

אביב תשפ"א

דחיסת נתונים ואותות תשפ"א (0510.7104)

זמן ומקום: ימי שלישי 16:00-18:00, כניראה בזום..

מרצה: רמי זמיר, חדר 135 (בניין מעבדות חשמל). טלפון: 6406273
e-mail: zamir@eng.tau.ac.il home page: www.eng.tau.ac.il/~zamir

דרישות קדם מומלצות: תורת האינפורמציה, העברה סיפרתית של אותות (תחליף אפשרי לקורס השני – עיבוד אותות סטאטיסטי ו-או מערכות תקשורת).

מטרת הקורס להרחיב את הידע בנושאים של דחיסת מקורות אינפורמציה (טקסט, קבצים, דיבור, שמע, תמונה ווידאו), תוך מתן רקע מתורת שאנון (אנטרופיה, פונקציה קצב-עיוות), ויישום באלגוריתמים מעשיים. דגשים מיוחדים: למידה והסתגלות, שילוב ברשתות תקשורת.

הקורס יכלול את הנושאים הבאים:

1. **קידוד מקור ללא-עיוות (lossless source coding):** אנטרופיה; קידוד שאנון-פאנו, הופמן, אליאס, קידוד אריתמטי, Tunstall, context, דוגמאות מקידוד פקס ותמונה.
2. **קידוד אוניברסלי (universal compression):** קידוד משפחה של מקורות, חסמים ליתירות, סדרות אינדיבידואליות, קידוד למפל-זיו; קוונטיזציה מסתגלת. יישומים.
3. **קידוד מקור עם עיוות (lossy source coding):** מדדי עיוות, תורת קצב-עיוות של שאנון, קוונטיזציה אופטימלית ותנאי לויד, קוונטיזציה ברזולוציה גבוהה, קוונטיזציה עם קידוד אנטרופיה, קוונטיזציה וקטורית (VQ), סריגים (VQ), Lattices, ברזולוציה גבוהה.
4. **שיטות עיבוד בזמן ובתדר לקוונטיזציה של אותות עם זיכרון:** קידוד התמרה (transform coding) ופסי תדר (sub-band coding), קידוד עם חיזוי (predictive coding) = DPCM; יישומים ותקנים (דחיסת דיבור, קול, תמונה ווידאו).
5. **לימוד והסתגלות בדחיסה: ADPCM,** שיערוך פרמטרי מקור, למידה עמוקה.
6. **קידוד מקור לערוצי-ורשתות-תקשורת:** קידוד משותף מקור-ערוץ; דחיסה בתנאי אי-וודאות; הגנת שגיאות לא אחידה; שידור אנלוגי-דיגיטאלי; עידון בשלבים (successive refinement); ריבוי תאורים; דחיסה מבוזרת, דחיסה עם מידע-צד במקודד/מפענח. חיפוש במידע דחוס ו-digital fingerprint.

דרישות הקורס כוללות הגשת תרגילי בית (כ-20%), בוחן אמצע (כ-30%), השתתפות פעילה בכיתה ועבודת גמר - אפשרות עם סמינר (כ-50%).

ספרות רקע:

1. "Digital Compression for Multimedia", Gibson, Berger, Lookabaugh, Lindbergh, Baker, 1998.
2. "Vector Quantization and Signal Compression", Gersho & Gray, 1992
3. "Network Information Theory", El Gamal and Kim, 2011.

דוגמאות לנושאים לעבודת גמר (רשימה ראשונית)

1. אלגוריתם הופמן דינמי.
2. זיהוי סוג קובץ על ידי שיערוך אנטרופיה.
3. גישה אקראית לבסיס נתונים דחוס.
4. דחיסת גנום.
5. דחיסת וידאו מתקדמת: תקני ה- H, *26.MPEG4, ושיחות ועידה (H.320).
6. דחיסת דיבור בקצב נמוך (G.728, STC) LD-CELP וכו'.
7. "טביעת אצבע" ספרתית לוידאו, לצורך חיפוש במידע דחוס, אימות זהות וכו'.
8. הגבול בין רזולוציה גבוהה ונמוכה בקוונטיזציה.
9. דחיסה עם גלונים ובסיסים מורחבים (wavelets, over-complete expansions).
10. חיזוי אוניברסלי.
11. קוונטיזציה עם סריגים מקוננים.
12. התמרת Burrows-Wheeler לקידוד מותנית הקשר (context-based coding).
13. הצדקה תאורטית לדחיסה מסתגלת (backward-adaptive lossy compression).
14. דחיסת מוסיקה (Dolby, Audio MP3, וכו'.
15. יישומים של למידה עמוקה לדחיסת וידאו (deep video compression).