

EÖTVÖS LORÁND TUDOMÁNYEGYETEM  
TERMÉSZETTUDOMÁNYI KAR

---

Pálffy Zsófia

ROBOTKAROK VIZUÁLIS VEZÉRLÉSE

BSc Alkalmazott Matematikus Szakdolgozat

Témavezető:

Dr. Szeghy Dávid

Geometria Tanszék



Budapest, 2018

# Köszönetnyilvánítás

Ez a dolgozat nem jöhetett volna létre konzulensem, Szeghy Dávid szakmai és emberi hozzáértése nélkül. Ha új ötlet vagy kérdés merült fel bennem, bátran fordulhattam hozzá, tudva, hogy meghallgat, segít végigvezetni a gondolatot, s akár többször, részletesen, türelmesen elmagyaráz mindent. Töretlen optimizmusával az utolsó pillanatokig tartotta bennem a lelket. Ezért végtelenül hálás vagyok.

# Tartalomjegyzék

Köszönetnyilvánítás .....	1
Bevezetés .....	3
1. A fényképezés matematikája .....	4
1.1. Centrális vetítés .....	4
1.2. Homogén koordinátázás.....	7
1.3. Egy speciális eset homogén koordinátázása .....	8
1.4. Projektív leképezések .....	8
1.5. Fotó és valóság .....	9
2 Valós reprezentáció .....	12
2.1. Képfüggvény, mintavétel.....	12
2.2. Leképezés.....	13
2.3. Téglatest és fényképe .....	15
2.4. Éldetektálási módszerek.....	15
2.3. Alakzatfelismerés.....	19
2.4. Zajcsökkentés.....	20
3. Robotgeometria .....	23
3.1. Robotkarok matematikai leírása .....	23
3.2. Robotikai feladatok.....	26
3.3. A síkbeli RRR robotkar .....	29
3.4. A térbeli PRRR robotkar inverz kinematikája .....	32
Összegzés .....	35
Irodalomjegyzék.....	36

# Bevezetés

A *robot* szó manapság nem jelent újdonságot, mindenkinek van fogalma róla, mit értünk alatta. A legtöbben talán először egy emberszerű gépre gondolnak, melyet filmben láttak, legyen ez jövőből érkező terminátor, földönkívüli autobot, vagy akár android egy messzi-messzi galaxisból; köszönhetően Karel Čapek 1921-es *R.U.R. (Rossumovi Univerzálni Robots – Rossum univerzális robotjai)* című színdarabjának, ahol először jelentek meg szintetikus anyagból készített robotoknak nevezett lények, amelyek a szó szláv eredetének megfelelően (szolgamunka) boldogan dolgoznak az emberekért. (Bár később fellázadnak és kitörlik az emberiséget.)

A robot tényleges mai definíciójában is megjelenik ez a szemlélet: elektromechanikai szerkezetet értünk alatta, amely képes különböző feladatok végrehajtására, de erre -még- csak előzetes programozás, vagy közvetlen emberi irányítás révén képes. Messze állunk attól, hogy a mesterséges intelligenciák az emberhez hasonló komplexitással gondolkozzanak, hiszen már egy olyan feladat megoldása, mint egy futószalagról szelektálás is rengeteg számolás eredménye, de az ilyenekhez hasonló robotkarok már egyre inkább elterjednek. Ennek oka, hogy a gyárakban ilyen relatív monoton, de nagy precizitást igénylő munkákban hatékonyabbak, mint az emberek. Emellett kisebb helyen elférnek és nem szükséges az extra védelmükre figyelni, mert nem fognak letérni a számukra kijelölt útvonalról.

Szakdolgozatom célja a fent említett selejtezési feladat egy megvalósítása mögötti matematika bemutatása. Tegyük fel, hogy egy gyárban dominókat gyártanak. A robot feladata, hogy azokat, amelyeknek rossz a festése – amikor egy hibahatáron túl eltér a teljesen tökéletes megvalósítástól -, kivegyje, és a kijelölt selejtgyűjtőbe helyezze. Ennek ellátásához szükség van az emberi látáshoz hasonló képességre, hiszen az érkező tárgyakról nem lehet előre tudni, pontosan hol és milyenek lesznek. A fényképezés modellje a centrális vetítés, ezért először ezt ismertetem. A kamera viszont nem helyezkedhet el pontosan a tárgy felett, tehát a megjelenő képből homogén koordinátázással, projektív transzformáció segítségével nyerjük vissza az eredeti alakot. Ezen az új képen kell megtalálni az éleket és a dominó pöttyeit, így felismerni az alakzatot. Ezután a megszerzett információk alapján a robotkar áthelyez a megfelelő helyre.

Végül összegzem a bemutatásra kerültek folyamatban betöltött szerepét, majd kitekintünk további megoldási, illetve felhasználási lehetőségekre.

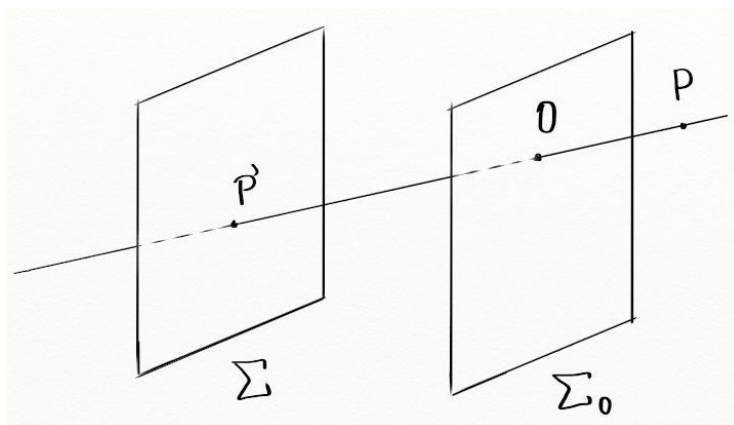
## 1. A fényképezés matematikája

A bemutatásra kerülő folyamat első lépése az, hogy a valós tárgyról digitális kép készül, amelyből aztán információkat nyerünk a tárgy formájáról. Ehhez először is ismernünk kell a háromdimenziós tárgy és kétdimenziós fotója közötti összefüggéseket, a leképezést, amely egyiket a másikba viszi.

### 1.1. Centrális vetítés

A fényképezés matematikai modellje a centrális vetítés, projektív térben. Ez egy tökéletes esetet modellez, ahol „hátrafelé is lát” a gép, és végtelen széles fotó készülhet.

**Definíció:** Adott  $O$  pontot és egy  $O$ -t nem tartalmazó  $\Sigma$  síkot véve a valós térben (ezt innenől képsíknak nevezzük) definiálhatunk egy  $\Phi_c: \mathbb{R}^3 \rightarrow \Sigma$  centrális vetítést úgy, hogy minden  $P$  ponthoz hozzárendeljük azt a pontot, ahol az  $OP$  egyenes döfi  $\Sigma$ -t.

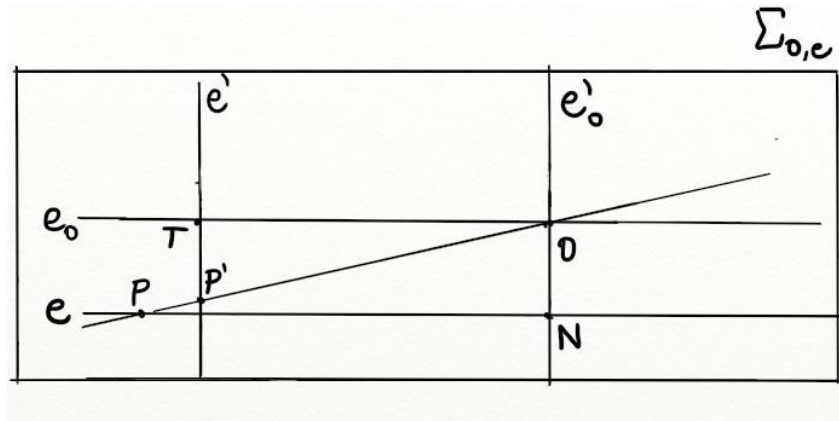


1. ábra Centrális vetítés

Ez nem mindenhol értelmezhető, hiszen az  $OP$  egyenes csak akkor létezik, ha  $O \neq P$ , de még ha létezik, akkor is csak akkor van metszéspont, ha nem párhuzamos a képsíkkal. Az ilyen egyenesek pontosan az  $O$ -ra illeszkedő, képsíkkal párhuzamos síkban helyezkednek el, ezt a továbbiakban  $\Sigma_0$ -val jelöljük.

**Definíció:** Ha a valós háromdimenziós síkot kibővítjük ideális térelemekkel úgy, hogy minden párhuzamos egyenesek által meghatározott ekvivalenciaosztályhoz rendelünk egy-egy ideális pontot úgy, hogy nem párhuzamos egyenesekhez különböző pontot rendelünk, akkor megkapjuk a projektív teret. Ebben a kibővített térben így a közönséges mellett ideális pontok is vannak. Az egyenesek a közönséges egyenesek az ideális pontjukkal, illetve egy-egy ideális egyenes tartozik a párhuzamos síkok ekvivalenciaosztályaihoz (ez egy tetszőleges síkkal párhuzamos egyenesek ideális pontjainak halmaza). A síkok a közönséges síkok, egyeneseik ideális pontjaival, emellett létezik az ún. ideális sík, amely az ideális pontokat tartalmazza.

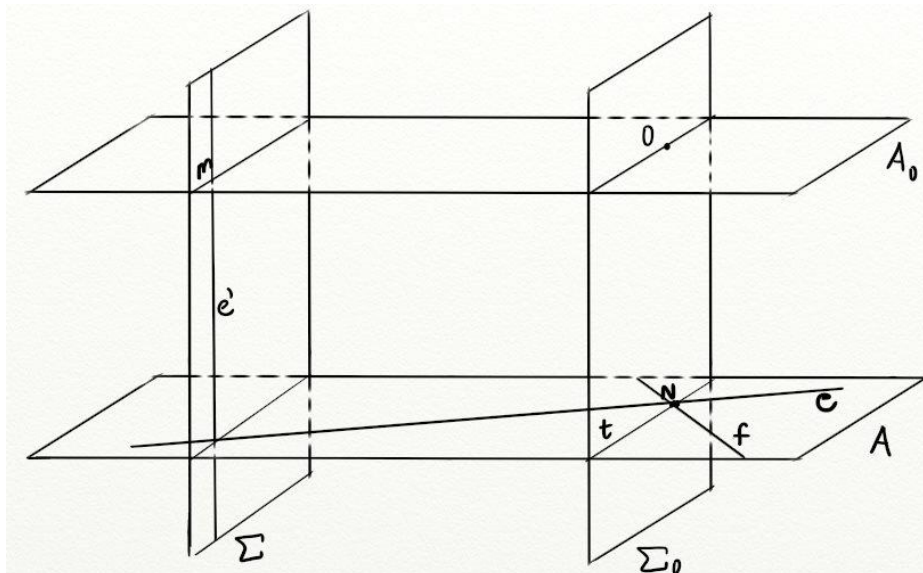
Projektív térre a következő módon terjeszthető ki a centrális vetítés:



2. ábra Egyenes képe

Vizsgáljuk az  $O$  vetítési centrumra és egy  $(O \notin)e$  egyenesre illeszkedő síkot, melyet  $\Sigma_{O,e}$ -vel jelölünk. Ebből a  $\Sigma$  képsík kimetsz egy egyenest, ezt jelöljük  $e'$ -vel. A továbbiakban az  $e$ -vel párhuzamos,  $O$ -ra illeszkedő egyenesre  $e_0$ -ként, az  $e'$ -vel párhuzamos,  $O$ -ra illeszkedő egyenesre  $e_0'$ -ként hivatkozunk.  $T$  jelöli  $e'$  és  $e_0$  metszéspontját,  $N$  az  $e$  és  $e_0'$  egyenesekét,  $e_\infty$  az  $e$  ideális pontját,  $e'_\infty$  az  $e'$ -ét. Ezen jelölések mellett látható, hogy  $e$  képe  $e'$  úgy, hogy  $N$  képe  $e'_\infty$ ,  $e_\infty$  képe pedig  $T$ .

Ha  $O$  része  $e$ -nek, akkor az egyenes képe az egyetlen  $\Sigma$ -val vett metszéspont. Ha  $e$  párhuzamos  $\Sigma$ -val, akkor ez nem más, mint  $e$  ideális pontja.

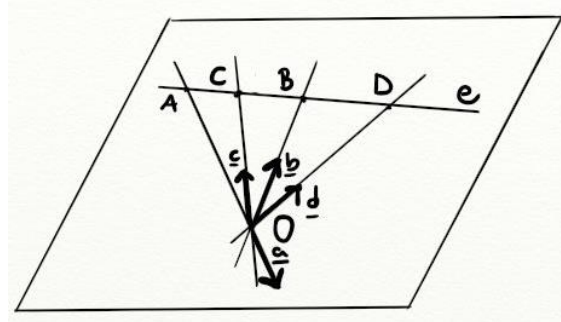


3. ábra Sík képe

Hasonló módon, ha egy  $A$  sík képét szeretnénk megkapni, a benne levő egyenesek képeit kell vizsgálni. Speciálisan a  $\Sigma$ -val párhuzamos,  $O$ -n átmenő sík ( $\Sigma_0$ ) és  $A$  metszete a  $t$ -vel jelölt egyenes, ennek képe a  $\Sigma$  ideális egyenese lesz, hiszen a tetszőleges pontját  $O$ -val összekötő egyenes párhuzamos  $\Sigma$ -val. Az adott  $N \in t$  ponton átmenő egyenesek képei így párhuzamosak

lesznek, mert ugyanaz az ideális pont tartozik hozzájuk. Hasonlóan, ha az  $A$ -val párhuzamos,  $O$ -t tartalmazó  $A_O$  sík és  $\Sigma$  metszetét tekintjük, vagyis az  $m$  egyenest, ott olyan pontok képei jelennek meg, melyek az  $A$ -beli ideális pontokhoz tartoznak, mert ezek  $O$ -val vett egyenesei párhuzamosak  $A$ -val. Ezzel  $O \notin A$  esetben egyértelmű a megfeleltetés  $A$  és  $\Sigma$  között.

Ha  $O \in A$ , akkor  $A' = m$ , ami  $A \parallel \Sigma$  esetén az ideális egyenesük.



4. ábra Kettőviszony

Definíció:  $A, B, C$  és  $D$  egy egyenes különböző pontjai a projektív térben,  $O$  tetszőleges közös pont. Legyenek  $\underline{a}, \underline{b}, \underline{c}$  és  $\underline{d}$  vektorok az  $AO, BO, CO, DO$  egyenesek irányvektorai, így semelyik kettő sem párhuzamos, tehát  $\underline{a}$  és  $\underline{b}$  bázist alkotnak, akkor ebben felírhatjuk  $\underline{c}$ -t és  $\underline{d}$ -t a következő módon:

$$\underline{c} = \gamma_1 \underline{a} + \gamma_2 \underline{b}, \underline{d} = \delta_1 \underline{a} + \delta_2 \underline{b}$$

Ekkor

$$(ABCD) = \frac{\gamma_2}{\gamma_1} : \frac{\delta_2}{\delta_1}$$

az  $A, B, C, D$  pontok kettőviszonya.

Számolással igazolható, hogy  $(ABCD)$  jól definiált, vagyis nem függ sem az  $O$  pont, sem az  $\underline{a}, \underline{b}, \underline{c}, \underline{d}$  vektorok megválasztásától.

$(ABCD)$  nem vehet fel 0-t vagy 1-t:

$$- (ABCD) = \frac{\gamma_2}{\gamma_1} : \frac{\delta_2}{\delta_1} = 0$$

pontosan akkor teljesül, ha  $\gamma_2 = 0$  vagy  $\delta_2 = 0$ , ami azt jelentené, hogy  $A = C$  vagy  $B = D$  az egyenesek irányvektorainak párhuzamossága miatt.

$$- (ABCD) = 1$$

esetén  $\frac{\gamma_2}{\gamma_1} = \frac{\delta_2}{\delta_1}$ , így

$$\frac{1}{\gamma_1} \cdot \underline{c} = \underline{a} + \frac{\gamma_2}{\gamma_1} \cdot \underline{b} = \underline{a} + \frac{\delta_2}{\delta_1} \cdot \underline{b} = \frac{1}{\delta_1} \cdot \underline{d}$$

tehát  $\underline{c}$  és  $\underline{d}$  párhuzamos,  $C = D$ .

Ezek ellentmondanak a definíciónak.

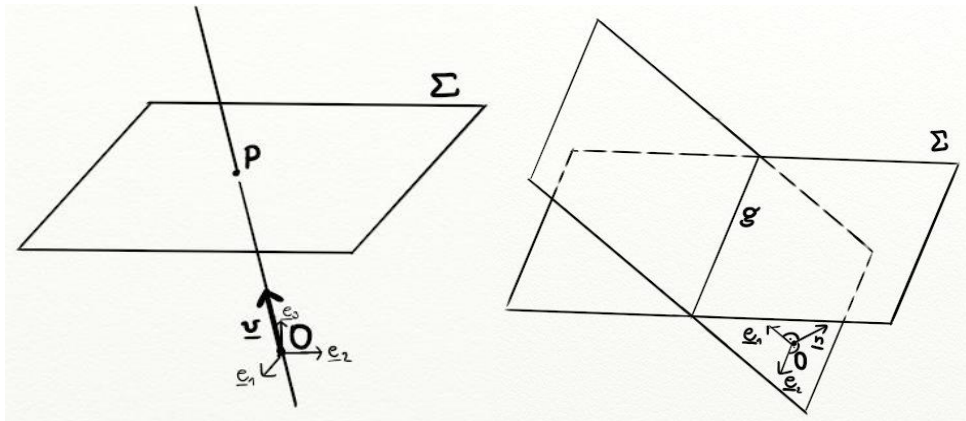
Ha adott  $\lambda = (ABCD) = \frac{\gamma_2}{\gamma_1} \cdot \frac{\delta_2}{\delta_1}$ , ahol  $\lambda \neq 0, 1$ , akkor  $\frac{\delta_2}{\delta_1} = \frac{\gamma_2}{\gamma_1} \cdot \frac{1}{\lambda}$ , ahol  $\frac{\gamma_2}{\gamma_1}$  csak az  $A, B, C$  pontoktól függ (közönséges pontok esetén ez az osztóviszonyuk). Tehát ha  $A, B, C$  ismert, akkor  $\frac{\gamma_2}{\gamma_1}$  is, így kiszámítható  $\frac{\delta_2}{\delta_1}$ , amely megadja  $\underline{d}$ -t, így az  $OD$  egyenes irányát is. Ennek az  $A, B$  és  $C$  egyenesével vett metszéspontja egyértelműen megadja  $D$ -t.

**Definíció:** Legyenek  $a, b, c$  és  $d$  négy különböző egyenes egy projektív síkon, melyek illeszkednek egy  $O$  pontra. Legyen  $e$  egy új egyenes ebben a síkban, amely nem illeszkedik  $O$ -ra.  $A, B, C, D$  legyenek rendre az egyenesek metszéspontjai  $e$ -vel. Ekkor  $(abcd) = (ABCD)$  az egyenesek kettősviszonya.

Papposz tétele segítségével igazolható, hogy  $(abcd)$  jól definiált, vagyis nem függ  $e$  választásától.

## 1.2. Homogén koordinátázás

Tekintsük  $V = \mathbb{R}^3 \setminus \{0\}$  vektorok halmazát. Ezen vezessük be azt az ekvivalenciarelációt, mely szerint  $v \sim w$  pontosan akkor, ha van olyan  $\lambda \neq 0$  skalár, melyre  $\underline{v} = \lambda \cdot \underline{w}$ . Jelölje  $[\underline{v}]$  a  $\underline{v}$  ekvivalenciaosztályát.



5. ábra Homogén koordinátázás

$\Sigma$  projektív sík a projektív térben,  $O$  közönséges pont,  $O \notin \Sigma$ . Minden síkbeli ponthoz hozzárendeljük az  $OP$  egyenest, ez kölcsönösen egyértelmű hozzárendelés a  $\Sigma$  pontjai és az  $O$ -ra illeszkedő egyenesek között. Az egyenesekhez kölcsönösen egyértelműen rendelhetjük az irányvektoraik ekvivalenciaosztályát, így ezeket kölcsönösen egyértelműen rendelhetjük a pontokhoz:  $P = [\underline{v}]$  jelöléssel. Amennyiben  $\underline{e}_1, \underline{e}_2, \underline{e}_3$  a vektortér egy bázisa, és  $\underline{v} = v_1 \underline{e}_1 + v_2 \underline{e}_2 + v_3 \underline{e}_3$ , akkor a  $P$  pont homogén koordinátái a  $(O, \underline{e}_1, \underline{e}_2, \underline{e}_3)$  rendszerben  $(\underline{v}_1 : \underline{v}_2 : \underline{v}_3)$ .



Ha ekvivalensnek tekintjük  $\lambda \neq 0$  esetben  $(\underline{v}_1: \underline{v}_2: \underline{v}_3)$  és  $(\lambda \underline{v}_1: \lambda \underline{v}_2: \lambda \underline{v}_3)$ -t, akkor egyértelműség is fennáll. Megjegyzendő, hogy a  $(0:0:0)$  hármas nem lesz homogén koordináta definíció szerint.

Ha  $g$  egyenes a  $\Sigma$  síkban, egyértelműen létezik  $\Sigma_{O,g}$  sík, amely  $O$ -ra és  $g$ -re illeszkedik. Ha ennek  $\underline{n}$  a normálvektora, akkor a  $g \leftrightarrow [\underline{n}]_e = [n_1: n_2: n_3]$  egyértelmű hozzárendelés adja meg a  $g$  egyenes homogén koordinátáit.

A fenti felállásban a  $(O, e_1, e_2, e_3)$  rendszert homogén koordinátarendszernek nevezzük.

### 1.3. Egy speciális eset homogén koordinátázása

Speciálisan, ha  $O$ -t  $(0,0,0)$ -nak választjuk és  $\Sigma = \{z = 1\}$ , akkor a  $\Sigma$ -beli  $P(x, y, 1)$  kölcsönösen egyértelműen megfeleltethető a Descartes-féle sík  $(x, y)$  pontjának. Szintén kölcsönösen egyértelmű megfeleltetés az  $(x: y: 1)$  homogén koordináta is, összefoglalva:

$$(x, y) \leftrightarrow (x, y, 1) \leftrightarrow (x: y: 1)$$

$$\text{Descartes} - \text{sík} \leftrightarrow \Sigma \text{ közönséges része}$$

Ha valamely pont homogén koordinátája  $(x: y: z)$ , akkor

- ha  $z \neq 0$ , akkor a  $(\frac{x}{z}, \frac{y}{z})$  közönséges pontnak,
- ha  $z = 0$ , akkor a  $\underline{v}(x, y)$  irányú egyenes ideális pontjának felel meg.

Ha  $e \equiv ax + by + cy = 0$  egy egyenes, akkor  $e$  homogén koordinátái  $(a: b: c)$  ebben a rendszerben.

### 1.4. Projektív leképezések

$\Sigma_1$  és  $\Sigma_2$   $\mathbb{P}^3$ -beli síkok,  $O_1 \notin \Sigma_1, O_2 \notin \Sigma_2$  közönséges pontok,  $(\underline{e}_1, \underline{e}_2, \underline{e}_3)$  bázis.  $[\underline{v}]_i$ -vel jelöljük egy  $P \in \Sigma_i$  pont a  $(\Sigma_i, O_i, \underline{e}_1, \underline{e}_2, \underline{e}_3)$  rendszerben felírt homogén koordinátáit.

Legyen  $L: V \rightarrow V$  egy lineáris leképezés, amely invertálható, így  $GL(3)$ -beli, a fenti bázisban  $3 \times 3$ -as mátrixként felírható. Ez definiál egy  $[L]_{1,2}: \Sigma_1 \rightarrow \Sigma_2, [\underline{v}]_1 \mapsto [L\underline{v}]_2$  leképezést.

Definíció: Ha  $\Phi: \Sigma_1 \rightarrow \Sigma_2$  projektív síkok között olyan leképezés, amely egyenestartó, (ezért valós esetben) invertálható és kettősvizonytartó, akkor projektív leképezés.

Állítás: Ha  $L$  invertálható lineáris leképezés, akkor  $[L]_{1,2}$  projektív leképezés lesz.

A linearitásból könnyen látszik az egyenestartás, ahonnan a kettősviszonytartás definíció alapján közvetlenül megvalósul, az invertálhatóság pedig  $L$  invertálhatóságából jön.

**Tétel:** Ha  $\Phi: \Sigma_1 \rightarrow \Sigma_2$  projektív leképezés és  $O_1, O_2$  közös pontok, amelyek az azonos indexű síkoknak nem részeik, akkor létezik olyan  $L: V \rightarrow V$  invertálható lineáris leképezés, amelyre  $\Phi = [L]_{1,2}$ .

Ezt szokás a projektív geometria alaptételének is nevezni. (Bizonyítást ld. [1].) Megjegyzendő, hogy  $L$  és  $M$  pontosan akkor valósítja meg ugyanazt a leképezést, ha egymás skalárszorosai. Ez a vektorok ekvivalenciájából könnyen látszik.

**Definíció:** Jelölje az  $n$  dimenziós tér projektív leképezéseinek csoportját  $PGL(n+1)$ . Ekkor  $PGL(n+1) \cong GL(n+1)/\sim$

ahol

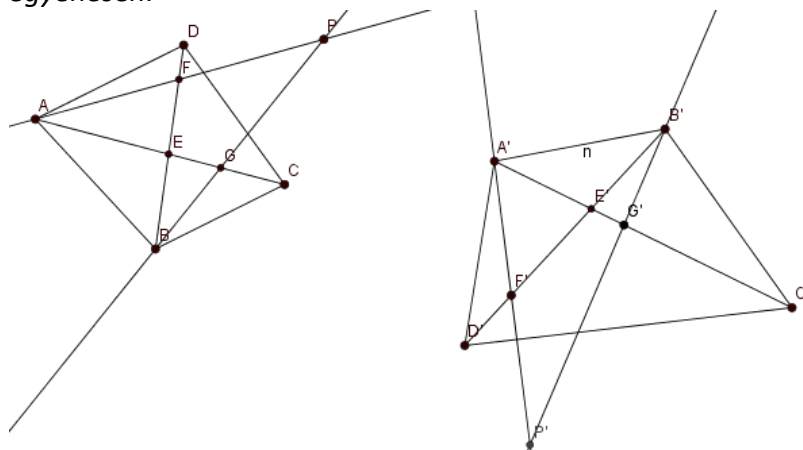
$$N \sim L \Leftrightarrow M = \alpha L \quad (0 \neq \alpha \in \mathbb{R})$$

Ezek szerint egy  $\Phi: \Sigma_1 \rightarrow \Sigma_2$  projektív síkok közötti projektív leképezés leírható egy  $L \in \mathbb{R}^{3 \times 3}$  mátrixszal való szorzásként.

### 1.5. Fotó és valóság

Az előzőekben megismerhettük a leképezést, amely a valóságot levetíti egy síkra, és láttuk, hogy síkok pontjait megfeleltethetjük egymásnak úgy, hogy fennálljon a kettősviszonytartás. Belátható, hogy ha tudjuk, hogy projekció áll fenn, akkor ha van négy általános helyzetű pontunk (ld. alább), amiknek ismerjük a helyzetét mindkettő síkon – például egy fényképen és a valóság felülnézeti merőleges vetületén, egy térképen -, akkor tetszőlegesen választva pontot az egyik síkból, meg tudjuk mutatni, hogy az hol helyezkedik el a másikon.

**Definíció:** A  $P_1, P_2, \dots, P_n$  ( $n \geq 3$ ) pontok általános helyzetűek, ha közülük semelyik három nincs egy egyenesen.



6. ábra Ötödik pont projektív képének szerkesztése négy ismert esetén

A 6. ábrán látható módon helyezkedik el az  $ABCD$  négyszög, melynek képe az  $A'B'C'D'$  négyszög.  $P$  pont képét keressük. Mivel a projektív leképezés egyenes-, és metszésponttartó, ezért behúzza az  $AC$  és  $BD$  átlókat, a metszéspontjukban levő  $E$  pont képe egyértelműen adódik. Az  $AP$  és  $BP$  egyenesek metszéspontjai rendre  $F$  és  $G$ . Ha kiszámoljuk a  $BD$  egyenesen levő pontok kettősviszonyát, az megegyezik a  $B'D'$  egyenesen levőkével, így az egyértelműség révén megkapjuk  $F'$ -t. Hasonlóan kapjuk  $G'$ -t az  $AP$  egyenes alapján. Innen  $P'$  már egyszerűen jön, hiszen az  $A'F'$  és  $B'G'$  egyenesek metszéspontjában található.

Ezt a példánkban úgy használjuk majd ki, hogy a dominóról készült fotón azonosítva a felső lap négy csúcsát, rekonstruáljuk a felülnézeti képet. Ebből tudjuk majd ellenőrizni, jó-e a festés (a megfelelő felülnézeti mintákhoz hasonlítva).

Ha egy konkrét  $\Phi: \Sigma_1 \rightarrow \Sigma_2$  projektív leképezés mátrixszát szeretnénk megkapni, ahol ismerjük  $A, B, C, D \in \Sigma_1$  pontok  $A', B', C', D' \in \Sigma_2$  képeit, akkor elég az  $A, B, C$  pontokat  $A', B', C'$ -be vivő mátrixot megtalálni, hiszen a háromdimenziós térben 3-3 pont ismerete elég egy invertálható lineáris leképezés megadásához.

Legyen  $O \notin \Sigma_i$  ( $i = 1, 2$ ),  $(\underline{e}_1, \underline{e}_2, \underline{e}_3)$  ortonormált bázis. Mivel három dimenzióban vagyunk, ezért a 4 helyvektor összefügg. Ekkor van olyan lineáris kombinációja  $\underline{a} = \overrightarrow{OA}$ ,  $\underline{b} = \overrightarrow{OB}$  és  $\underline{c} = \overrightarrow{OC}$ -nek, melyre

$$\lambda_1 \underline{a} + \lambda_2 \underline{b} + \lambda_3 \underline{c} = \underline{a}^* + \underline{b}^* + \underline{c}^* = \underline{d}$$

Ekkor homogén koordinátáknak megfelelnek az  $\underline{a}^*, \underline{b}^*, \underline{c}^*$  fenti skalárszorosokból képzett vektorok is, hiszen páronként ugyanabba az ekvivalenciaosztályba tartoznak. Ugyanígy választhatjuk  $A', B', C'$  homogén koordinátáit is úgy, hogy a (vesszővel jelölt) irányvektoraik összege kiadja  $D'$  irányvektorát. Mivel a keresett  $L$  leképezés lineáris, ezért

$$\underline{d}' = \underline{a}' + \underline{b}' + \underline{c}' = L\underline{a}^* + L\underline{b}^* + L\underline{c}^* = L(\underline{a}^* + \underline{b}^* + \underline{c}^*) = L\underline{d}$$

tehát ezzel a negyedik pont is az előírt helyére kerül.

Legyenek  $M = (\underline{a}^* \quad \underline{b}^* \quad \underline{c}^*) \in \mathbb{R}^{3 \times 3}$ ,  $N = (\underline{a}' \quad \underline{b}' \quad \underline{c}') \in \mathbb{R}^{3 \times 3}$  a megfelelő oszlopvektorok révén képzett mátrixok. Ekkor  $M^{-1}$  annak a leképezésnek a mátrixa, amely az  $(\underline{a}^*, \underline{b}^*, \underline{c}^*)$  vektorokat az  $(\underline{e}_1, \underline{e}_2, \underline{e}_3)$  bázisvektorokba,  $N$  pedig azé, amelyik ezeket  $(\underline{a}', \underline{b}', \underline{c}')$ -be viszi, tehát  $L = NM^{-1}$  lesz a keresett transzformációs mátrix.

A dominó festésének ellenőrzéséhez tehát a következőket kell tennünk: egy mintadarabkal kalibrálunk, ezzel bemérjük a felső lap síkját, ami minden dominónál megegyezik. Ebből a síkból szeretnénk felülnézeti képet csinálni. A mintadarab négy felső csúcsának ismerjük mindkét síkbeli helyzetét, így egy tetszőleges további pont átvihető oda. Elkészíthetjük a fenti  $L$  transzformációs mátrixot is, amely majd az új dominók felső lapjának pontjait átviszi ebbe a felülnézeti síkba.

Ezt az új képet fogjuk összevetni minden lehetséges helyes festés képével. Ez olyan dominók esetén, melyeken mindkét oldalon 0-6 pötty lehet,  $\frac{7 \cdot 6}{2}$  olyan lehetőséget jelent, ahol különböző számok vannak a két oldalon. Ezek fordítva is állhatnak, tehát a szimmetriaközéppontra középpontosan tükrözöttek még egyszer ennyit vizsgálatot jelentenek, így az egyforma pöttyszámúakkal együtt összesen  $42 + 7 = 49$  mintával kell végrehajtani az ellenőrzést. Minden ilyenre pixelenkénti összehasonlítást végzünk, közben számolva az egyezéseket. Ezt leosztjuk az összpixelszámmal, hogy megkapjuk a hasonlóság mértékét, s akkor jelöljük a dominót selejtesnek, ha nincs olyan kép, amellyel legalább 97%-os egyezést mutatna. Amint találtunk legalább ilyen jót, az algoritmus leállhat, elfogadjuk a festés helyességét.

## 2 Valós reprezentáció

Ahhoz, hogy a matematikai modellt átvigyük a valóságba, figyelembe kell venni az olyan lényeges különbségeket elmélet és gyakorlat között, melyek a kamera fizikai megvalósításából adódnak: a képsík véges, az adatokat diszkrét módon tároljuk, pixeleken.

### 2.1. Képfüggvény, mintavétel

**Definíció:** Képfüggvénynek nevezzük a fénykép matematikai reprezentációját, ez általában vektorértékű, kevés argumentummal.

Speciálisan (a mi esetünkben is) diszkrét képfüggvényt használunk, ahol a változók és a függvényérték is egészek. Ugyanazt a képet több függvény is leírhatja, használatuk aszerint választandó, milyen tulajdonság a legfontosabb a vizsgálat szempontjából. A legtöbb függvény kétdimenziós vektorváltozójú,  $f(\underline{x}) = f(x, y)$  alakú, ahol  $(x, y)$  a fénykép, mint sík koordinátázásából adódik (diszkrét módon, táblázatszerűen osztjuk fel). Az  $f(x, y)$  érték jelentheti a (fekete-fehér) kép szürkeségét egy feketétől fehérig terjedő skálán.

Ahhoz, hogy a diszkrét adatokat kezelni tudjuk, le kell tudnunk írni a folytonos függvényeket diszkrét módon. Ehhez a delta disztribúciót használjuk, mely a következő módon definiált:

$$\delta(x) = \begin{cases} 0, & \text{ha } x \neq 0 \\ \infty, & \text{ha } x = 0 \end{cases}$$
$$\int_{-\infty}^{\infty} \delta(x) dx = 1$$

Ezt megkaphatjuk  $\delta_n: \mathbb{R} \rightarrow \mathbb{R}$  függvények  $n \rightarrow \infty$  limeszeként, ahol

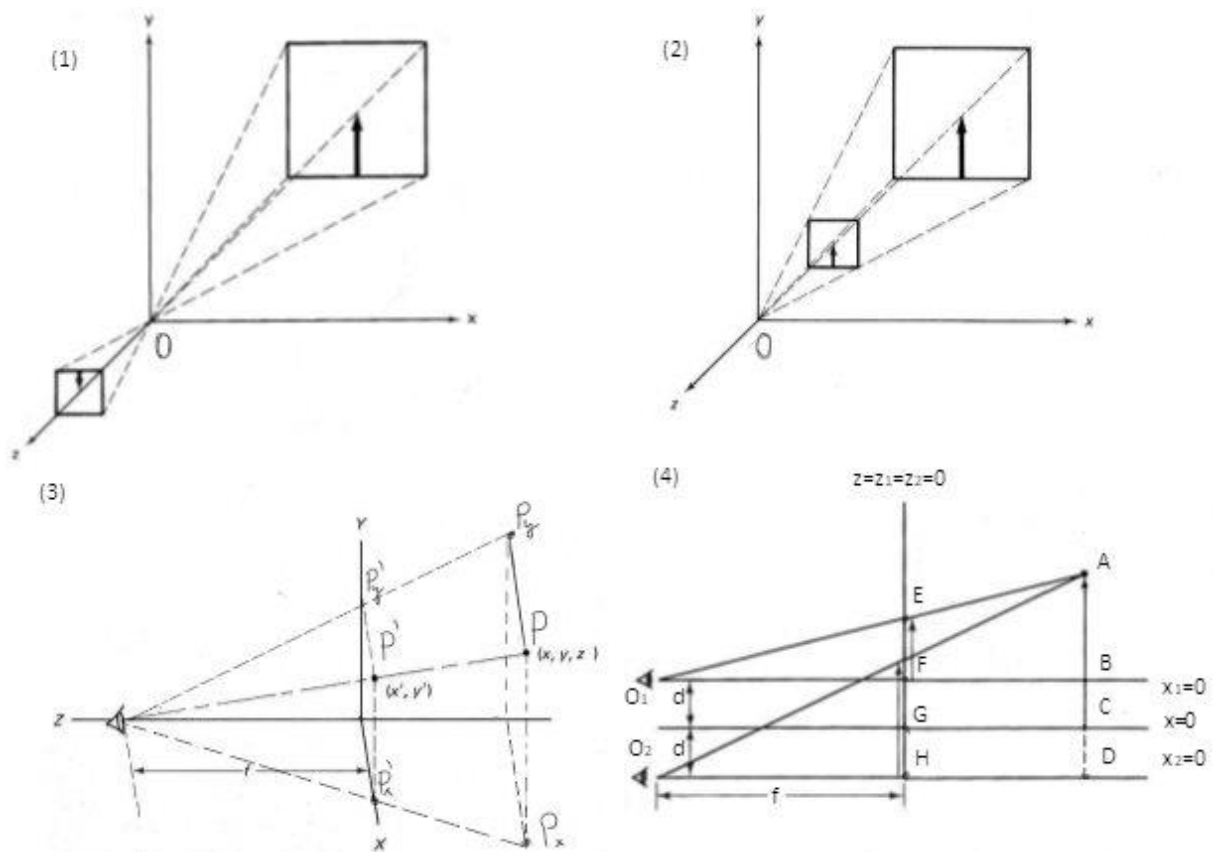
$$\delta_n(x) = \begin{cases} n, & \text{ha } |x| < \frac{1}{2n} \\ 0, & \text{ha } |x| \geq \frac{1}{2n} \end{cases}$$

Ennek segítségével diszkrét mintákat nyerhetünk, mivel egy  $f: \mathbb{R} \rightarrow \mathbb{R}$  folytonos függvény  $a$  pontbeli értékét a következő integrál adja meg:

$$\int_{-\infty}^{\infty} f(x)\delta(x - a) dx = f(a)$$

Ez több dimenzióra is hasonló módon oldható meg, ekkor  $\mathbb{R}^n$ -en kell integrálni.

## 2.2. Leképezés



7. ábra Kamera koordinátázási lehetőségek. (1) A fókuszpont az origó, a kép megfordul (2) A fókuszpont az origó, a kép nem fordul meg (3) A fókuszpont koordinátái  $(0,0,f)$  (4) Két nézőpont, a fókuszpontok  $(d,0,f)$  és  $(-d,0,f)$  [3]

Az előző fejezetben bemutatott projektív leképezés megvalósítása a gyakorlatban a következő módon zajlik (ld. [2]):

- A fényképezőgép fókuszpontja a vetítési középpont.
- Ettől  $f$  (fókusz)távolságra található a CCD, vagyis a képsík, ez adja a  $z$  tengely irányát.
- A 7. ábrán látható (1)-es modellen a fókuszpont az origó,  $\Sigma = \{z = f\}$  és a kép fejjel lefelé fordul, a (2)-es számmal jelzett, módosított verzióban a kép egyenesen áll, tükröztünk az origóra. Ezt eltolhatjuk, hogy fókuszunk a  $(0,0,f)$  pontot véve a kép a  $\Sigma = \{z = 0\}$  síkba kerüljön, ez látható a (3)-as számmal jelzett modellen.

Ekkor  $OPP_x$  és  $OP'P'_x$  háromszögek hasonlóságából, és ezek  $\{y = 0\}$ , illetve  $\{x = 0\}$  síkra vetítéséből adódnak a következő összefüggések:

$$\frac{y}{f-z} = \frac{y'}{f} \text{ és } \frac{x}{f-z} = \frac{x'}{f}$$

Bár a  $z$  komponens a képtérben mindenhol 0, ezen is végrehajthatjuk ugyanazt a leképezést, mint az  $x$  és  $y$  irányokban, így az előzőek alapján  $(x, y, z)$  képe:

$$(x', y', z') = \left( \frac{fx}{f-z}, \frac{fy}{f-z}, \frac{fz}{f-z} \right)$$

Ez azért hasznos, mert így információt tárolunk az eredeti pont fókuszról vett távolságáról is, ami alapvetően nem mondható el egy projektív leképezés képét vizsgálva, hiszen egy  $P \in \Sigma$  pont az  $OP$  egyenes tetszőleges pontjának képe lehetne. Ez viszont csak akkor működik, ha egy olyan síkról készítünk fotót, amely a CCD-vel párhuzamosan helyezkedik el.

De ha az emberi szemhez hasonlóan kettő nézőpontból tekintünk egy tárgyra, akkor kialakul a „mélységérzet”, vagyis számolhatók a  $z$  irányú távolságok.

Ez a következő módon valósítható meg (ld. 7.ábra (4)-es képe):

- Az első vetítés centruma  $O_1 = (d, 0, f)$ , a másodiké  $O_2 = (-d, 0, f)$ , mindkettő képe  $\Sigma_1 = \Sigma_2 = \Sigma = \{z = 0\}$ .
- $\Sigma_i$ -t úgy koordinátázzuk, hogy az  $x_i$  tengely az  $O_i$ -n áthaladó  $x$  tengellyel párhuzamos egyenes legyen. Az  $y_i$  tengelyt ehhez igazítjuk úgy, hogy  $z = z_1 = z_2$  választással  $x_i, y_i, z_i$  jobbsodrású derékszögű koordinátarendszer legyen.

Ekkor  $A$  koordinátáira a különböző rendszerekben a következők teljesülnek:

$$x_1 = \frac{(x-d)f}{f-z} \text{ és } x_2 = \frac{(x+d)f}{f-z}$$

mivel  $ABO_1$  és  $EFO_1$ , illetve  $ADO_2$  és  $EHO_2$  hasonló háromszögek. Másképp:

$$(f-z)x_1 = f(x-d)$$

$$(f-z)x_2 = f(x+d)$$

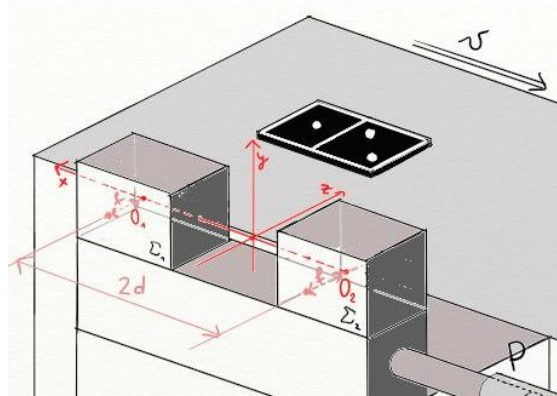
Ezeket egymásból kivonva  $z$  kifejezhető:

$$(f-z)(x_1 - x_2) = f(2d)$$

$$z(x_2 - x_1) = f(x_2 - x_1) - 2fd$$

$$z = f - \frac{2fd}{x_2 - x_1}$$

### 2.3. Téglatest és fényképe



8. ábra A kamerák elhelyezése

Tegyük fel most, hogy van két egyforma kameránk, ezekkel szeretnénk egy téglatest alakú tárgyról pontosan megmondani, hogy hol van – vagyis milyen a dominó futószalagon való elhelyezkedése. Ehhez először is úgy kell elhelyeznünk a két kamerát, ahogy az előző fejezetben láttuk: egy síkba essenek a CCD-k. Válasszuk úgy, hogy ez párhuzamos legyen a futószalag haladási irányával. Ekkor egy egyenesre esnek a fókuszok is, távolságuk  $2d$ . Ez az egyenes párhuzamos lesz az  $x$  tengellyel, elérve, hogy a képsík az  $XY$  síknak megfelelően legyen koordinátázva, még hozzá diszkrétén: egy pixel szélessége lesz egy egység, egész koordinátájú pontokként hivatkozhatunk a képpontokra. Innen megkapjuk  $d$ -t,  $f$ -et.

Az előző alfejezet alapján a megfelelő csúcsokat azonosítva azok valós  $(x, y)$  koordinátái kiszámolhatók a két képhez tartozókból  $((x_1, y_1)$  és  $(x_2, y_2))$ :

$$z = f - \frac{2fd}{x_2 - x_1}$$

$$x = \frac{(f - z)x_1}{f} + d = \left(1 - \frac{z}{f}\right)x_1 + d$$

$$y = \frac{y_1 x}{x_1}$$

Ezek azok a valóságbeli koordináták, melyekre az 1.5 fejezetben számolt  $L$  leképezést haddatjuk majd, hogy felülnézeti képet kapjunk.

### 2.4. Éldetektálási módszerek

Az eddigiek alapján tudjuk, hogy ha egy fotón felismerünk egy alakzatot, az hol helyezkedik el a térben. De azt még nem tisztáztuk, hogy veszi észre egy gép, hogy hol van a képen a lényegi rész. Erről szól ez az alfejezet.



A fő gondolat az, hogy ha a kép egy sávjában levő szürkeségi értékekből képzett függvényben ugrás figyelhető meg, akkor ott jó eséllyel valaminek a szélét látjuk. Ahhoz, hogy ezt az ún. lokális élet megkapjuk, azt is meg kell vizsgálni, milyen szögben a legnagyobb az eltérés, mert ez lesz az él ottani érintőjére merőleges irány.

Különböző matematikai operátorokat - vagyis azok számítógépes megvalósításait - használhatjuk ezen ugrások megtalálására. Ezek közül a leggyakrabban használtak azok, amelyek a *gradiens* vagy a *Laplace* operátort közelítik, hiszen ezekben különböző (iránymenti) deriváltak szerepelnek, amelyek egy függvény különböző pontok körüli megváltozásáról adnak vissza információt.

Definíció: Egy  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  függvény gradiense a következő differenciáloperátor:

$$\text{grad}(f) = \nabla f = \sum_{i=1}^n \partial_i f \underline{e}_i \in \mathbb{R}^n$$

ahol  $\partial_i f$  az  $f$  függvény  $i$ . parciális deriváltját,  $\underline{e}_i$  pedig az  $i$ . koordinátának megfelelő irányú egység hosszú vektort jelöli. A leképezés csak azon pontokban értelmezhető, ahol minden parciális derivált létezik.

A diszkrét megvalósításhoz legyenek

$$\begin{aligned} \Delta_1 &= f(x+n, y) - f(x, y) \\ \Delta_2 &= f(x, y+n) - f(x, y) \end{aligned}$$

ahol  $n$  egy kicsi pozitív szám, a lépésköz. Törekedni kell, hogy úgy válasszuk meg, ne legyen túl kicsi, mert akkor érzékeny lesz  $f(x)$  lényegtelen változásaira is, de nagy se, mert akkor nem találja meg a tényleges éleket. Leggyakrabban  $n = 1$ -et választunk, amikor gyakorlatilag pixelenként haladunk végig a képen, mivel ez a legegyszerűbb, de ez alapvetően a kép minőségétől függ, hiszen például egy homályosabb kép esetén érdemesebb nagyobbat lépni az él megvastagodása miatt, de akkor is, ha szemcsés a kép, ami például fénymásolatoknál, vagy szkennelt állományok esetén gyakoribb a porszemek, odahulló hajszálok stb. miatt. Ha viszont tudjuk, hogy részletgazdag, jó minőségű fényképen dolgozunk, és minden élet meg szeretnénk tartani, akkor megfelelő az egyesével haladás.

Definíció: Egy  $f(x)$  képfüggvényre a gradiens  $s(x)$  nagysága és  $\Phi(x)$  iránya a következő módon számolható:

$$\begin{aligned} s(x) &= (\Delta_1^2 + \Delta_2^2)^{\frac{1}{2}} \\ \Phi(x) &= \text{atan2}(\Delta_2, \Delta_1) \end{aligned}$$

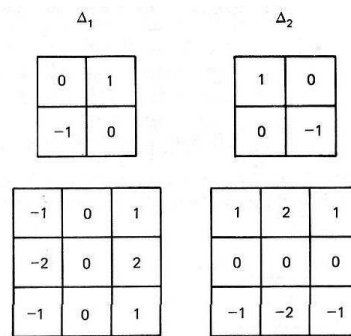
ahol az  $\text{atan2}$  azt a programozásban használatos függvényt jelöli, amely visszaadja  $\theta$ -t, az  $XY$  sík  $(x, y) \neq (0,0)$  pontjához tartozó vektor és az  $x$  tengely által bezárt szöget a következő képlet alapján:

$$\text{tg}(\theta) = \frac{y}{x}$$

úgy, hogy figyelembe veszi, a pont melyik síknegyedbe esett:

- ha  $x > 0$ , akkor  $\text{atan2}(y, x) = \arctan\left(\frac{y}{x}\right) \in \left]-\frac{\pi}{2}; \frac{\pi}{2}\right]$
- ha  $x < 0$  és  $y < 0$ , akkor  $\text{atan2}(y, x) = \tan^{-1}\left(\frac{y}{x}\right) \in \left]-\pi, -\frac{\pi}{2}\right]$
- Ha  $x < 0$  és  $y > 0$ , akkor  $\text{atan2}(y, x) = \tan^{-1}\left(\frac{y}{x}\right) \in \left]\frac{\pi}{2}, \pi\right]$

Itt  $s(x)$  tulajdonképpen a távolságoperátor,  $\Delta_1$  és  $\Delta_2$  pedig a legegyszerűbb közelítései az  $x$ , illetve  $y$  irányú deriváltaknak. Ehelyett használhatók olyanok, amelyek kevésbé érzékenyek a „zajra”, vagyis a kép egyes részeinek indokolatlan részletességére, nem valós élekre.



9. ábra A Roberts és a Sobel operátorok [3]

Ilyen a *Roberts operátor* (ld. [2], [3]), amely pixelek  $2 \times 2$ -es négyzetein alapul, és az átlók menti első deriváltat közelíti:

$$\Delta_1 = f(x + 1, y) - f(x, y)$$

$$\Delta_2 = f(x, y + 1) - f(x, y)$$

A *Sobel operátor* (ld. [2], [3]) már  $3 \times 3$ -as négyzet alapú, és bár súlyozottak az  $x$  és  $y$  irányok, az átlókat is figyelembe veszi:

$$\Delta_1 = (f(x + 1, y - 1) + 2f(x + 1, y) + f(x + 1, y + 1)) - (f(x - 1, y - 1) + 2f(x - 1, y) + f(x - 1, y + 1))$$

$$\Delta_2 = (f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1)) - (f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1))$$

Ezek kiküszöbölik azt, hogy egy-egy porszem fényképét lényegesnek tekintsünk, de egy valóban lyukacsos felületen elveszíthetjük a lyukak körvonalát, mint éleket. Ezért fontos úgy beállítanunk a lépésközt, hogy felhasználjuk előzetes ismereteinket a fényképezett környezetről, amennyiben vannak ilyenek.

Definíció: A  $\Delta$  Laplace operátor egy  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  függvényhez a tiszta második deriváltak összegét rendeli:

$$\Delta f = \operatorname{div}(\operatorname{grad}(f)) = \sum_{i=1}^n \partial_i^2 f \in \mathbb{R}$$

Ennek egy diszkrét megvalósítása az  $f(x, y)$  képfüggvény esetén a következő:

$$L(x, y) = \frac{(f(x, y+1) + f(x, y-1) + f(x+1, y) + f(x-1, y))}{4} - f(x, y)$$

Ez a második deriváltak közelítéséből jön, mivel:

$$\partial_x f\left(x + \frac{1}{2}, y\right) \approx \frac{f(x+1, y) - f(x, y)}{(x+1) - (x)} = f(x+1, y) - f(x, y) \text{ és}$$

$$\partial_x f\left(x - \frac{1}{2}, y\right) \approx \frac{f(x, y) - f(x-1, y)}{(x) - (x-1)} = f(x, y) - f(x-1, y)$$

$$\partial_x^2 f(x, y) \approx \frac{\partial_x f\left(x + \frac{1}{2}, y\right) - \partial_x f\left(x - \frac{1}{2}, y\right)}{\left(x + \frac{1}{2}\right) - \left(x - \frac{1}{2}\right)} = f(x+1, y) - 2f(x, y) + f(x-1, y)$$

hasonlóan  $y$ -ra:

$$\partial_y^2 f(x, y) \approx (f(x, y+1) - 2f(x, y) + f(x, y-1))$$

Innen

$$\Delta f(x, y) \approx -4f(x, y) + f(x, y+1) + f(x, y-1) + f(x+1, y) + f(x-1, y)$$

amit ha normálásként leosztunk 4-gyel, megkapjuk  $L(x, y)$ -t.

Az eljárás hátránya, hogy egyrészt nem ad irányinformációt, másrészt a második deriváltak becslése miatt érzékenyebb a zajra, mint a kizárólag első deriváltakkal operálók.

Viszont ezzel vékonyabbá tehetjük a gradiensmódszerrel talált éleket, ha azokat a pixeleket tartjuk csak meg, ahol  $L(x, y)$  egy adott korlátnál kisebb értéket vesz fel, ugyanis a gradiens lokális szélsőhelyein a Laplace 0-t vesz fel.

A gradiensmódszer egy további igen speciális esete az, amikor úgy tekintjük, az élek a pixelek között, lépcsőszerűen helyezkednek el, vagyis minden képpontról eldönthető, hogy pontosan melyik tartományba esik. Az irányok  $\frac{\pi}{2}$  többszöröseit, az él súlyát az  $|f(\underline{x}) - f(\underline{y})|$  különbség adja, ahol  $\underline{x}$  és  $\underline{y}$  az élet határoló pixelek koordinátái. Hátrány, hogy a kapott élek ívei elvesznek, nyersnek lesznek.

### 2.3. Alakzatfelismerés

Általánosan ahhoz, hogy egy robotkar meg tudja valósítani feladatát, vagyis végrehajthassa a selejtezési folyamatot, fel kell tudni ismerni a tárgyat, ami elé került. Vagyis nem elég az éleket megtalálni, hanem meg kell tudni mondani, pontosan milyen alakzatot látunk.

Ennek egy módszere az ún. momentumok kiszámítása (ld. [2]). Ezek a kép egy régióján értelmezettek, és kétdimenziós alakzatok régióon belüli elhelyezkedésének, formájának karakterizálására használhatók. Következésképp számítható egy  $(x, y)$  diszkrét pontokat tartalmazó  $R$  régió:

$$m_{i,j} = \sum_{(x,y) \in R} f(x,y) x^i y^j$$

Ha élhetünk a feltételezéssel, hogy a kép tisztán fekete-fehér, és binárisan kódolva a háttér 0-s, az alakzat 1-es értéket kapott, akkor ez a következőre egyszerűsödik:

$$m_{i,j} = \sum_{(x,y) \in R} x^i y^j$$

Eszerint  $m_{0,0}$  pontosan egyszer számolja össze a régió pixeleit, tehát annak területét adja meg, míg  $m_{1,0}$  az  $x$  irányú szélességet,  $m_{0,1}$  az  $y$  irányút. Így

$$\bar{x} = \frac{m_{1,0}}{m_{0,0}} \text{ és } \bar{y} = \frac{m_{0,1}}{m_{0,0}}$$

a régióhoz tartozó ún. centroid koordinátái, ami a képtérben való elhelyezkedéséről nyújt információt. Ez tulajdonképpen a tömegközéppontot jelenti, ha a régióra kétdimenziós merev testként tekintünk. Ekkor  $m_{0,0}$  a test tömegének felel meg.

Mivel az előzőekben  $m_{i,j}$  értéke függ attól, hogy hol helyezkedik el a régió a képtérben (minél messzebb az origótól, annál nagyobb), ezért helyette gyakran a centrális momentumokat használják, melyek a következő módon számíthatók:

$$\mu_{i,j} = \sum_{(x,y) \in R} (x - \bar{x})^i (y - \bar{y})^j$$

Ez már invariáns az eltolásra.

A fizikai hozzáállásnál maradva a következő mátrix a test középpontjának  $x, y$  tengelyekre vett tehetetlenségi nyomaték tenzora:

$$I = \begin{pmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{pmatrix}$$

Mivel ez valós, szimmetrikus mátrix, ezért diagonalizálható, így a következő alakra hozható:

$$I = \begin{pmatrix} I_1 & 0 \\ 0 & I_2 \end{pmatrix}$$

ahol  $I_1$  és  $I_2$  -  $I$  sajátértékei - a fő tehetetlenségi nyomatékok. A diagonalizáló derékszögű koordináta-rendszer tengelyeit tehetetlenségi főtengeknek hívjuk.

Ha a testnek vannak szimmetriatengelyei, akkor ezek egybeesnek a főtengekkel. Ez sok esetben lehet hasznos, például üvegek helyzetének megállapításához vagy az elkapási szélesség számolásához. Jelen esetben nem erre használjuk, mert ahhoz az 1.5. fejezetben kifejtett módszer elegendő információt ad.

#### 2.4. Zajcsökkentés

Az előző fejezetekben felmerült annak a problémája, hogy mivel általában a képek nem teljesen egyárnyalatú részekből állnak, a függvényérték megváltozása szinte sehol nem lesz nulla. Így feltételezhető lenne, hogy a kapott kisebb értékek elhanyagolhatók éldetektálási szempontból, tehát ha az él ereje az  $(x, y)$  pontban kisebb valamely küszöbnél, akkor elhanyagoljuk. Ez számítástechnikailag könnyen megvalósítható, viszont a gyakorlat szerint nem a legjobb. Ehelyett Frei és Chen módszerét (ld. [4]) érdemes használni:

Itt is  $3 \times 3$ -as maszkokkal dolgozunk, tehát szükségünk van az ekkora képrészletek formai leírására, amit  $\mathbb{R}^{3 \times 3}$ -as mátrixokkal tehetünk meg. Ezeket természetes módon átalakíthatjuk  $\mathbb{R}^9$ -beli vektorokká:

$$b = \begin{pmatrix} b_4 & b_3 & b_2 \\ b_5 & b_0 & b_1 \\ b_6 & b_7 & b_8 \end{pmatrix} \leftrightarrow \underline{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_8 \end{pmatrix}$$

A  $3 \times 3$ -as képrészletek kilencdimenziós vektortérén, amit jelölhetünk  $V$ -vel, bázist alkotnak a következő mátrixok:

$$G_1 = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{pmatrix}$$

$$G_2 = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{pmatrix}$$

$$G_3 = \frac{1}{2\sqrt{2}} \begin{pmatrix} 0 & -1 & \sqrt{2} \\ 1 & 0 & -1 \\ -\sqrt{2} & 1 & 0 \end{pmatrix}$$

$$G_4 = \frac{1}{2\sqrt{2}} \begin{pmatrix} \sqrt{2} & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -\sqrt{2} \end{pmatrix}$$

$$G_5 = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$G_6 = \frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

$$G_7 = \frac{1}{6} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$$

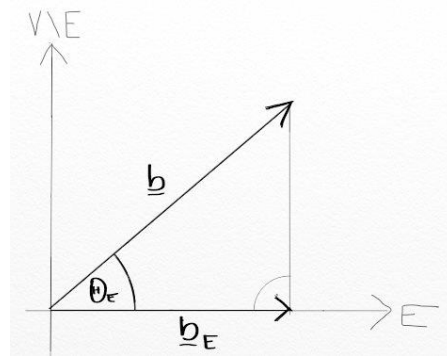
$$G_8 = \frac{1}{6} \begin{pmatrix} -2 & 1 & -2 \\ 1 & 4 & 1 \\ -2 & 1 & -2 \end{pmatrix}$$

$$G_9 = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Itt a  $G_1, G_2, G_3$  és  $G_4$  mátrixok  $\underline{v}_1, \dots, \underline{v}_4$ -gyel jelölt vektor reprezentációik kifeszítik az éltek alterét, ez legyen  $E$ . Ez azt jelenti, hogy ha egy  $3 \times 3$ -as képrészleten megjelenik egy él, akkor a 9 dimenziós reprezentációja ebben az altérben, vagy a közelében helyezkedik el. Ezért az lesz az eljárás, hogy erre levetítjük a  $\underline{b}$  vektorral reprezentált képrészletet, majd az eredeti

és a projekció ( $\underline{b}_E$ ) egymással bezárt szögének nagysága alapján döntünk, hogy tartalmazott-e élet. A szöget a következő összefüggésből kaphatjuk meg:

$$\cos(\theta_E) = \frac{|\underline{b}_E|}{|\underline{b}|} = \frac{\sqrt{\sum_{i=1}^4 \langle \underline{v}_i; \underline{b} \rangle^2}}{\sqrt{\sum_{i=1}^9 \langle \underline{v}_i; \underline{b} \rangle^2}}$$



ahol  $\langle \underline{a}; \underline{b} \rangle$  a szokásos skalárszorzást jelöli.

Minél pontosabb az egyezés, annál kisebb a szög, tehát az algoritmus használatakor rögzítünk egy  $\tau$  küszöbértéket, végighaladunk a kép belső pontjain (a rájuk illesztett  $3 \times 3$ -as mozaikokon), és ezekre kiszámoljuk  $\theta_E$ -t. Amelyik pontra  $\tau \geq \theta_E$ , azt élbelinek tekintjük.

Ugyanez a módszer a  $v_5, \dots, v_8$  vektorok által kifeszített  $L$  altérre használva egyenesek detektálására is alkalmas, ekkor a bezárt szöget a

$$\theta_L = \cos^{-1} \sqrt{\frac{\sum_{i=5}^8 \langle \underline{v}_i; \underline{b} \rangle^2}{\langle \underline{b}; \underline{b} \rangle}}$$

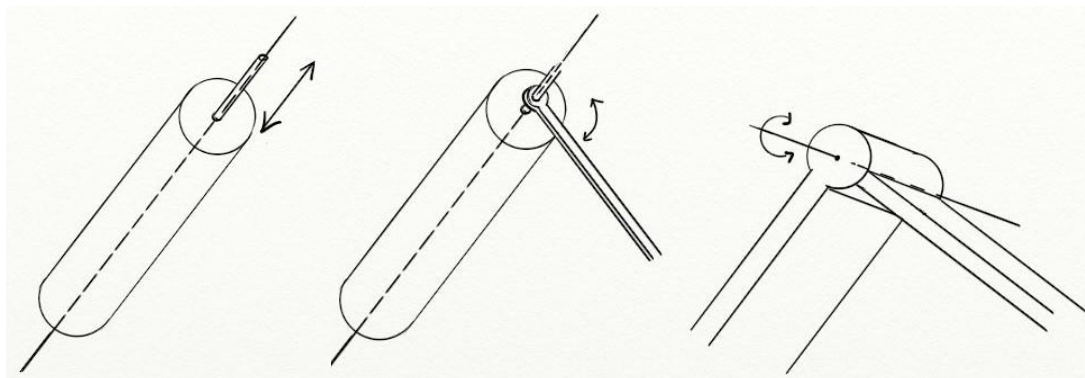
formulából kapjuk, s egy  $T$  küszöb felett fogadjuk el, hogy a képrészleten áthalad egy egyenes.

### 3. Robotgeometria

A selejtezés tényleges megvalósítását, tehát a hibás tárgyhoz kinyúlást, annak megfogását, majd egy jelölt helyre vitelét egy robotkar végzi. Ebben a fejezetben meglátjuk, milyen matematikai modellt használhatunk ezek leírására, hogyan paraméterezzük és irányítsuk, hogy végrehajtsa ezt a feladatot.

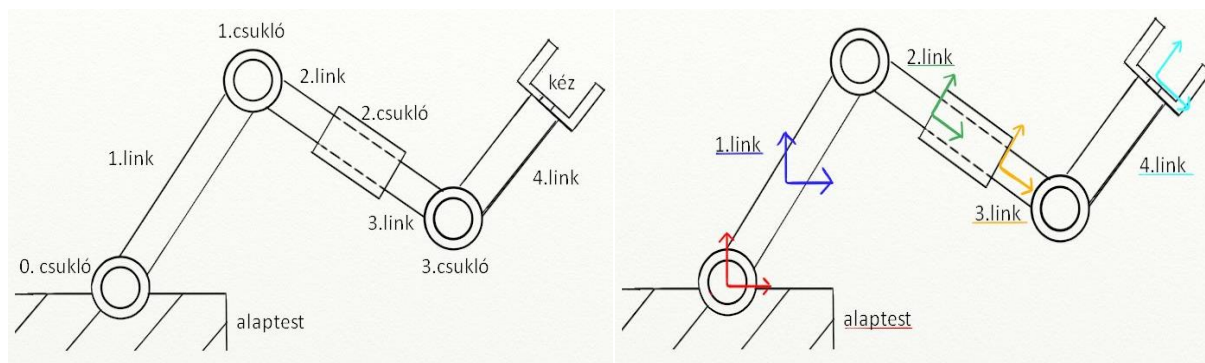
#### 3.1. Robotkarok matematikai leírása

**Definíció:** Robotkarnak nevezünk egy olyan vezérelhető tárgyat, amely merev alaptestéhez csuklókon keresztül merev szegmensek (linkek) csatlakoznak. Az alaptestet a 0. csukló köti az 1. linkhez, amely másik végén az 1. csukló a 2. linkhez stb. Az utolsó linket a robot kezének nevezzük. A robotkar nyílt vagy zárt, aszerint, hogy a kéz megegyezik-e az alaptesttel. Mi a nyílt esettel fogunk foglalkozni.



11. ábra Ízülettipusok

Egy csukló lehet forgató ízület (R), ez egy ún. hatástengely körüli forgatást biztosít, vagy eltoló ízület (P), amelynek csak párhuzamosság erejéig egyértelmű a tengelye. Speciálisan egy gömbcsukló felfogható 3 mozgató ízületként, melyeket 0 hosszú link köt össze.



12. ábra Robotkar részei és illesztett koordinátarendszerek

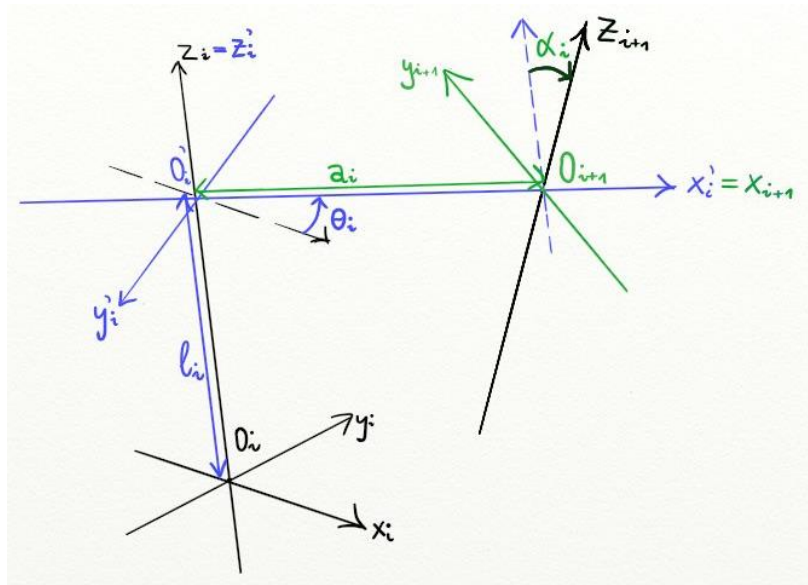
Ahhoz, hogy bármilyen feladatot megoldhassunk, precízen le kell tudni írni a robotkar helyzetét. Ennek leggyakrabban használt módja a Denavit- Hartenberg konvenció (ld. [1]). Eszerint olyan koordinátarendszereket veszünk, melyekre a következők teljesülnek:



- minden merev részhez tartozik egy jobbsodrású, páronként merőleges tengelyekből álló koordináta-rendszer, az alaptesthez a  $(O_0, x_0, y_0, z_0)$ , az  $i$ -hez  $(O_i, x_i, y_i, z_i)$ ,
- az  $i$ . koordináta-rendszerben a  $z_i$  tengely az  $i$ . csukló hatástengelye,
- az  $x_i$  tengely merőleges  $(y_i$  és  $z_i$ -n kívül)  $z_{i-1}$ -re is, metsző módon.

Állítás: Minden robotkarhoz létezik ilyen koordinátázás.

Bizonyítás: Indukcióval. A 0. csukló hatástengelyét véve megkapjuk  $z_0$  irányított tengelyt, ezen tetszőlegesen felvehető az  $O_0$  origó, melyen keresztül  $z_0$ -ra merőlegest húzunk, ezt irányítva kapjuk  $x_0$ -t. Innen egyértelműen kapjuk meg  $y_0$ -t, amely merőleges a tengelyekre, áthalad  $O_0$ -on és a koordináta-rendszert jobbsodrásúvá teszi. Ha az  $i - 1$ . koordináta-rendszer már adott, a következő módon készíthető el az  $i$ .:  $z_i$ -t meghatározza az  $i$ . csukló hatástengelye. Ha  $z_{i-1}$  és  $z_i$  kitérők (nem párhuzamosak), akkor egyértelműen létezik egy normál transzverzálisuk, amely mindkettőt metszi és merőleges rájuk, ez lesz  $x_i$ , a metszéspont pedig  $O_i$ . Innen  $y_i$  egyértelműen adódik a feltételekből. Ha  $z_{i-1}$  és  $z_i$  párhuzamosak, akkor több lehetőség is van  $x_i$ -re, általában úgy választjuk, hogy az adott karhoz legkézenfekvőbbben illeszkedjen (pl. szimmetriatengely); majd  $O_i$  és  $y_i$  az előző módon kaphatók. Metsző tengelyek esetén a metszésponton keresztül egyértelműen húzható merőleges egyenes, innen a szokott módon kapjuk  $O_i$  és  $y_i$ -t. Így sorban felépíthetők a koordináta-rendszerek minden robotkar esetén.



13. ábra Koordináta-rendszerek közötti áttérés

Ezek a koordináta-rendszerek úgy hasznosak a gyakorlatban, ha a kéznél levőben megadott koordináták és az alaptesthez viszonyított koordináták egymásból számíthatók. Ehhez először vizsgáljuk meg, hogyan térhetünk át az  $(O_i, x_i, y_i, z_i)$  rendszerből az  $(O_{i+1}, x_{i+1}, y_{i+1}, z_{i+1})$ -be:

Ehhez segéd koordináta-rendszert definiálunk: legyen  $O_i'$  az  $x_{i+1}$  és  $z_i$  metszéspontja, ekkor  $x_i' = x_{i+1}$  és  $z_i' = z_i$  mellett  $y_i'$  az egyértelműen létező  $O_i'$ -n átmenő tengely, mellyel  $x_i', y_i', z_i'$  jobbsodrású. Az  $(O_i, x_i, y_i, z_i) \rightarrow (O_i', x_i', y_i', z_i')$  áttérés megadható az  $O_i$  és  $O_i'$  előjeles távolságával ( $l_i$ ) vett eltolás és az  $x_i$  és  $x_i'$  tengelyek szögével ( $\theta_i$ ) vett forgatás kombinációjaként. Az  $(O_i', x_i', y_i', z_i') \rightarrow (O_{i+1}, x_{i+1}, y_{i+1}, z_{i+1})$  áttérés  $O_i'$  és  $O_{i+1}$  előjeles távolságával ( $a_i$ ) vett eltolásnak és  $z_i'$  és  $z_{i+1}$  szögével ( $\alpha_i$ ) vett forgatásnak felel meg. Tehát  $(O_i, x_i, y_i, z_i) \rightarrow (O_{i+1}, x_{i+1}, y_{i+1}, z_{i+1})$  megadható 4 adattal:  $(l_i, \theta_i, a_i, \alpha_i)$ . Ezek közül 3 adat fix, hiszen a hatástengelyek  $a_i$  távolsága, illetve  $\alpha_i$  szöge sem eltoláskor, sem forgatásnál nem változik. Ha egy ízület forgató, akkor az őt körülvevő merev szegmensek  $x_i$  tengelyei nem fognak egymástól távolodni, így  $l_i$ -t tudjuk, ha eltoló,  $\theta_i$  nem változik; mert ez a hozzá kapcsolódó szegmensek szöge, ami szintén már a konstrukcióból adódik.

Így ha  $(u_{i+1}, v_{i+1}, w_{i+1})$  egy pont koordinátái az  $(O_{i+1}, x_{i+1}, y_{i+1}, z_{i+1})$  rendszerben, akkor a

$$\begin{pmatrix} 1 & 0 & 0 & a_i \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha_i & -\sin \alpha_i & 0 \\ 0 & \sin \alpha_i & -\cos \alpha_i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_{i+1} \\ v_{i+1} \\ w_{i+1} \\ 1 \end{pmatrix} = \begin{pmatrix} u_i' \\ v_i' \\ w_i' \\ 1 \end{pmatrix}$$

képlettel adható meg az  $(O_i', x_i', y_i', z_i')$  rendszerben, hiszen  $\alpha_i$ -vel forgatunk az első tengely körül, majd  $a_i$ -vel tolunk el ugyanez mentén. Innen az  $(O_i, x_i, y_i, z_i)$  rendszerbe a

$$\begin{pmatrix} \cos \theta_i & -\sin \theta_i & 0 & 0 \\ \sin \theta_i & -\cos \theta_i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & l_i \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_i' \\ v_i' \\ w_i' \\ 1 \end{pmatrix} = \begin{pmatrix} u_i \\ v_i \\ w_i \\ 1 \end{pmatrix}$$

képlet segítségével írhatók át a koordináták, mert  $\theta_i$ -vel forgatunk a harmadik tengely körül, majd ementén  $l_i$ -vel kell eltolni. Tehát ha  $A$ -val jelöljük az első leképezésbeli mátrixok szorzatát és  $B$ -vel a másodikban levőket, akkor az áttérés a következő módon adható meg:

$$\begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = BA \begin{pmatrix} x_{i+1} \\ y_{i+1} \\ z_{i+1} \\ 1 \end{pmatrix}$$

ahol

$$BA = \begin{pmatrix} \cos \theta_i & -\cos \alpha_i \sin \theta_i & \sin \theta_i \sin \alpha_i & a_i \cos \theta_i \\ \sin \theta_i & \cos \alpha_i \cos \theta_i & -\cos \theta_i \sin \alpha_i & a_i \sin \theta_i \\ 0 & \sin \alpha_i & \cos \alpha_i & l_i \\ 0 & 0 & 0 & 1 \end{pmatrix} = M_i(l_i, \theta_i, a_i, \alpha_i)$$

az áttérési mátrix. Az alaprendszerbeli koordináták így számolhatók, ha az  $n$ . link a robot keze:

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{pmatrix} = M_0(l_0, \theta_0, a_0, \alpha_0) M_1(l_1, \theta_1, a_1, \alpha_1) \dots M_{n-1}(l_{n-1}, \theta_{n-1}, a_{n-1}, \alpha_{n-1}) \begin{pmatrix} x_n \\ y_n \\ z_n \\ 1 \end{pmatrix}$$

### 3.2. Robotikai feladatok

Az előzőekben láttuk, hogy  $4n$  paraméter kódolja a robotkar állását. Ez a megoldása az ún. direkt kinematikai feladatnak, vagyis így számolható a csuklók helyzetét megadó paramétereiből a robotkéz pozíciója.

Az inverz kinematikai feladat célja ezzel szemben a kéz pozíciójából kikövetkeztetni a csuklók paramétereit. Ez nem mindig egyértelmű. Látni fogjuk, hogy például a két egy síkba eső forgató ízületből álló kar esetén is, ha az első csukló nem kinyújtott helyzetben van, akkor a kezét és 0. csuklót összekötő egyenesre tükrözött karállás azonos kézhelyzetet eredményez.

Sebességkinematika esetén adott a csuklóparaméterek egy sima függvénye, azaz a paraméterek változtatásával mozgást valósítunk meg, a feladat az, hogy megadjuk a kéz sebességét egy adott pillanatban.

Az inverz sebességkinematikai feladatnál egy adott pozícióban van a kéz, ismerjük a csuklóparamétereket. Úgy szeretnénk megváltoztatni a paramétereket, hogy a kéz pont egy adott sebességgel mozduljon el. Ennek folytonos változata, ha a kéz útját egy időtől függő  $\gamma(t)$  görbével adjuk meg, és keressük az  $l_1(t), \theta_1(t), a_1(t), \alpha_1(t), l_2(t), \theta_2(t), a_2(t), \alpha_2(t), \dots, l_n(t), \theta_n(t), a_n(t), \alpha_n(t)$  paraméterek folytonos függvényeit, melyekre a robot keze bejárja a  $\gamma(t)$  pályát. A következőkben láthatjuk, hogy nincs szükség az összes kiszámítására.

Állítás: Az inverz kinematikai feladat megoldásához elég megoldani a sebességkinematikát.

Bizonyítás: Konstrukcióval.

A direkt feladat megoldása a csuklóparaméterek függvényében ilyen alakú:

$$\begin{pmatrix} x_0(t) \\ y_0(t) \\ z_0(t) \\ 1 \end{pmatrix} = M_0(l_0(t), \theta_0(t), a_0(t), \alpha_0(t)) \cdot \dots \cdot M_{n-1}(l_{n-1}(t), \theta_{n-1}(t), a_{n-1}(t), \alpha_{n-1}(t)) \begin{pmatrix} x_n \\ y_n \\ z_n \\ 1 \end{pmatrix}$$

ahol a  $4n$  paraméterből csak  $n$  változhat, hiszen egy link hossza nem változik, egy adott ízület pedig vagy forog, ami  $\theta$  változását jelenti, vagy eltol, ami  $l$ -ét. Mindez a jobboldali szorzások után felírható ilyen alakban:

$$\begin{pmatrix} x_0(t) \\ y_0(t) \\ z_0(t) \\ 1 \end{pmatrix} = \begin{pmatrix} F_x(l_0(t), \dots, \alpha_{n-1}(t)) \\ F_y(l_0(t), \dots, \alpha_{n-1}(t)) \\ F_z(l_0(t), \dots, \alpha_{n-1}(t)) \\ 1 \end{pmatrix}$$

A kar sebességét a koordinátafüggvények deriváltjaiból kapjuk:

$$\begin{pmatrix} x_0'(t) \\ y_0'(t) \\ z_0'(t) \\ 0 \end{pmatrix} = \sum_{i=0}^{n-1} M_0(\dots) \cdot \dots \cdot \frac{d}{dt} M_i(l_i(t), \theta_i(t), a_i(t), \alpha_i(t)) \cdot \dots \cdot M_{n-1}(\dots) \cdot \begin{pmatrix} x_n \\ y_n \\ z_n \\ 1 \end{pmatrix},$$

ez a másik felírás szerint, a többváltozós deriválás szabályai alapján:

$$\begin{aligned} \begin{pmatrix} x_0(t) \\ y_0(t) \\ z_0(t) \\ 1 \end{pmatrix}' &= \begin{pmatrix} F_x(l_0(t), \dots, \alpha_{n-1}(t)) \\ F_y(l_0(t), \dots, \alpha_{n-1}(t)) \\ F_z(l_0(t), \dots, \alpha_{n-1}(t)) \\ 1 \end{pmatrix}' = \\ &= \begin{pmatrix} \frac{dF_x}{dl_0}(l_0(t), \dots, \alpha_{n-1}(t)) & \frac{dF_x}{d\theta_0}(l_0(t), \dots, \alpha_{n-1}(t)) & \dots & \frac{dF_x}{d\alpha_{n-1}}(l_0(t), \dots, \alpha_{n-1}(t)) \\ \frac{dF_y}{dl_0}(l_0(t), \dots, \alpha_{n-1}(t)) & \frac{dF_y}{d\theta_0}(l_0(t), \dots, \alpha_{n-1}(t)) & \dots & \frac{dF_y}{d\alpha_{n-1}}(l_0(t), \dots, \alpha_{n-1}(t)) \\ \frac{dF_z}{dl_0}(l_0(t), \dots, \alpha_{n-1}(t)) & \frac{dF_z}{d\theta_0}(l_0(t), \dots, \alpha_{n-1}(t)) & \dots & \frac{dF_z}{d\alpha_{n-1}}(l_0(t), \dots, \alpha_{n-1}(t)) \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} l_0'(t) \\ \theta_0'(t) \\ \vdots \\ \alpha_{n-1}'(t) \end{pmatrix} = \\ &= \begin{pmatrix} J(t) \\ \underline{0} \end{pmatrix} \cdot \underline{v}(t) \end{aligned}$$

ahol  $J$   $3 \times 4n$ -es Jacobi mátrix  $t$  függvényében, amely minden  $n$  csuklós robotkarra ilyen alakú, csupán a függvényeibe írt paraméterek változnak. Sőt,  $\underline{v}(t) \in \mathbb{R}^{4n}$ -ből  $3n$  tag 0, hiszen

ezek konstansok deriváltjai, csak a maradék  $n$  db módosulhat. Így a Jacobi mátrixnak is csak a nemnulla tagokhoz tartozó  $n$  db oszlopára van szükségünk. Ezzel megoldottuk a direkt sebességkinematikai feladatot.

Az inverz kinematikai feladat megoldásához tegyük fel, hogy  $(l_0, \theta_0, \dots, \alpha_{n-1})$  a kar paraméterei, és el akarjuk vinni az  $(x_1, y_1, z_1)^T$  pontba. A direkt kinematikai feladatból tudjuk a jelenlegi helyzetét, ezt jelölje  $(x_0, y_0, z_0)^T$ . Legyen  $\gamma(t) = (\gamma_x(t), \gamma_y(t), \gamma_z(t))$  olyan görbe, amire  $\gamma(0) = (x_0, y_0, z_0)$  és  $\gamma(1) = (x_1, y_1, z_1)$ . Tekintsük a következő differenciálegyenletet:

$$\begin{pmatrix} \gamma_x'(t) \\ \gamma_y'(t) \\ \gamma_z'(t) \end{pmatrix} = J(l_0(t), \theta_0(t), \dots, \alpha_{n-1}(t)) \cdot \begin{pmatrix} l_0'(t) \\ \theta_0'(t) \\ \vdots \\ \alpha_{n-1}'(t) \end{pmatrix} = J(t) \cdot \underline{v}(t)$$

ahol  $J(t) \in \mathbb{R}^{3 \times 4n}$ , de valójában csak  $n$  oszlop számít, hiszen bár  $\underline{v}(t) \in \mathbb{R}^{4n}$ , de ebből csak  $n$  tag nemnulla, hiszen egy csukló mozgásakor csak egy paramétere változik meg a hozzá tartozó 4-ből, a többi konstans  $t$  függvényében. Így ez egy elsőrendű differenciálegyenlet  $n$  függvényre,  $t = 0$  kezdeti feltétel mellett, mert a kezdőpillanat paraméterei ismertek. Ez akkor biztosan megoldható, ha minden pillanatban megoldható az inverz sebességkinematikai feladat, azaz az  $n$  db nemnulla paraméterhez tartozó oszlop kifeszíti  $\mathbb{R}^n$ -et. Ekkor létezik a paramétereknek olyan folytonos függvénye, amely a robotkezet a kívánt helyzetbe viszi.

Definíció: Azon pontok halmazát, ahova a robot keze elér, munkatérnek nevezzük.

Definíció: Az olyan  $4n$ -es  $(l_0, \theta_0, \dots, \alpha_{n-1})$  paramétervektort, amihez tartozó helyzetből a robotkar nem minden irányba mozdítható, szinguláris pozíciónak nevezzük.

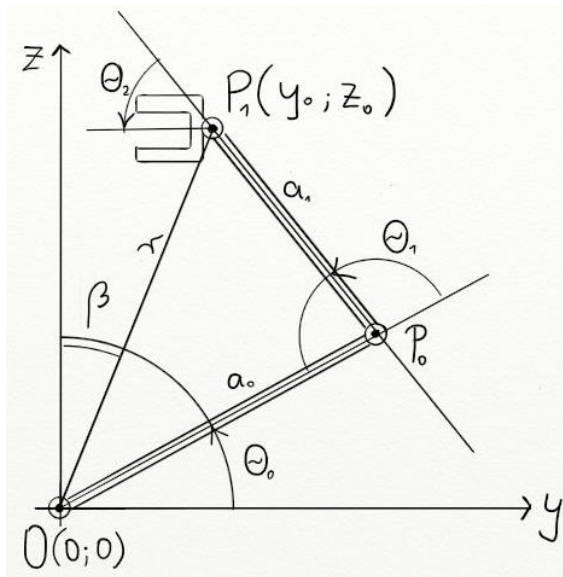
Tehát ha a kéz görbéje  $\gamma(t) = (\gamma_x(t), \gamma_y(t), \gamma_z(t))^T$  olyan, hogy

- nem halad át szinguláris pozíción, vagy
- ha a  $t_0$  pillanatban szinguláris pozíción megy át, akkor ott a  $(\gamma_x'(t_0), \gamma_y'(t_0), \gamma_z'(t_0))^T$  vektorra megoldható az inverz sebességkinematika,

akkor megoldható a differenciálegyenlet, így megoldottuk az inverz kinematikai feladatot.

A szinguláris pozíciók és irányuk meghatározása viszont nehéz feladat.

### 3.3. A síkbeli RRR robotkar



14. ábra A síkbeli RRR robotkar

Mivel példánkban az YZ síkban leszünk, már ebben a fejezetben is így tekintünk rá. Csak forgatási ízületeink vannak, tehát a direkt kinematika áttérési mátrixai a következő módon néznek ki:

$$M_i(0, \theta_i, a_i, 0) = \begin{pmatrix} \cos \theta_i & -\sin \theta_i & 0 \\ \sin \theta_i & \cos \theta_i & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & a_i \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos \theta_i & -\sin \theta_i & a_i \cos \theta_i \\ \sin \theta_i & \cos \theta_i & a_i \sin \theta_i \\ 0 & 0 & 1 \end{pmatrix}$$

Innen a 0 paramétereket elhagyva:

$$\begin{aligned} & M_0(\theta_0, a_0)M_1(\theta_1, a_1) = \\ & = \begin{pmatrix} \cos \theta_0 \cos \theta_1 - \sin \theta_0 \sin \theta_1 & -\cos \theta_0 \sin \theta_1 - \sin \theta_0 \cos \theta_1 & a_1 \cos \theta_0 \cos \theta_1 - a_1 \sin \theta_0 \sin \theta_1 + a_0 \cos \theta_0 \\ \sin \theta_0 \cos \theta_1 + \cos \theta_0 \sin \theta_1 & -\sin \theta_0 \sin \theta_1 + \cos \theta_0 \cos \theta_1 & a_1 \sin \theta_0 \cos \theta_1 + a_1 \cos \theta_0 \sin \theta_1 + a_0 \sin \theta_0 \\ 0 & 0 & 1 \end{pmatrix} = \\ & = \begin{pmatrix} \cos(\theta_0 + \theta_1) & -\sin(\theta_0 + \theta_1) & a_1 \cos(\theta_0 + \theta_1) + a_0 \cos(\theta_0) \\ \sin(\theta_0 + \theta_1) & \cos(\theta_0 + \theta_1) & a_1 \sin(\theta_0 + \theta_1) + a_0 \sin(\theta_0) \\ 0 & 0 & 1 \end{pmatrix} = \\ & = \begin{pmatrix} R(\theta_0 + \theta_1) & F_{y,z}(a_0, \theta_0, a_1, \theta_1) \\ \underline{0} & 1 \end{pmatrix} \end{aligned}$$

ahol

$$F_{y,z}(a_0, \theta_0, a_1, \theta_1) = \begin{pmatrix} F_y(a_0, \theta_0, a_1, \theta_1) \\ F_z(a_0, \theta_0, a_1, \theta_1) \end{pmatrix} = \underline{v}_0(a_0, \theta_0) + \underline{v}_1(a_1, \theta_1) = (y, z)^T$$

határozzák meg a kéz  $(y, z)$  koordinátáit.

Ezek alapján a szinguláris pozíciók meghatározásához a paramétereket  $t$ -től függővé tesszük:

$$\begin{pmatrix} y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} a_1 \cos(\theta_0(t) + \theta_1(t)) + a_0 \cos(\theta_0(t)) \\ a_1 \sin(\theta_0(t) + \theta_1(t)) + a_0 \sin(\theta_0(t)) \end{pmatrix}$$

majd deriválunk, és kiemeljük a paraméterfüggvények deriváltjait:

$$\begin{aligned} \begin{pmatrix} y'(t) \\ z'(t) \end{pmatrix} &= \begin{pmatrix} -a_1 \sin(\theta_0(t) + \theta_1(t)) (\theta_0'(t) + \theta_1'(t)) - a_0 \sin(\theta_0(t)) \theta_0'(t) \\ a_1 \cos(\theta_0(t) + \theta_1(t)) (\theta_0'(t) + \theta_1'(t)) + a_0 \cos(\theta_0(t)) \theta_0'(t) \end{pmatrix} = \\ &= \begin{pmatrix} -a_1 \sin(\theta_0(t) + \theta_1(t)) - a_0 \sin(\theta_0(t)) & -a_1 \sin(\theta_0(t) + \theta_1(t)) \\ a_1 \cos(\theta_0(t) + \theta_1(t)) + a_0 \cos(\theta_0(t)) & a_1 \cos(\theta_0(t) + \theta_1(t)) \end{pmatrix} \begin{pmatrix} \theta_0'(t) \\ \theta_1'(t) \end{pmatrix} = \\ &= (\underline{v}_1(t) + \underline{v}_0(t) \quad \underline{v}_1(t)) \begin{pmatrix} \theta_0'(t) \\ \theta_1'(t) \end{pmatrix} = (\underline{v}_2(t) \quad \underline{v}_1(t)) \begin{pmatrix} \theta_0'(t) \\ \theta_1'(t) \end{pmatrix} \end{aligned}$$

Pontosan akkor szinguláris a pozíció, ha  $\underline{v}_2(t) \parallel \underline{v}_1(t)$ , mert ekkor nem feszítik ki  $\mathbb{R}^2$ -t.

$$\begin{aligned} \underline{v}_2(t) \parallel \underline{v}_1(t) &\Leftrightarrow \underline{v}_1(t) + \underline{v}_0(t) \parallel \underline{v}_1(t) \Leftrightarrow \underline{v}_0(t) \parallel \underline{v}_1(t) \Leftrightarrow \\ &\Leftrightarrow a_0 \begin{pmatrix} -\sin(\theta_0(t)) \\ \cos(\theta_0(t)) \end{pmatrix} \parallel a_1 \begin{pmatrix} -\sin(\theta_0(t) + \theta_1(t)) \\ \cos(\theta_0(t) + \theta_1(t)) \end{pmatrix} \Leftrightarrow \\ &\Leftrightarrow \begin{pmatrix} \sin(-\theta_0(t)) \\ \cos(-\theta_0(t)) \end{pmatrix} \parallel \begin{pmatrix} \sin(-(\theta_0(t) + \theta_1(t))) \\ \cos(-(\theta_0(t) + \theta_1(t))) \end{pmatrix} \Leftrightarrow -\theta_0 = -(\theta_0 + \theta_1) + k\pi \Leftrightarrow \theta_1 = k\pi \end{aligned}$$

ahol  $k \in \mathbb{Z}$ . Ez azt jelenti, hogy pontosan akkor vagyunk szinguláris pozícióban, ha a második csukló teljesen visszazár, vagy teljesen kinyúl. Ez alapján ha  $a_0 > a_1$ , akkor a munkatéren, azaz a  $W = \{(y, z) \in \mathbb{R}^2 \mid (a_0 - a_1) \leq r \leq (a_0 + a_1), \text{ ahol } r = \sqrt{y^2 + z^2}\}$  körgyűrűn szigorúan belül nincsen ilyen, hiszen a feltétel pontosan a határcöröket jelenti.

Az inverz kinematikát, amihez általában differenciálegyenlet szükséges, itt geometriai megfigyelések alapján is megoldhatjuk (ld. 11.ábra):

Itt adott  $(y_0, z_0) \in \mathbb{R}^2$ , amely az origóból induló robotkar kezének kívánt helyzete, és  $a_0$ , illetve  $a_1$ , melyek a robotkar fix paraméterei. Keressük  $\theta_0$  és  $\theta_1$  szögeket, melyekkel ez előáll. A következő jelöléseket és összefüggéseket használjuk:

- $r$  az  $\overline{OP_1}$  szakasz hossza, így

$$r = \sqrt{y_0^2 + z_0^2}$$

- $\beta$  az  $OP_1$  és a  $z$  tengely által bezárt szög, tehát

$$\operatorname{tg}(\beta) = \frac{y_0}{z_0} \Rightarrow \beta = \operatorname{arctg}\left(\frac{y_0}{z_0}\right)$$

- az  $OP_0P_1$  háromszög  $O$  csúcsához eső szögének nagysága  $\frac{\pi}{2} - (\theta_0 + \beta)$ , a  $P_0$ -hoz tartozó  $(\pi - |\theta_1|)$ .
- az  $OP_0P_1$  háromszögre alkalmazva a koszinusz-tételt:

$$\begin{aligned} a_1^2 &= a_0^2 + (y_0^2 + z_0^2) - 2a_0\sqrt{y_0^2 + z_0^2} \cos\left|\frac{\pi}{2} - (\theta_0 + \beta)\right| \Rightarrow \\ &\Rightarrow \left|\frac{\pi}{2} - (\theta_0 + \beta)\right| = \operatorname{arccos}\left(\frac{a_0^2 + y_0^2 + z_0^2 - a_1^2}{2a_0\sqrt{y_0^2 + z_0^2}}\right) \end{aligned}$$

- hasonlóan,  $\theta_1$ -hez:

$$\begin{aligned} (y_0^2 + z_0^2) &= a_0^2 + a_1^2 - 2a_0a_1 \cos(\pi - |\theta_1|) \Rightarrow \\ &\Rightarrow |\theta_1| = \pi - \operatorname{arccos}\left(\frac{a_0^2 + a_1^2 - y_0^2 - z_0^2}{2a_0a_1}\right) \end{aligned}$$

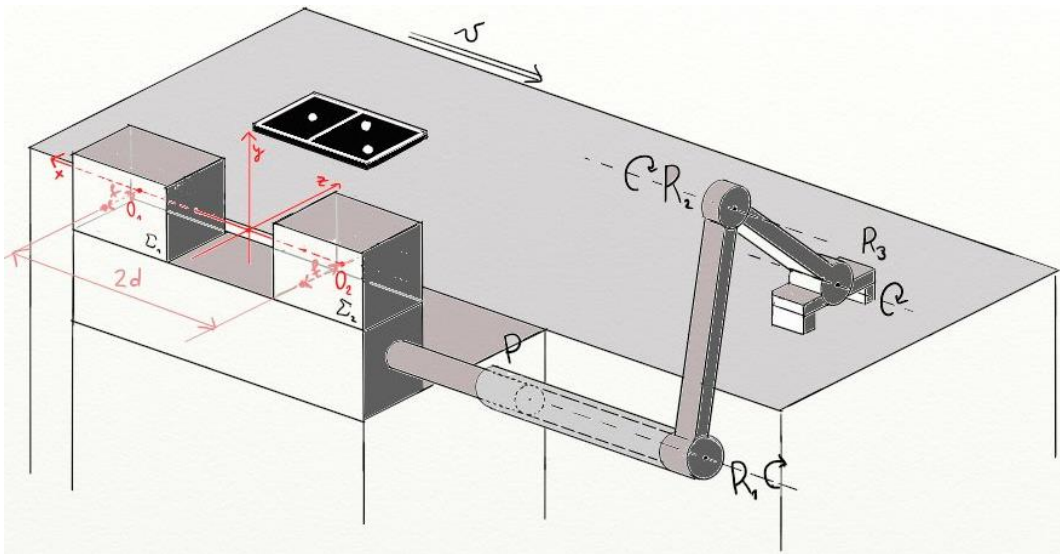
Innen még több megoldás lehetne, de ez valóságos is, hiszen  $OP_1$ -ra tükrözve a robotkar állását, ugyanezt a helyzetet kapjuk – annyi kell, hogy a szögek előjele megegyezzen. Konkrét esetben úgy választjuk, hogy  $\theta_0$  és  $\theta_1$  pozitívak legyenek, mert ez kódolja azt az állást, amikor biztosan nincs ütközés a futószalaggal.

A harmadik forgató ízület feladata az, hogy a kéz minden esetben „lefelé nézzen”,  $-y$  irányba, a 9. ábrán látható módon. Ehhez az kell, hogy mindig fennálljon a következő összefüggés:

$$\theta_0 + \theta_1 + \theta_2 = \pi$$



### 3.4. A térbeli PRRR robotkar inverz kinematikája



15. ábra A selejtezés modellje

Bár a dominót térben kell elkapni, nem síkban, mivel az első eltoló ízülettel elérjük a kívánt  $x_0$  koordinátát, innen a következő három forgató ízület szokványos síkbeli RRR-ként működik ebben az  $\{x = x_0\}$  síkban, így az előző fejezetben kifejtetteket hasznosíthatjuk. Az utolsó forgató ízület feladata az, hogy a kéz mindig az elkapáshoz megfelelő helyzetben álljon.

Az első ízület tehát az  $x$  tengely negatív irányában tud eltolni. Alaphosszát ( $a_1$ ) úgy határozzuk meg, hogy amikor a kamera fotójának elkészülése után lezajlanak a szükséges számítások, és eljut az információ a karhoz, hogy ki kell-e venni a tárgyat, akkor az pont abban a síkban legyen, ahova ezidő alatt ért az építőkocka. Tehát ha  $t_0$  ideig tartanak a számítások és  $v$  sebességgel mozog a futószalag, akkor

$$a_0 = (t_0 v) - x^*$$

ahol  $x^*$  jelöli a kamerák tartójának  $-x$  irányú kiterjedését. Ha  $y^*$  az a távolság, amennyivel lejjebb helyezkedik el az első csukló a kamera horizontvonalához képest, akkor a robotkar „origójának”, azaz az alaptest és az első csukló találkozásának a koordinátái  $(-x^*, 0, -y^*)$ .

Innen, amikor szükséges,  $v_0$  sebességgel tolja a kart, vagyis

$$l_0(t) = a_0 + vt$$

ahol  $t \in [t_k, t_v]$ , ahol  $t_k = t_f + t_0$ , ha  $t_f$ -fel jelöljük a fotózás időpontját,  $t_k$ -vel a robotkar mozgásának kezdő-,  $t_v$ -vel a végső pillanatát. Ezt azért fontos követni, mert a dominók folyamatosan érkeznek, tehát hamar újra szükség lehet a selejtezésre, adott idő alatt vissza

kell érnie az alaphelyzetbe. Az időmérést minden dominóhoz külön végezhetjük, így amint visszaért a kar, lenullázzuk  $t_f$ -et, ezzel periodikussá téve a folyamatot.

A második, harmadik és negyedik ízületnek a síkbeli RRR-nek megfelelően a

$$\theta_1(t_v) = \frac{\pi}{2} - \arccos\left(\frac{a_1^2 + y_0^2 + z_0^2 - a_2^2}{2a_1\sqrt{y_0^2 + z_0^2}}\right) - \arctg\left(\frac{z_0}{y_0}\right) \text{ és}$$

$$\theta_2(t_v) = \pi - \arccos\left(\frac{a_1^2 + a_2^2 - y_0^2 - z_0^2}{2a_1a_2}\right)$$

$$\theta_3(t_v) = \pi - (\theta_1(t_v) + \theta_2(t_v))$$

paramétereket kell felvenniük ahhoz, hogy a végpillanatban a kívánt  $(x_0, y_0, z_0)$  pontba kerüljön a kéz. Itt  $y_0$ -hoz már hozzászámoltuk az eltolást, vagyis igazából az  $(x_0, y_0 + y^*, z_0)$  pontban valósul meg az elkapás.

A kezdőpillanatbeli

$$\theta_1(t_k) = 0 \text{ és } \theta_2(t_k) = \frac{\pi}{2}$$

választás mellett a következő függvény leírja a kar egy lehetséges mozgását  $t$  függvényében:

$$\theta_i(t) = \theta_i(t_k) + t \cdot \theta_i(t_v)$$

ahol  $i \in \{1; 2\}$  és  $t \in [t_k; t_v]$ . Ehelyett magasabbrendű függvényeket használnak, hiszen azt szeretnénk, a kar gyorsan mozogjon, de figyelembe véve, hogy ez lineáris esetben túlságosan hirtelen megállást, illetve indulást eredményezne, olyat választunk, aminek a kezdő- és végpontbeli deriváltja 0. Ellenkező esetben a kar túlságosan kilengene, nagyobb terhelésnek lennének kitéve és súlyosabb mértékben kopnának az alkatrészek.

Az utolsó ízülettől elvárhatjuk, hogy végig az elkapáshoz szükséges állapotban legyen:

$$\theta_3(t) = \pi - (\theta_1(t) + \theta_2(t))$$

ha  $t \in [t_k, t_v]$ . Eszerint a kezdőpillanatbeli érték:

$$\theta_3(t_k) = \pi - \left(0 + \frac{\pi}{2}\right) = \frac{\pi}{2}$$

Így elértük, hogy a robotkéz oda kerüljön, ahol a dominó van. Egy érzékelőnek köszönhetően az ujjak automatikusan összezárnak. Ezután a selejtggyűjtőhöz kell juttatni, amit fix koordináták jelölnek. Ezt célszerű a robotkar alaphelyzetéhez köthetően elhelyezni. Vegyük azt az esetet, amikor pont az alaphelyzet síkjában van, a futószalaggal átellenes oldalon. Ekkor miközben az eltoló kar visszahúz az  $\{x = x_0\}$  síkba, a forgók feladata, hogy az  $y$  tengelyre tükrös helyzetbe kerüljön az RRR rész, vagyis az alapállapothoz választott szögek ellentettjéhez kell átfordulni. Itt elenged a kéz, és majd visszatérünk az alapállapotba. Ezeket a mozgásokat is az elkapáshoz jelzett szempontok szerint választott függvénnyel érdemes leírni, vagyis adott végpontbeli függvény- és deriváltértékekhez – utóbbi 0 – keresünk magasabbrendű függvényt. Hogy pontosan melyiket, azt az olyan körülmények, mint a munkatér vagy a felhasználható idő korlátoltsága adják meg egy konkrét konstrukció esetén.

# Összegzés

Mivel dolgozatom témája egy összetett feladat volt, így több részletének megoldása is bemutatásra került. Most, hogy az olvasó már megszerezte a szükséges háttér-információkat, célszerű összegezni, hogyan is valósult meg ez a bizonyos selejtezés.

Az első fejezetben megismerkedtünk a fényképezés matematikai modelljével, ami egy ideális eset volt, illetve a projektív leképezésekkel, amely segített a valós síkok fényképpé transzformálását megérteni. Erre volt szükség ahhoz, hogy az oldalhelyzetű kameraképből visszanyerjük a felülnézetet. Itt tudtuk ugyanis a minta helyességét ellenőrizni.

A második fejezetben gyakorlati megvalósításokat láttunk, itt már azt elemeztük, hogyan tudjuk a fényképen megjelent alakzatokat értelmezni. Szó esett az élek detektálásáról, ezzel találtuk meg a dominót és felső lapjára került mintáját.

A harmadik fejezetben először általánosan a robotkarokat leíró matematikáról lehet olvasni, utána egy konkrét típusú karra kiszámoltuk, hogyan állítsuk be, hogy felkapja a dominót, ha szükséges.

Ezen témák közül mindegyik gazdag szakirodalommal rendelkezik, és további mélységekben lehetne róluk beszélni. Folyamatosan érkeznek újabb feladatok, ahol valamit fel kell ismerni, ezekhez készülnek az újabb, gyorsabb, pontosabb algoritmusok. A szelektálás más konstrukciókkal, pl. merőleges szegmensekből álló PPP karral is megvalósítható lenne, vagy akár a kamera is lehet mozgó, a kézhez rögzített.

Számomra későbbi cél lehet a robotkarok vezérlési számítógépének, további alakzatfelismerési módszereknek, mint a telefonok ujjlenyomat- ill. arcfelismerője, vagy akár a különböző képszerkesztő filterek megismerése.

# Irodalomjegyzék

- [1] D. H. Ballard és C. M. Brown, *Computer Vision*, Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1982, pp. 16-22, 75-81.
- [2] B. Siciliano, L. Sciavicco, L. Villani és G. Oriolo, *Robotics Modelling, Planning and Control*, London: Springer, 2009, pp. 406-417.
- [3] D. Szeghy, *Alkalmazott Modul Jegyzet - Geometriai transzformációk*, ELTE, 2013.
- [4] J. N. Wilson és G. X. Ritter, *Handbook of Computer Vision Algorithms in Image Algebra*, USA: CRP Press, 2000, pp. 96-98.