

Ψηφιακή Επεξεργασία και Ανάλυση Εικόνας

Ακαδημαϊκό Έτος 2017-2018 (Εαρινό Εξάμηνο)

Εργαστηριακή Άσκηση - Μέρος Β

Διαδικαστικά Ζητήματα

1. Υπενθυμίζουμε ότι οι εργασίες είναι ατομικές. Κάθε άτομο θα πρέπει να επιλέξει ένα από τα θέματα που ακολουθούν.
2. Η τεχνική αναφορά που θα υποβάλλετε θα πρέπει να περιλαμβάνει συνοπτική περιγραφή του θέματος, επαρκή σχολιασμό των αποτελεσμάτων τα οποία προκύπτουν σε κάθε περίπτωση και επίσης θα πρέπει να συνοδεύεται από τον κώδικα που γράψατε.
3. Προτεινόμενο περιβάλλον ανάπτυξης για τις διάφορες επεξεργασίες είναι το περιβάλλον Matlab, με απαραίτητα τα toolboxes signal και image.
4. Η παράδοση της άσκησης θα γίνει ηλεκτρονικά στο eclass μέσω της ενότητας Εργασίες, ενώ διευκρινίσεις θα δίνονται μέσω φόρουμ στο eclass: markopoul@ceid.upatras.gr (θέμα 1), ampeliot@ceid.upatras.gr (θέμα 2,3) και trigka@ceid.upatras.gr (θέμα 4).
5. Προθεσμία παράδοσης ορίζεται η 13/07/2018 12.00μμ, σε περίπτωση που επιθυμείτε να σταλεί βαθμός τον Ιούλιο. Η προθεσμία παράδοσης για τον Σεπτέμβριο θα ανακοινωθεί αργότερα.

Θέμα 1: Αντιστοίχιση Εικόνων

Αντιστοίχιση Εικόνων (Image Registration) είναι η διαδικασία ‘ευθυγράμμισης’ δύο ή περισσότερων εικόνων της ίδιας σκηνής, οι οποίες έχουν ληφθεί σε διαφορετικές χρονικές στιγμές ή από διαφορετικά σημεία ή από διαφορετικούς αι-σθητήρες [1.1], [1.2]. Εξαιτίας των παραπάνω λόγων, κάθε εικόνα παρουσιάζει μια παραμόρφωση σε σχέση με την άλλη. Η παραμόρφωση αυτή, μπορεί να είναι μετατόπιση (οριζόντια και κατακόρυφη), περιστροφή, κλιμάκωση (scaling) καθώς και πιο γενικές παραμορφώσεις ή συνδυασμός των ανωτέρω.

Αντικείμενο της συγκεκριμένης εργασίας είναι :

Να υλοποιηθούν τεχνικές αντιστοίχισης εικόνων οι οποίες χρησιμοποιούν τις μετρικές (α) της ελαχιστοποίησης του αθροίσματος τετραγώνων των διαφορών των εντάσεων, (β) της ελαχιστοποίησης του αθροίσματος των απολύτων διαφορών των εντάσεων των εικόνων και (γ) της μεγιστοποίησης του συντελεστή συσχέτισης ανάμεσα στις 2 εικόνες.

Στη συνέχεια, κάθε μια από αυτές τις τεχνικές θα πρέπει να εφαρμοστεί για την αντιστοίχιση των ζευγών εικόνων (A, B) (A1, B1) και (A2, B2) που δίνονται στο αρχείο registration.zip.

Αντικείμενο της εργασίας αποτελεί επίσης και μια βιβλιογραφική παρουσίαση των κυριότερων τεχνικών αντιστοίχισης εικόνων.

[1.1] L. Brown, “A survey of image registration techniques”, *ACM Computing Surveys*, 24 (1992), no. 4, 325-376.

[1.2]B. Zitova and J. Flusser, “Image registration methods: A survey”, *Elsevier Image and Vision Computing* 21 (2003), 977-1000.

Θέμα 2: Ανακατασκευή Εικόνων με χρήση Τεχνικών Συμπιεσμένης Καταγραφής

Τα τελευταία χρόνια παρουσιάζεται έντονο ενδιαφέρον από την επιστημονική κοινότητα για τη μελέτη «αραιών σημάτων». Ως αραιά σήματα ορίζουμε εκείνα για τα οποία υπάρχει κάποιος κατάλληλος μετασχηματισμός, ο οποίος τα μεταφέρει σε ένα χώρο που ορίζεται από ένα κατάλληλο πλήθος σημάτων βάσης, και στον οποίο χώρο τα αραιά σήματα δίνονται ως ένας γραμμικός συνδυασμός λίγων σημάτων βάσης. Σε ένα τέτοιο χώρο, το αρχικό σήμα αντιστοιχεί (γενικά) σε ένα διάνυσμα συντελεστών, όπου τα περισσότερα στοιχεία του διανύσματος έχουν μηδενική (ή γενικότερα, μικρού πλάτους) τιμή. Αν και η περιγραφή αυτή για τα αραιά σήματα φαίνεται αρχικά μάλλον περιοριστική, έχει αποδειχθεί τα τελευταία χρόνια πως πολλά από τα σήματα τα οποία μας περιβάλλουν στην καθημερινή μας ζωή μπορούν να χαρακτηριστούν ως αραιά. Οι τεχνικές Συμπιεσμένης Καταγραφής (Compressed Sensing) αποτελούν έναν τρόπο για να ανακατασκευάσουμε σήματα που έχουν αραιή αναπαράσταση, χρησιμοποιώντας ένα σχετικά μικρό πλήθος μετρήσεων.

Για παράδειγμα, ας θεωρήσουμε μια εικόνα x και το δισδιάστατο διακριτό μετασχηματισμό συνημιτόνου της (2D-DCT) $X = DCT_2\{x\}$. Θεωρούμε πως η εικόνα εισόδου είναι τέτοια, ώστε το σήμα X να έχει ένα μεγάλο πλήθος συντελεστών ίσους με μηδέν. Επομένως, η αρχική εικόνα είναι ένα αραιό σήμα, καθώς υπάρχει κάποιος μετασχηματισμός (εδώ ο 2D-DCT) που το μετασχηματίζει όπως αναφέραμε. Προχωρώντας στο πρόβλημα της ανακατασκευής, θεωρήστε πως αναδιατάσσουμε τα στοιχεία του $X \in R^{M \times N}$ σε ένα διάνυσμα $i \in R^K, K = MN$, και χρησιμοποιώντας κάποιο κατάλληλο πίνακα $F \in R^{k \times K}, k \ll K$, λαμβάνουμε ένα διάνυσμα μετρήσεων $y \in R^k$, σύμφωνα με τη σχέση:

$$y = Fi.$$

Είναι γνωστό από τη γραμμική άλγεβρα πως αν γνωρίζουμε τις τιμές του διανύσματος y και τον πίνακα F , δεν μπορούμε να καθορίσουμε με μοναδικό τρόπο τα στοιχεία του διανύσματος i , καθώς το πλήθος των αγνώστων K είναι πολύ μεγαλύτερο από το πλήθος των εξισώσεων k που έχουμε στη διάθεσή μας. Ωστόσο, δεν λάβαμε υπόψιν μας το γεγονός πως γνωρίζουμε πως το διάνυσμα i είναι αραιό, δηλαδή πως έχει ένα μεγάλο πλήθος στοιχείων ίσα με μηδέν. Για να λάβουμε υπόψη μας την «αραιότητα» αυτή, θα μπορούσαμε να λύσουμε το πρόβλημα

$$\min_i (||y - Fi||_2) \text{ s.t. } ||i||_0 = l$$

δηλαδή να αναζητήσουμε εκείνο το διάνυσμα i το οποίο ελαχιστοποιεί το σφάλμα, ανάμεσα σε όλα τα διανύσματα που έχουν l μη-μηδενικά στοιχεία (για ένα διάνυσμα π.χ. v , το $|v|_0$ είναι το πλήθος των μη-μηδενικών του στοιχείων), θεωρώντας πως γνωρίζουμε και το πλήθος των μη-μηδενικών στοιχείων του διανύσματος i . Αν και με τον τρόπο αυτό καταφέρνουμε να ξεπεράσουμε το πρόβλημα της μη-μοναδικής λύσης, το νέο πρόβλημα δεν μπορεί να λυθεί με αποδοτικό τρόπο, καθώς έγινε συνδυαστικό. Ωστόσο, αν ορίσουμε το πρόβλημα

$$\min_i (||y - Fi||_2) \text{ s.t. } ||i||_1 \leq \delta$$

τότε αυτό προσεγγίζει το προηγούμενο και μπορεί να λυθεί με αποδοτικό τρόπο, όπου δ σταθερά που καθορίζει πόσο αραιό είναι το διάνυσμα i και $||i||_1$ είναι η νόρμα-1 του διανύσματος, δηλαδή το άθροισμα των απολύτων τιμών των στοιχείων του. Τα τελευταία χρόνια έχουν αναπτυχθεί διάφοροι αλγόριθμοι για την επίλυση του προβλήματος αυτού (ή διαφόρων παραλλαγών του). Οι αλγόριθμοι αυτοί είναι γνωστοί ως αλγόριθμοι αραιής κωδικοποίησης ή αραιής προσέγγισης (sparse coding / approximation algorithms). Στα πλαίσια της άσκησης, θα ασχοληθούμε με τον αλγόριθμο αραιής κωδικοποίησης Orthogonal Matching Pursuit (OMP) [2.1].

Για τους σκοπούς της άσκησης δίνεται το αρχείο **Fruits.mat** το οποίο, όταν φορτωθεί στο Matlab, δίνει μια εικόνα μεγέθους 128×128 εικονοστοιχείων. Χρησιμοποιώντας την εικόνα αυτή, απαντήστε τα ακόλουθα:

1. Υπολογίστε το διδιάστατο διακριτό μετασχηματισμό X της εικόνας x που δόθηκε, και επιβεβαιώστε πως η εικόνα αυτή είναι ένα «αραιό σήμα». Υπολογίστε το πλήθος και το ποσοστό των μη-μηδενικών συντελεστών. Λόγω αριθμητικής αρκετοί συντελεστές δεν θα είναι ακριβώς μηδέν. Χρησιμοποιείτε ένα κατώφλι (π.χ. 0,0001) και μηδενίστε τους συντελεστές που είναι κάτω από αυτό.
2. Θεωρήστε τώρα πως δεν έχουμε στη διάθεσή μας την εικόνα (δηλαδή δεν γνωρίζουμε ούτε το x , ούτε το μετασχηματισμό X , ούτε την αναδιάταξη i των συντελεστών του μετασχηματισμού σε διάνυσμα). Ωστόσο, γνωρίζουμε έναν πίνακα F ο οποίος περιέχει τυχαίες τιμές από την Κανονική κατανομή $N(0,1)$ και ένα διάνυσμα μετρήσεων $y = Fi$ το οποίο περιέχει k μετρήσεις. Για το ερώτημα αυτό, υλοποιήστε τον αλγόριθμο OMP και προσπαθήστε να ανακατασκευάσετε το διάνυσμα i (άρα και το X , μέσω αναδιάταξης και την εικόνα x , μέσω εφαρμογής του αντίστροφου μετασχηματισμού DCT), για διάφορα πλήθη μετρήσεων k (π.χ. 4000, 5000, 6000). Επίσης, υπολογίστε πόσο αραιή είναι η ανακατασκευη σε κάθε περίπτωση, και δώστε τις ανακατασκευασμένες εικόνες. Για την υλοποίηση του αλγορίθμου OMP, χρήσιμη είναι η εργασία [2.5].

Βιβλιογραφία:

[2.1] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," in *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231-2242, Oct. 2004.

[2.2] Emmanuel Candes, "Compressive Sensing – A 25 Minute Tour", 2010.

[2.3] J. A. Tropp and A. C. Gilbert, "Signal Recovery from Random Measurements Via Orthogonal Matching Pursuit", *IEEE Transactions On Information Theory*, Vol. 53, No. 12, December 2007.

[2.4] Joel A. Tropp and Stephen J. Wright “Computational Methods for Sparse Solution of Linear Inverse Problems”, Proceedings of the IEEE, Vol. 98, No. 6, June 2010.

[2.5] Rubinstein, Ron, Michael Zibulevsky, and Michael Elad. "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit." *Cs Technion* 40.8 (2008): 1-15.

Θέμα 3: Ψηφιακή Υδατογράφιση Εικόνων

Η Ψηφιακή Υδατογράφιση Εικόνων (Digital Image Watermarking) είναι η διαδικασία ένθεσης ενός αδιόρατου σήματος (υδατογράφημα) σε μία εικόνα, με τέτοιο τρόπο ώστε, από τη μία, να είναι δύσκολο να αφαιρεθεί, και από την άλλη, να μη δημιουργεί προβλήματα στην ποιότητα της εικόνας. Επιπλέον στόχος είναι η αξιόπιστη ανίχνευση ύπαρξης του σήματος αυτού σε μία δεδομένη εικόνα. Το σήμα υδατογραφήματος μπορεί να θεωρηθεί έτσι ως μία “ψηφιακή υπογραφή” η οποία συνδέει μια εικόνα με τον κάτοχο πνευματικών δικαιωμάτων της.

Αντικείμενο της εργασίας είναι η υλοποίηση δύο συστημάτων εισαγωγής προσθετικού υδατογραφήματος σε μια εικόνα, στο χωρικό πεδίο και στο πεδίο των συντελεστών του μετασχηματισμού DCT της εικόνας. Το υδατογράφημα που θα χρησιμοποιηθεί θα είναι ένα ψευδοτυχαίο σήμα, συγκεκριμένης ισχύος και ίδιων διαστάσεων με την εικόνα, το οποίο θα προστίθενται στην αρχική εικόνα: α) στο χωρικό πεδίο και β) στο πεδίο των συντελεστών DCT.

Η ανίχνευσή του θα γίνεται και στις δύο περιπτώσεις με την τεχνική της γραμμικής συσχέτισης και τη χρήση κατωφλίου. Ως εικόνα εισόδου χρησιμοποιήστε την εικόνα **Εικόνα 1** και θεωρείστε ότι η συγκεκριμένη εικόνα δεν θα είναι διαθέσιμη στον αλγόριθμο ανίχνευσης του υδατογραφήματος (blind watermarking).

Η διαδικασία της ανίχνευσης του υδατογραφήματος θα πρέπει να αξιολογηθεί κάτω από ορισμένες “επιθέσεις” οι οποίες θα γίνουν στην υδατογραφημένη εικόνα τόσο στο χωρικό πεδίο όσο και στο πεδίο συντελεστών. Πιο συγκεκριμένα θα πρέπει να εξεταστούν οι περιπτώσεις όπου στην υδατογραφημένη εικόνα

α) προστίθεται λευκός θόρυβος,

β) εφαρμόζεται ένας γεωμετρικός μετασχηματισμός της επιλογής σας, π.χ. scaling, cropping, rotation.

Βιβλιογραφία:

[3.1] Rafael C. Gonzalez and Richard E. Woods, “Digital Image Processing”, 3rd edition 2007.

[3.2] [G.C. Langelhaar](#), I. Setyawan, R. L. Lagendijk, “Watermarking digital image and video data. A state-of-the-art overview”, IEEE Signal Processing Magazine, vol. 17, no. 5, 2000. pp. 20-46, Sept. 2000.

[3.3] Shinfeng D. Lin and Chin-Feng Chen, “A Robust DCT-Based Watermarking For Copyright Protection”, IEEE Transactions on Consumer Electronics, 2000.

Θέμα 4: Ανάκτηση εικόνας από βάση δεδομένων

Σκοπός αυτού του θέματος είναι να εξοικειωθείτε με τη διαδικασία της ανάκτησης εικόνων. Για το λόγο αυτό σας δίνεται μία βάση εικόνων (**DataBase.rar**), που είναι ουσιαστικά ένας φάκελος με έναν αριθμό από εικόνες, από τις οποίες μερικές μοιάζουν μεταξύ τους και θα πρέπει να υλοποιηθούν κατάλληλες ρουτίνες για την ανάκτησή τους, με χρήση της Ανάλυσης Κύριων Συνιστωσών μιας εικόνας (Principal Component Analysis - PCA).

Ο σκοπός της PCA είναι να μειώσει την μεγάλη διάσταση του χώρου των δεδομένων, αναπαριστώντας τα σε ένα μικρότερο χώρο βασικών. Αυτός ο χώρος μπορεί να έχει πολύ μικρότερη διάσταση, ειδικά σε περιπτώσεις που υπάρχει υψηλή συσχέτιση μεταξύ των παρατηρούμενων μεταβλητών. Ο πυρήνας της PCA είναι ουσιαστικά μια διάσπαση ιδιοτιμών και επιμέρους συνιστωσών που ονομάζονται ιδιο-εικόνες. Ο σκοπός είναι, αρχικά να δημιουργήσετε ένα σύστημα το οποίο θα υπολογίζει την PCA των εικόνων της συγκεκριμένης βάσης δεδομένων.

Στην συνέχεια να πραγματοποιήσετε ανάκτηση των εικόνων δηλαδή: δίνοντας μια εικόνα ως είσοδο σε αυτό το σύστημα θα πρέπει αυτό να την εντοπίζει στην βάση δεδομένων.

Ζητούμενα:

α) Στο φάκελο test σας δίνονται 10 εικόνες με τις οποίες θα πειραματιστείτε για τον έλεγχο του συστήματος PCA που θα κατασκευάσετε. Σκοπός είναι για αυτές τις εικόνες να εξάγετε ποσοστά επιτυχούς ανάκτησης. Η επιτυχής ανάκτηση αναφέρεται στην εύρεση της καλύτερης δυνατής προσέγγισης, δηλαδή της εγγύτερης στη δοθείσα βάση εικόνων. Για να εξάγεται τα ποσοστά επιτυχούς ανάκτησης πειραματιστείτε για το 100%, 50% και 10% του αριθμού των συνιστωσών που θα κρατήσετε για κάθε εικόνα. Είναι σημαντικό για κάθε εικόνα να αναγνωρίζονται οι κύριες συνιστώσες για λόγους υπολογιστικής πολυπλοκότητας. Σχολιάστε τα αποτελέσματα.

β) Δοκιμάστε να κάνετε την ίδια διαδικασία σε εικόνες αλλοιωμένες από Γκαουσιανό θόρυβο με $SN R = [0, 10, 20]dB$ και σχολιάστε τα αποτελέσματα.

Βιβλιογραφία:

[4.1] Lindsay I Smith, “A tutorial on Principal Components Analysis”, Cornell University, 2002.

[4.2] Jonathon Shlens, “A tutorial on Principal Components Analysis”, arXiv preprint, 2014.

[4.3] Turk, M. and Pentland A., “Eigenfaces for Recognition”, Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-86, 1991.

[4.4] W.Zhao, R.Chellappa, A.Rosenfeld and P.J.Phillips, "Face Recognition: A Literature Survey", ACM Computing Surveys, pp.399-458, 2003.