A black background with a black square

Description automatically generated with medium confidenceΕθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Εαρινό Εξάμηνο 2023-2024

Τεχνολογια αναλυσησ εικονα και βιντεο

Λύσεις Θεμάτων

Ιωάννης (Χουάν) Τσαντήλας

03120883

Contents

[Κανονική 23 2](#_Toc169119727)

[Θέμα 1 2](#_Toc169119728)

[Θέμα 2 2](#_Toc169119729)

[Θέμα 3 3](#_Toc169119730)

[Θέμα 4 3](#_Toc169119731)

Κ22, Ε22

Κ21

Κ20

Κ16, Ε16

# Κανονική 23

Θέμα 1

**Ερώτημα 1**

Εξετάζουμε τον Λαπλασιανό τελεστή. Για κάθε εικονοστοιχείο υπολογίζει:

1. Την παράγωγο στην κατακόρυφη διεύθυνση.
2. Την παράγωγο στην οριζόντια διεύθυνση.
3. Το μέγεθος και την κατεύθυνση των ακμών.

**Ερώτημα 2**

Εξετάστε μια εικόνα στην οποία εφαρμόζεται ο Λαπλασιανός τελεστής για ανίχνευση ακμών. Η ύπαρξη μιας ακμής στην αρχικά εικόνα, στην φιλτραρισμένη εικόνα αντιστοιχεί:

1. Σε διάβαση από το μηδέν.
2. Σε μηδενισμό.
3. Σε τοπικό μέγιστο.
4. Σε τοπικό ελάχιστο.

**Ερώτημα 3**

Εάν η εικόνα είναι θορυβώδης, τι είδους φίλτρο πρέπει να εφαρμοστεί στην εικόνα πριν από την εφαρμογή του Λαπλασιανού τελεστή;

1. Ένα ανωδιαβατό φίλτρο
2. Ένα κατωδιαβατό φίλτρο.
3. Ένα ζωνοπερατό φίλτρο.

**Ερώτημα 4**

Μεταβάλλοντας τη διάσταση αυτού του φίλτρου, μπορούμε να εντοπίσουμε ακμές με:

1. Διαφορετικό contrast.
2. Διαφορετικό μήκος.
3. Διαφορετικό resolution.

Θέμα 2

Θεωρήστε το χωρικό φίλτρο Η:

|  |  |  |
| --- | --- | --- |
| 1 | -3 | 0 |
| -3 | 0 | 4 |
| 0 | 4 | 1 |

1. Αυτό εφαρμόζεται σε μία εικόνα gray scale I με pixels των 5-bits. Ποια είναι η μέγιστη και ελάχιστο δυνατή τιμή των pixels της εικόνας J που προκύπτει μετά την εφαρμογή του φίλτρου; Μην κάνετε κανονικοποίηση.
2. Ποιος Μ/Σ επιπέδου του γκρι πρέπει να εφαρμοστεί στα pixels της J για να φέρει τις τιμές τους πάλι στα 5-bits;

Θέμα 3

Εξετάστε μια εφαρμογή βιντεοτηλεφωνίας χρησιμοποιώντας τη Σύσταση ITU-T H.261. Η ακολουθία βίντεο κωδικοποιείται με χωρική ανάλυση CIF (352x288 δείγματα για τη φωτεινότητα, 4:2:0), ρυθμό καρέ 10 Hz και κανάλι με σταθερό bitrate 64 kbits/s. Ο buffer εξόδου έχει μέγεθος 12,800 bits. Απαντήστε στις παρακάτω ερωτήσεις, εξηγώντας τις απαντήσεις σας.

1. Ποιος είναι ο μέσος αριθμός bits κάθε εικόνας που κωδικοποιείται;
2. Ποιος είναι ο μέγιστος και ποιος ο ελάχιστος αριθμός bits που μπορούν να χρησιμοποιηθούν για την κωδικοποίηση κάθε εικόνας;
3. Ποιος είναι ο μέγιστος αριθμός bits που μπορεί να χρησιμοποιηθεί για την κωδικοποίηση μιας εικόνας εάν ο buffer είναι γεμάτος όταν η εικόνα αρχίζει να κωδικοποιείται;
4. Ποιος είναι ένας τυπικός αριθμός διανυσμάτων κίνησης που μπορούν να χρησιμοποιηθούν για την κωδικοποίηση μιας εικόνας;
5. Λαμβάνοντας υπόψη ότι χρησιμοποιείται ένα κανάλι σταθερού ρυθμού bits, ποιο τμήμα του κωδικοποιητή λαμβάνει τις πιο σημαντικές αποφάσεις για να επιτρέψει στον κωδικοποιητή να ελέγχει τον αριθμό των bits που δαπανώνται για κάθε εικόνα;

Θέμα 4

1. Εάν τα δεδομένα εισόδου σε ένα επίπεδο συνελικτικού νευρωνικού δικτύου έχουν διάσταση 200x200x128 και εφαρμόζουμε φίλτρα διαστάσεων 7x7x128, ποια είναι η διάσταση της εξόδου του επιπέδου αυτού, αν αποτελείται από 10 activation maps;
2. Σωστό ή Λάθος:
   * Τα CNNs μπορούν να μάθουν να αναγνωρίζουν ένα αντικείμενο σε μια εικόνα ανεξάρτητα από το αν το αντικείμενο μετατοπίζεται οριζόντια ή/και κάθετα.
   * Τα CNNs μπορούν να μάθουν να αναγνωρίζουν ένα αντικείμενο σε μια εικόνα ανεξάρτητα από το αν το αντικείμενο περιστρέφεται (στο επίπεδο της εικόνας).
3. Σχεδιάζετε ένα σύστημα βαθιάς μάθησης για τη διάγνωση του καρκίνου του στήθους με ανάλυση ακτινογραφιών. Ποια πιστεύετε ότι μπορεί να είναι η καταλληλότερη μέτρηση αξιολόγησης: accuracy, precision ή recall; Γιατί;
4. Εργάζεστε ως μηχανικός προσπαθώντας να σχεδιάσετε ένα μοντέλο που προβλέπει εάν ένα άτομο θα κάνει κλικ σε μία δεδομένη διαφήμιση. Δοκιμάζετε 2 προσεγγίσεις:
   * Εκπαιδεύετε ένα μεγάλο μοντέλο, το οποίο χρειάζεται πολλές ώρες για να εκπαιδευτεί.
   * Εκπαιδεύετε 10 λίγο μικρότερα μοντέλα και υπολογίζεται τον μέσο όρο των αποτελεσμάτων τους.

Και τα 2 μοντέλα σας δίνουν ακρίβεια 90% στο σετ ελέγχου σας. Πρέπει να διαλέξετε μία από τις 2 προσεγγίσεις για τη χρήση σε παραγωγική διαδικασία μέσω μιας μηχανής με μία GPU που πρέπει να παρέχει προβλέψεις σε πραγματικό χρόνο, όσο πιο γρήγορα γίνεται. Ποια προσέγγιση είναι πιο κατάλληλη και γιατί;

# Κανονική 22

Θέμα 1

**Ερώτημα 1**

Να θεωρήσετε την ακόλουθη εικόνα 5x5 με τις αντίστοιχες τιμές Pixels:

A black and white square

Description automatically generated

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 180 | 160 | 160 | 140 | 120 |
| 110 | 110 | 120 | 140 | 120 |
| 110 | 140 | 120 | 120 | 140 |
| 120 | 160 | 160 | 170 | 170 |
| 170 | 120 | 110 | 140 | 110 |

Δώστε τον πίνακα συχνοτήτων για κάθε σύμβολο. Ποιος είναι ο ελάχιστος αριθμός bits ανά σύμβολο για αυτό το μήνυμα; Σχεδιάστε τον κωδικό Huffman για κάθε σύμβολο. Υπολογίστε τον μέσο αριθμό bits ανά σύμβολο.

**Ερώτημα 2**

Γιατί οι υπηρεσίες βίντεο χρησιμοποιούν συνήθως υποδειγματοληψία χρώματος 4:2:0;

Θέμα 2

**Ερώτημα 1**

Να θεωρήσετε τον Μ/Σ DFT των 2 εικόνων που φαίνονται παρακάτω.

A black and white rectangular object with white letters

Description automatically generated

Ποια είναι η σχέση της φάσης των μετασχηματισμών DFT των 2 εικόνων; Να εξηγήσετε σύντομα.

**Ερώτημα 2**

Ένας μαθητής έγραψε την ακόλουθη πρόταση: «Το φίλτρο ενδιάμεσης τιμής μπορεί να μειώσει τον αριθμό των κατειλημμένων τιμών (bins) σε ένα ιστόγραμμα. Δεν θα αυξήσει ποτέ τον αριθμό τους. Συμφωνείτε ή διαφωνείτε; Γιατί;

**Ερώτημα 3**

Γιατί τα πρότυπα MPEG-1 και MPEG-2 χρησιμοποιούν τη δομή GOP με περιοδικά πλαίσια Ι; Για βιντεοδιάσκεψη ή εφαρμογές τηλεφώνων βίντεο, μπορεί ο κωδικοποιητής να εισάγει περιοδικά I-frames; Να δικαιολογήσετε την απάντηση σας.

Θέμα 3

Να αντιστοιχίσετε τις εικόνες 1-4 με τα υπογράμματα α-δ, δίνοντας μία σύντομη εξήγηση.

A close-up of a graph

Description automatically generated

Θέμα 4

**Ερώτημα 1**

Η είσοδος σε ένα νευρωνικό δίκτυο (όχι συνελικτικό) είναι μια έγχρωμη εικόνα 300x300. Το 1ο κρυμμένο επίπεδο του δικτύου έχει 100 νευρώνες πλήρως συνδεδεμένους στην είσοδο. Πόσες άγνωστες παραμέτρους έχει το επίπεδο αυτό;

**Ερώτημα 2**

Στα συνελικτικά δίκτυα η χρήση «parameter sharing» επιτρέπει τα επόμενα:

* Μείωση του συνολικού αριθμού άγνωστων παραμέτρων.
* Εφαρμογή ανίχνευσης features σε πολλαπλές θέσεις των εικόνων εισόδου

Το sharing των παραμέτρων που υπολογίστηκαν για μια κατηγορία εικόνων και σε άλλες διαφορετικές κατηγορίες (transfer learning). Συμφωνείτε ή όχι; Γιατί;

**Ερώτημα 3**

Ένα συνελικτικό δίκτυο εφαρμόζεται σε εικόνες προσώπων και βγάζει στην έξοδο Ν χαρακτηριστικά σημεία (landmarks) για το εν λόγω πρόσωπο. Πόσες εξόδους έχει το δίκτυο: Ν2, Ν, 2Ν, 3Ν;