost. delo**  **Individ. work** | |  | **ECTS** |
| **45** |  | | | **30** | | |  | | | |  | | | **75** | |  | **6** |
|  | | | | | | | | | | | | | | | | | |
| **Nosilec predmeta / Lecturer:** | | | | | France Mihelič | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | |
| **Jeziki /**  **Languages:** | | **Predavanja / Lectures:** | | | | slovenski | | | | | | | | | | | |
| **Vaje / Tutorial:** | | | | slovenski | | | | | | | | | | | |
| **Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:** | | | | | | | | |  | **Prerequisits:** | | | | | | | |
| vpis v letnik | | | | | | | | |  | Enrolment in the year of the course. | | | | | | | |
| **Vsebina:** | | | | | | | |  | | **Content (Syllabus outline):** | | | | | | | |
| Uvod: opis področja, kratek zgodovinski opis razvoja govornih in slikovnih tehnologij, značilnosti sistemov za razpoznavanje vzorcev in predstavitev govornih in slikovnih tehnologij v luči razpoznavanja vzorcev.  Osnovne značilnosti slušnega zaznavanja ter govorna komunikacija pri ljudeh. Predstavitve govora in kodiranje informacije z govorom.  Obdelava govora: zajem in predobdelava, značilke govornega signala, razčlenjevanje govornega signala, govorne podatkovne zbirke.  Procesiranje govora  Razpoznavanje govora: vrste sistemov za razpoznavanje, statistično modeliranje akustične in jezikovne predstavitve govora ter njegova pomenska analiza.  Umetni govor: zgradba sistemov za tvorjenje umetnega govora, grafemsko-fonemska pretvorba, modeliranje prozodije, načini tvorjenja umetnega govornega signala.  Dialog: zgradba sistemov za vodenje dialoga, načini vodenja dialoga, vrednotenje delovanja.  Slikovne tehnologije: osnovni pojmi, primeri uporabe, osnovne transformacije slikovnih podatkov, barvni prostori in kodiranje slik.  Obdelava slikovnih podatkov: obdelava slik v slikovnem in frekvenčnem prostoru, modeli šuma in obnavljanje, morfološke operacije in algoritmi, iskanje robov.  Napredni algoritmi: krajevni deskriptroji in njihovo uporaba, detekcija objektov v sliki, razpoznavanje objektov, pod-prostori za predstavitev podatkov.  Segmentacija slik: vrste rojenja in njihova uporaba pri segmentaciji, mean-shift. | | | | | | | |  | | Introduction: description of the field, short outline of the historical develoment of speech and image technologies.  Basic characteristics of visual and auditory perception and human speech-based communication. Representation of speech and image patterns.  Pattern recognition: structural description, pattern recognition systems in general, feature extraction, learning, classification and clustering in pattern recognition systems.  Speech processing: acquisition and preprocessing, speech features, speech signal segmentation, databases of speech.  Speech recognition: types of speech-recognition systems, statistical modelling, acoustic and langauge modelling, semantic analysis of speech.  Artificial speech: systems for speech synthesis in general, grapheme-to-phoneme conversion, prosody modelling, speech-synthesis procedures.  Dialogue: automated dialogue systems in general, approached to designing human-computer dialogue systems, assessment of dialogue systems.  Image technologies: terminology, use-cases, basic image transformations, color images and color spaces, image coding.  Image processing: image processing in the spatial and frequency domains, noise models and image restoration, morphological operations and algorithms, edge detection  Advanced algorithms, local descriptors and their applications, object detection in images, object recognition from image data, subspaces for data representation.  Image segmentation: clustering techniques and thier application to image segmentation, mean-shift. | | | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Temeljni literatura in viri / Readings:** | | | | | |
| 1. Mihelič F., Žibert J., Hajdinjak M., Štruc V., Skripta za predmet Govorne in slikovne tehnologije, 1. Izdaja, Ljubljana, Fakulteta za elektrotehniko, 2012. 2. Mihelič F., Signali, Založba FE in FRI, Ljubljana, 2006. 3. Pavešić N., Razpoznavanje vzorcev: uvod v analizo in razumevanje vidnih in slušnih vzorcev, 3. Popravljena in dopolnjena izdaja, Založba FE in FRI, Ljubljana, 2012. 4. Rabiner L., Schafer R., Theory and Applications of Digital Speech Processing, Prentince Hall, 1. Ed., 2010. 5. Gonzales R. C., Woods, R.E., Digital Image Processing, 3 izdaja, Prentice Hall, 2007. 6. R.C. Gonzales, R.E. Woods, S.L. Eddins, Digital image processing using Matlab, 2 izdaja. Gatesmark Publishing, 2009. | | | | | |
| **Cilji in kompetence:** | |  | | **Objectives and competences:** | |
| Seznanjanje s področjem govornih in slikovnih tehnologij, spoznavanje samodejnih postopkov za izvajanje različnih nalog s tega področja. | |  | | The aim of this course is to acquaint students with the field of speech and image technologies and introduce various algoritms, techniques, and methods to acomplish tasks related to this field. | |
| **Predvideni študijski rezultati:** | | |  | **Intended learning outcomes:** | |
| Znanje o načinu opisov, predstavitev in umetnem tvorjenju ter razpoznavanju govornih in slikovnih signalov. Razumevanje kompleksnosti in interdisciplinarnosti področja. Znanje in razumevanje o zgradbi, načinu delovanja in zmogljivosti sistemov govornih in slikovnih tehnologij. | | |  | Knowledge about the representation, description, synthesis and recognition of speech and image signals. Understanding the complexity and interdisciplinarity of the field of speech and image technologies. Knowledge and understanding of the structure and capabilities of speech- and image-based technologies. | |
|  | | |  |  | |
| **Metode poučevanja in učenja:** | | |  | **Learning and teaching methods:** | |
| Predavanja  Sodelovalno učenje  Laboratorijske vaje | | |  | Lectures  Interactive teaching  Practical assignements | |
| **Načini ocenjevanja:** | Delež (v %) /  Weight (in %) | | | | **Assessment:** |
| Ocena pri predmetu se sestavi na podlagi treh kriterijev:  Pisnega izpita  Ustnega izpita, in  uspešnosti pri laboratorijskih vajah.  Kandidat, ki na pisnem izpitu zbere vsaj 50 % možnih točk in v okviru laboratorijskih vaj zbere vsaj 30% točk, lahko pristopi k ustnemu izpitu.  Kandidat lahko opravi pisni izpit tudi z dvema kolokvijema, pri čemer mora na vsakem kolokviju doseči vsaj 50 % možnih točk.    Kriterij po kategorijah:  Ustni izpit  Pisni izpit (2xkolokvij)  Laboratorijske vaje | 50%  20%  30% | | | | The grade of the course is constracted based on three criteria:  A written exam  An oral exam, and  Practical assignements  To qualify for the oral exam, each student has to score above 50% in the written eaxm and gather more than 30% of the available points during the practical assignements.  Students can skip the written exam by taking part in two colloquia during the school year. Here, a score of more than 50% has to be achieved at each colloqium to earn the right to take the oral exam.  Contribution of each criterium to the final grade:  Oral exam  Written exam (2x colloqium)  Practical assignments |
| **Reference nosilca / Lecturer's references:** | | | | | |
| 1. GAJŠEK, Rok, MIHELIČ, France, DOBRIŠEK, Simon. Speaker state recognition using an HMM-based feature extraction method. *Computer speech & language*, ISSN 0885-2308, Jan. 2013, vol. 27, no. 1, str. 135-150.  2. DOBRIŠEK, Simon, GAJŠEK, Rok, MIHELIČ, France, PAVEŠIĆ, Nikola, ŠTRUC, Vitomir. Towards efficient multi-modal emotion recognition. International journal of advanced robotic systems, ISSN 1729-8814, 2013, vol. 10, no. 53, str. 1-10.  3. VESNICER, Boštjan, ŽGANEC GROS, Jerneja, MIHELIČ, France. Fusion of discriminative and generative scoring criteria in GMM-based speaker verification. V: HABERNAL, Ivan (ur.), MATOUŠEK, Václav (ur.). Text, speech and dialogue : proceedings, (Lecture notes in computer science, ISSN 0302-9743, Lecture notes in artifical intelligence, 6836). Berlin; Heidelberg: Springer, cop. 2011, str. 139-146.  4. PAVEŠIĆ, Nikola, ŽGANEC GROS, Jerneja, DOBRIŠEK, Simon, MIHELIČ, France. Homer II - man - machine interface to internet for blind and visually impaired people. Computer communications, ISSN 0140-3664. [Print ed.], 2003, vol. 26, str. 438-443.  5. DOBRIŠEK, Simon, ŽIBERT, Janez, PAVEŠIĆ, Nikola, MIHELIČ, France. An edit-distance model for the approximate matching of timed strings. IEEE transactions on pattern analysis and machine intelligence, ISSN 0162-8828. [Print ed.], Apr. 2009, vol. 31, no. 4, str. 736-741. | | | | | |