

# Számítógépes Látórendszerek

## Vizsga kérdések – 2023

*Az alábbi kérdéssor a tárgy ismeret-ellenőrzéséhez kapcsolódó kérdésköröket ismerteti.*

*A vizsga e kérdéskörökre épül, ugyanakkor a kérdések nem feltétlen ebben a megfogalmazásban kerülnek a hallgatók elé.*

*A kérdések megalkotásánál fontos szempont volt, hogy részletesen szerepeljen, hogy mik azok a részletek, aminek az értését/tudását a vizsga során számonkérjük (ezért is meglehetősen hosszú a megfogalmazásuk).*

*A ZH során a kérdések a ZH előtti utolsó előadásig elhangzott részhalmlaza kerül számokérésre. (51. kérdés)*

*Az államvizsga tárgyként e tárgyat választó hallgatók számára a kérdéssor a tárgyat 2023 tavaszi félévben végző hallgatók államvizsgatételeinek alapjaként is szolgál.*

## I. HAGYOMÁNYOS LÁTÁS

### 1. Képi alapok

1. Alapfogalmak: számítógépes látás, képfeldolgozás. Alapvető feladatok: osztályozás, detektálás, szegmentálás (és fajtái). Nehézségek.
2. Képképzés, fotodióda működése, CCD és fajtái, CMOS (előnyök hátrányok). Képek felépítése és alapvető tulajdonságai, hibái. Kamerák típusai.
3. Színes képek előállítás. Három fontos színtér, színkomponensek értelmezése, a színtér alapvető előnye(i) és célja. Szürkeárnyalatosítás változatai.

### 2. Képjavítás

4. Intenzitás transzformációk (melyik mit csinál), hisztogram, felhasználása, és azon végezhető műveletek.
5. Képi zajok típusai, konvolúciós szűrők. Simító, élesítő, élkereső szűrők felismerése súlymátrix alapján. Lineáris vs rank szűrők, melyik mire jó, mi az előnye.
6. Képi matematika alapvető műveletei és céljuk. Interpoláció elve, különböző algoritmusai, ezek tulajdonságai, előnyei és hátrányai.

### 3. Frekvenciatartomány

7. 2D Fourier transzformáció elve, frekvenciaváltozók, amplitúdó és fázisspektrum értelmezése. Fázistorzítás hatása. Gyors Fourier-transzformáció elve.
8. Szűrések frekvenciatartományban. Ideális szűrők és típusaik, melyik mire jó, mi a hátrányuk. Butterworth és konvolúciós szűrők. Hogyan néz ki egy konvolúciós szűrő frekvenciatartományban.
9. DCT és FCT, hogyan számoljuk, előnyei. Hogyan használható tömörítésre, hogy működik a JPEG.
10. Hogyan használható az FFT periodikus zajok szűrésére, vagy a periodicitás kiemelésére.

Alakfelismerés megoldása FFT segítségével. Dekonvolúció és fajtái.

#### 4. Képjellemzők

11. Képjellemzők típusai, előnyök, hátrányok. Template Matching működési elve és korlátai (mikor tud jól működni, forgatás és skálázás hatása). Felhasználható mércék és közöttük lévő különbség.

12. Mi az az optikai áramlás, mik a feltételezései és limitációi, hogyan működik 1D-ben, intenzitás áramlás egyenlet levezetése. Megoldhatóság problémái, melyik irányban tudjuk meghatározni?

13. Lucas-Kanade és Farneback módszer alapelve és működése, gyakorlati kérdések, megoldhatóság. Iteratív és piramis módszerek. Ritka és sűrű áramlás.

14. Élkeresés, első és második derivált alapú, deriválás fokszáma szűrőablak alapján. Sobel és Prewitt operátorok, irányfüggőség. Canny algoritmus elve és lépései.

15. Mi az a Hough-transzformáció és hogyan működik? Miért jobb nekünk paraméteresen leírni az egyeneseket. Milyen alakzatokra használható még?

16. Sarokdetektálás: Mi a képi sarok, KLT és Harris alapelve, lokális struktúramátrix. Saroksági kritériumok, két detektor közti különbségek. Mire invariáns a két módszer?

17. SIFT módszer: hogyan detektálunk skálainvariáns módon, milyen szűrőt használunk, és hogyan tudunk ezzel egy skálát hozzárendelni a jellemzőhöz? Hogyan generáljuk a leíró, és hogyan oldjuk meg a négy fő transzformációra való invarianciát?

18. ORB módszerről ugyanez.

19. Követés változatai, nehézségei. Milyen jellemzőket használhatunk fel, hogyan lehet összehasonlítani őket (affinitás) és hogyan párosíthatjuk őket? Mi az a drift és miért van szükség újradetektálásra?

20. Rejtett Markov Modell felépítése és a rajta megoldható feladatok, hogyan használjuk követésre és miben jobb? Hogyan működik (és miért úgy) a Kalman-szűrő és miben nyújt többet a Markov modellnél?

#### 5. Szegmentáció

21. Intenzitás alapú szegmentálás: küszöbözés és eredménye. Hisztogram alapú küszöbözés. Miért nem olyan jó ez?

22. Klaszterezés: hogy működik a k-Means, és a Mixture of Gaussians? Mi a kettő közti kapcsolat (különbségek, hasonlóságok). Mi az Expectation Maximization módszer? Hogyan választjuk ki a klaszterek számát. Mean-shift működése, miért és hogyan oldja meg az előbbi problémát.

23. Klaszterezés hátránya, hogyan oldják ezt meg a régió alapú módszerek. Régiónövelés és a Split&Merge módszerek működése, előnyök-hátrányok.

24. Egyéb szegmentációs módszerek: Gráfágás, és Watershed alapú szegmentáció működése. Mozgás alapú szegmentáció, alapvető nehézség, gauss alapú háttérmodell, variancia alapú küszöbözés.

#### 6. Bináris képek

25. Erózió, dilatáció, hatásuk a méretre és formára (strukturáló elem), szürkeárnyaltos eset. Nyitás és zárás, olvasztásmentes dilatáció, kontúrkeresés.

26. Milyen szomszédosságok vannak, miért fontos ez (Jordan-tulajdonság), hogyan lehet feloldani a problémát? Mi az a csontváz és hogyan lehet meghatározni?

27. Objektumok számlálása és címkézése. Milyen módszerek vannak, (előnyök, hátrányok), és ezek hogyan működnek?
28. Objektumok tulajdonságai: projekció, középpont, befoglaló téglalap, orientáció. Melyik mire jó, hogyan számolható? Euler-szám, lenyomat, lánckód számítása. Szubpixeles pozíció meghatározása.

## 7. Döntéshozás

29. Mi a hagyományos mesterséges intelligencia (GOFAI) és a gépi tanulás közti különbség? Hogyan épülnek fel a gépi tanuló algoritmusok, milyen típusai vannak, és milyen általános szabályokra érdemes odafigyelni?
30. Hogyan működik a kNN algoritmus. Milyen problémái vannak? Mire jó a lineáris regresszió és hogyan működik? Miben tér el a TLS változat és azt hogyan lehet megoldani?
31. SVM alapfogalmai, alapvető működése. Becslés módszerének szemléletes magyarázata. Mi az a kernel függvény, és mit befolyásol a megválasztása?
32. Dimenzió redukció alapfeladata, mire jó, hogyan működik a főkomponens-analízis. Miért kell az adatokat nulla közepűvé tenni, és miért érdemes a skálázásra figyelni? Miben más a lineáris diszkriminancia analízis?
33. Mi az a Haar-féle jellemző, hogy számolható, miért számoljuk integrálkép segítségével? Hogyan lehet ezekből arcfelismerő algoritmust készíteni, és miért kaszkádosítjuk a megtanult jellemzőket?
34. Hogyan működik a Bag of Visual Words algoritmus? Mi az a vizuális szó, hogyan konstruálunk szótárat, és hogyan végzünk osztályozást? Mi az Eigenfaces és hogyan működik?

## II. DEEP LEARNING

### 8. Neurális hálózatok

35. Hogyan működik a Perceptron modell és mit jelentenek az egyes kimenetei? Hogyan néz ki a Perceptron döntési függvénye? Mi az a multi-layer perceptron (MLP), és milyen limitációt orvosol?
36. Milyen hibafüggvényekkel taníthatjuk a Perceptront? Mi a Hinge költség alapelve és alakja? Mi a kereszt-entrópia alapelve, hogyan érjük el, hogy a Perceptron kimenetén valószínűségek legyenek, és hogyan definiáljuk az „igazi” osztály valószínűségeket? Mi a két módszer előnye/hátránya?
37. Mi az a regularizáció és milyen módjai vannak? Miért van rá szükség, és hogyan kapcsolódik az overfitting jelenséghez?
38. Mi a gradiens-módszer alapelve, hogyan működik? Miért használunk minibatcheket, mit határoz meg a minibatch mérete? Milyen módszerekkel javítjuk a gradiens módszer hatékonyságát (momentum, skálázás) és ezek hogyan segítenek?
39. Mi a Newton-módszer alapelve, hogyan működik? Miért alkalmazzuk ritkán mély neurális hálók esetében? Mi az előnye/hátránya a gradiens módszerrel szemben, és hogyan kezelhetők a hátrányai (Levenberg-Marquardt alapelve, működése)?
40. Hogyan lehet mély neurális hálók esetében meghatározni a korábbi rétegek hiba szerinti deriváltját (backpropagation elve és működése)? Backpropagation kiszámolása minta gráfon. Láncszabály normál és vektoros esetben, Perceptron réteg deriváltjai. Neurális háló réteg általános interfésze.

## 9. Konvolúciós Neurális Hálók

41. Konvolúciós hálók réteg típusai. Konvolúciós réteg alapvető működése, paraméterei, lineáris réteggel szembeni előnyei. Pooling réteg és típusai, lehetséges helyettesítése. Aktiváció típusok, szigmoid és tanh (problémák), ReLU és változatai. Általános felépítés, a háló működésének szemléletes magyarázata.
42. Fontos neurális háló architektúrák, AlexNet és VGG (pontos felépítés nem kell), neurális háló rétegek paraméterszámának, memóriaigényének és számítási igényének számítása.
43. Inception réteg, miért jó, hogy párhuzamos mini-hálókat teszünk a neurális hálóba. Miért alkalmazunk bottleneck-et? Hogyan épül fel a GoogLeNet, és miért van több osztályozó kimenete?
44. Miért nehéz nagyon mély neurális hálókat tanítani (gradiens numerikus problémái), és hogyan oldja ezt meg a ResNet? Milyen fajtái léteznek (bottleneck, WRN, DenseNet) és ezek miben különböznek?
45. Hogyan tudjuk megmondani, hogy egyes neuronok aktivációját mely pixelek okozzák a képen? Hogyan lehet neurális hálók egyes neuronjait vizualizálni? Mi az a guided backpropagation, és miért alkalmazzuk ebben az esetben? Mik azok az ellenséges példák, és hogyan lehet őket generálni? Hogyan lehet védekezni ellenük?

## 10. Deep Learning a gyakorlatban

46. Mik azok az alapvető problémák amikkel neurális hálók tanítása esetén szembekerülhetünk? Melyik módszer melyik problémá(k)ra ad megoldást?
47. Miért fontos az inicializáció a konvergencia esetén? Hogyan függnek össze az egyes rétegek súlyai, aktivációi és a gradiens? Milyen numerikus problémák adódhatnak mély hálók esetében és ezek a súlymátrix milyen mértékével függnek össze (és hogyan)? Mi alapján működik a Xavier és a He inicializáció? Miért kell normalizálni a bemenetre adott adatokat?
48. Hogyan ellenőrizzük egy neurális háló általánosító képességét? Hogyan néz ki egy jól általánosító háló aktivációinak eloszlása? Hogyan oldja meg ezt a problémát a dropout, és hogyan kell a tesztelés során a hálót módosítani ebben az esetben? Hogyan segít a batch normalizálás ugyanezen a problémán, és milyen fajtái léteznek (ezeket mikor érdemes alkalmazni)? Mit old még meg?
49. Hogyan tudunk még az overfittingen segíteni (regularizáció, adat augmentáció)? Hogyan lehet kis adatbázisokkal jól tanítani nagy hálókat, és ez miért működik jól (transfer learning).
50. Hogyan érdemes hiperparamétereket keresni és miért? Milyen hatása van a tanuási ráta nagyságának, és hogyan érdemes ezt beállítani?
51. Milyen problémák adódnak neurális hálók installációja során, és ezekre milyen módszerek vannak? Hogyan működik a pruning, a weight sharing, valamint a modell együttesek módszere?

## 11. Visszacsatolt hálózatok

52. Hogyan néz ki egy visszacsatolt neurális háló felépítése? Hogyan oldható meg a tanítás a backpropagation módszer segítségével (időbeli kibontás elmagyarázása), és mi ennek a megoldásnak a legnagyobb problémája (eltűnő/felrobbanó gradiens), és miért fordul ez elő?
53. Hogyan lehet az eltűnő gradiensket mérsékelni? Hogyan végzi ezt el az LSTM cella. Mit jelentenek az egyes belső változók és hogyan működnek? További változatok (GRU).
54. Milyen sorozatfeldolgozási feladatok lehetségesek? Milyen alkalmazásokat ismer (kérdés-válasz, videók, képfeliratozás, puha figyelem, stb.), és ezek hogyan működnek (ábra, leírás)

## 12. Detektálás és Szegegmentálás

55. Milyen architektúra alkalmas a szemantikus szegegmentálás elvégzésére, milyen alapvető trükköket érdemes ebben alkalmazni (fel- és leskálázás, előrecsatolás, dilatált konvolúció). Milyen módszerek léteznek a felskálázásra (Unpooling, transzponált konvolúció és DUC) és ezek miben különböznek?

56. Hogyan lehet lokalizációt elvégezni neurális hálókkal? Hogyan működik a Régió-CNN? Milyen módszerek léteznek a gyorsításra (Fast-R-CNN és Faster-R-CNN), és ezek mit újítottak? Hogyan lehet ezt az architektúrát objektum szegegmentálásra kiterjeszteni (Mask-R-CNN)?

57. Hogyan működik a YOLO architektúra, mi az alapvető különbsége a régió alapú detektálással szemben? Mi a detektálás és a megfelelő kimenet kiválasztásának alapelve? Mi az az anchor box és mire használjuk? Mi a YOLO módszer előnye/hátránya?

## III. 3D LÁTÁS

### 13. Kameramodell és kalibráció

58. Hogy néz ki a Pinhole kameramodell (rajz, egyenlet, paraméterek)? Mik azok a homogén koordináták, és miért alkalmazzuk őket (linearizálás, skálainvariancia, irányított végtelen)? Hogyan néz ki a vetítés mátrixa és mi mit jelent benne (belső külső paraméterek, egyes elemek)? Hogyan változtatja meg a vetítést a valódi optikák alkalmazása? Mi az a radiális torzítás és mi okozza?

59. Milyen geometriai transzformáció fajták léteznek, és mi köztük a különbség (benn foglalt transzformációk és invarianciák)? Melyik írja le ezek közül a képalkotás folyamatát?

60. Hogyan végezhetünk lineáris regressziót (LS becslést)? Mi történik, ha a kimenet nulla, ebben az esetben milyen módszerekkel oldható meg a feladat (melyik jobb és miért)? Mi a két fajta LS becslés közti szemléletes különbség? Polinomiális regresszió.

61. Mi a 3D és a Sakktáblás kalibráció célja, mit akarunk vele meghatározni (melyik paramétereket)? Mely paraméterek állnak elő a kalibráció során? Mik az eljárás lépései, és milyen módon végezhető el? Melyik lépésekben különbözik a kétfajta kalibrációs mód és hogyan? Mi az az algebrai és geometriai hiba, és melyiket szeretnénk igazából minimalizálni? Hogyan tesszük meg ezt? Hogyan építjük bele a radiális torzítás meghatározását a kalibrációba? Lehetséges-e 1D kalibrációs objektummal kalibrálni? És kalibrációs objektum nélkül? Ha igen, akkor mi kell hozzá?

62. Epipoláris geometria, mátrixok, epipoláris egyenes és pont. Mire tudunk olyan egyenletet felírni, hogy a kalibrációt el tudjuk végezni? Hogyan működik a 8 és a 7 pontos módszer? Utóbbi esetben honnan vesszük azt a plusz egy információt, ami miatt 7 pontpár is elég?

### 14. 3D Rekonstrukció

63. Mik a 3D rekonstrukció lépései? Mi a sztereó kalibráció célja, és mit akarunk vele meghatározni? Mit kell ismerni hozzá? Hogyan tudjuk elvégezni, ha fix kameráink vannak, és hogyan, ha mozoghatnak a kamerák. Mit veszítünk az utóbbi esetben?

64. Mi az a rektifikáció és miért csináljuk? Milyen két műveletet szoktunk benne összevonni? A rekonstrukció melyik lépésében számítjuk ki a transzformációt és mikor alkalmazzuk? Hogyan szoktuk alkalmazni egy képre (LUT)?

65. Mi az a diszparitás, mikor van értelme és hogyan értelmezzünk egy diszparitás képet? Hogyan működik a Block Matching algoritmus, és melyik korábban tanult eljárásra hasonlít? Mi a legnagyobb problémája és hogyan orvosolja ezt az SGBM algoritmus (szomszédosságban

konstansokkal büntetés)? Hogyan működik a Belief Propagation, mi az a hit, mit tartalmaznak az üzenetek, és milyen módon változtatjuk meg a hitet a szomszédok üzenetei alapján? Mi a közös ebben az eljárásban és a korábban ismertetett Expectation Maximization módszerben? Mi a Belief Propagation előnye és hátránya?

66. Hogyan lehet ismert pontokból és vetítés mátrixokból az eredeti térbeli pont 3D koordinátáit meghatározni? Hogyan módosul a feladat, ha sztenderd elrendezésünk és diszparitásunk van? Mi a metrikus rekonstrukció és milyen feltételek esetén tudjuk ezt előállítani. Milyen tipikus elrendezések esetén tudunk metrikusan és mikor csak skálainvariáns módon rekonstrukciót elvégezni?

67. Mi az a Structure from Motion és SLAM? Milyen változatai lehetnek (szekvenciális/online, nem szekvenciális/offline)? Mit értünk SLAM esetében Bundle Adjustment alatt, és miért csináljuk? Hogyan kapcsolódik a kalibráció során megtett finomító lépéshez? Hogyan kerüljük el SLAM esetén a driftet (hurok detekció, adatbázis)?

## 15. 3D feldolgozás

68. Milyen módon tudunk 3D struktúrákat tárolni (voxel, pontfelhő, mesh, paraméteres felületek, spline)? Mi a voxeles megoldás problémája? Milyen típusú pontfelhőink lehetnek? Mi az a kd-Fa reprezentáció, hogyan konstruáljuk, és mire jó alapvetően?

69. Milyen szűrési módszereket alkalmazhatunk 3D-ben? Hogyan működik a doboz szűrő? Mik azok az outlier elemek, és milyen módszerrel szűrhetők?

70. Hogyan tudunk szegmentálni pontfelhőkben, működnek-e a 2D-ben megismert szegmentáló eljárások? Hogyan használható a RANSAC néhány paraméterrel leírható alakzatok felismerésére? Hogyan néz ki a jelöltek állítása és kiértékelése általánosságban? 3D alakzatok felismerésére hogyan specializálható? Általános paraméterbecslés esetén milyen tulajdonságokkal rendelkezik a RANSAC? Mi az előnye/hátránya más módszerekkel (pl. LS becslés) szemben?

71. Milyen módszerekkel lehet jellemzőket generálni pontfelhőkben (lokális, globális)? Hogyan tudunk normálisokat számolni, és milyen kapcsolatban áll ez a főkomponens-analízissel? Milyen módon lehet összetett lokális leírókat rendelni az egyes pontokhoz (pl. PFH)? Hogyan működik a GASD globális leíró? Hogyan állítja be az objektumot globálisan konzisztens irányra, majd milyen jellemző segítségével írja le?

72. Mi a regisztráció fogalma, hogyan végezhető el az ismert algoritmusok segítségével?

73. Milyen neurális hálós megoldásokat ismer 3D feldolgozás esetére? Mi az alapvető különbség ezek között, és milyen előnyökkel/hátrányokkal szolgálnak? Hogyan lehet pontfelhőn konvolúció-szerű műveletet végezni?

## IV. A VALÓSIDEJŰ LÁTÁS ESZKÖZEI ÉS PARADIGMÁI

### 16. GPU

74. Milyen hardvereket alkalmazhatunk látás algoritmusok gyorsítására, ezek hogyan viszonyulnak felxibilitás és sebesség szempontjából? Milyen alapvető paradigmákat valósítanak ezek meg (több processzor, csővezeték, adatfolyam-feldolgozás)? Mi az a SISD, SIMD és MIMD feldolgozás, és melyik hardver melyiket támogatja alapvetően ezek közül? Mik a GPU alapú feldolgozás legfontosabb szűk keresztmetszetei?

75. Hogyan épül fel fizikailag egy GPU? Mi az a Streaming multiprocesszor, és milyen hardver elemekben osztoznak az egy SM-be tartozó magok? Hogyan néz ki egy SM általános felépítése, milyen típusú egységek találhatóak benne? Mi az a Tensor Core, milyen műveletet képes elvégezni és milyen alkalmazások számára kiemelten fontos?

76. Milyen egységek találhatók a CUDA futási modelljében (thread, block, grid, warp), hogyan viszonyulnak egymáshoz, és miért bontjuk a futást 3 szintre (mi közös az egy blokkon belül lévő szálakban)? Hogyan jön be a fonat (warp)? Mi a divergens kód, és miért érdemes kerülni? Hogyan kapcsolódik a futási modell felépítése a GPU fizikai kialakításához (SM-blokk)? Hogyan alkalmazható a csővezetés technika több szál párhuzamos futtatására?

77. Milyen memóriatípusok találhatók a GPU-ban? Ezek közül melyik statikus és melyik dinamikus RAM? Melyek helyezkednek el a GPU fő memóriájában fizikailag, és melyik találhatók az SM-ben? Mi a különbség az egyes memóriatípusok között sebesség és szálak általi elérés tekintetében? Hova kerülnek a változók alapértelmezett esetben?

78. Milyen programozási nyelvek léteznek GPU-kra, és ezekben mi a közös, illetve miben különböznek? Milyen módszerekkel szinkronizálhatunk szálak között? Mi a textúra memória, és milyen szolgáltatásokat nyújt számunkra (indexelés, szál-kezelés). Hogyan néz ki a CUDA kód fordítási mechanizmusa (milyen fordító, hova érdemes a projektben tenni a CUDA függvényeket, stb.)?

## 17. TPU, FPGA

79. Mi az a TPU, és milyen művelet elvégzésére alkalmas? Mi az a szisztolikus tömb, és hogyan képes a mátrixszorzást hatékonyan megvalósítani? Hogyan néz ki a TPU általános felépítése, mi változott az egyes verziók közt és miért?

80. Hogyan lehet különböző logikai áramkörü elemeket programozható hardverként megvalósítani? Ismertesse egy FPGA általános felépítését. Mire jók az egyes elemek? Mit lehet egyetlen CLB segítségével megvalósítani? Milyen memória típusok állnak rendelkezésünkre, és ezek közt mi a különbség?

81. Hogyan lehet FPGA segítségével a <küszöbözés/szürkeárnyalatosítás/konvolúció> műveletét megvalósítani? Nagyságrendileg mekkora késleltetést okoznak ezek a műveletek? Hogyan lehet futás közben adaptív módon változtatni a feldolgozási csővezetés struktúráját? Hogy néz ki az FPGA tervezés folyamata (konkrét tervezési csővezeték, hardver-szoftver együtt tervezés)?