Studienprojekt: Mathematische Methoden der Computervision

Kevin von Bargen

September 12, 2019

1 Einleitung

In dieser Arbeit beschäftigen wir uns damit, wie die Algorithmen der Computervision funktionieren, welche benötigt waren um automatisiert die grundlegenen Informationen des Einzelspieler-Spieles Super Hexagon, von Terry Cavanagh, zu bestimmen.
Diese wurden weiter verwendet um es einem Neuronalen Netz zu ermöglichen das Spiel zu spielen und zu lernen besser zu werden.
Grundlagenkenntnisse der Linearen Algebra und analytischen Geometrie, sowie der Analysis sind hierbei ausreichend um die mathematische Ausarbeitung nachvollziehen zu können.
Um die dahinter liegenen Programme zu verstehen sollten grundlegene Kenntnisse einer Programmiersprache vorhanden sein, am besten in Python.
Die Progamme, sowieso Erläuterungen des Ablaufes und Demonstrationen finden Sie unter `https://github.com/kvonbargen/super-hexagon` .
Zum Verständnis des skizzierten Programmablaufes sind keine Programmierkenntnisse notwendig.

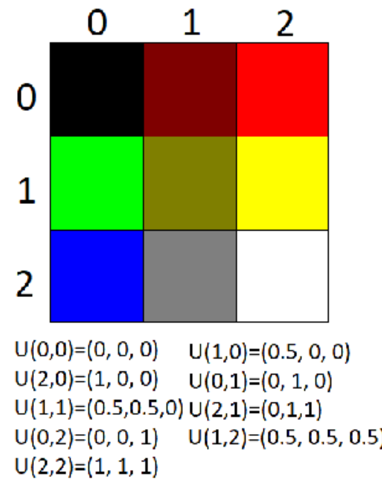
2 Grundlagen der Bildverarbeitung

2.1 Definition Diskretes Bild

Ein diskretes Bild ist eine Abbildung U: $\mathbb{Z}^2 \to \mathbb{R}^d$

2.2 Beispiele für Farbräume

RGB Bilder: $U : \mathbb{Z}^2 \to \mathbb{R}^3$, wobei in Anwedungen häufig $U : \mathbb{Z}^2 \to [0,1]^3 \subset \mathbb{R}^3$ verwendet wird, wobei 0 kein Anteil eines Farbwertes und 1 den maximalen Anteil darstellt.
HSV Bilder: $U : \mathbb{Z}^2 \to (H, S, V) \in [0, 360) \times [0, 100] \times [0, 100] \subset \mathbb{R}^3$ wobei H=Hue den Farbton, S=Saturation bzw Sättigung und V=Value den Hellwert angibt.
Graustufen Bilder: : $U : \mathbb{Z}^2 \to [0,1] \subset \mathbb{R}$ bzw in Anwendung eher $U : \mathbb{Z}^2 \to [0, 255] \subset \mathbb{N}$
Binäre Bilder: : $U : \mathbb{Z}^2 \to 0, 1$ wobei 0 Schwarz und 1 Weiß meint.



2.3 Transformation eines RGB Bildes zu einem graustufen Bild

Für die Bestimmung einer Farbe gibt es die Maßeinheit Luma. Das menschliche Auge nimmt beispielsweise maximal helle Blau $(0, 0, 1)$ deutlich dunkler wahr als das maximal helle Grün $(0, 255, 0)$, weshalb hier die graustufen Transformation nicht durch das arithmetische Mittel bestimmt wird, sondern nach dem CCIR601: $G = 0.2989 * R + 0.587 * G + 0.114 * B$

2.4 Definition Histogramm

Sei U ein diskretes Graustufen Bild mit $U : \Omega \to$ F, $\Omega = \{1, \dots, N\} \times \{1, \dots, M\}$, $F = \{0, \dots, 255\}$ so ist $H_U(k) = \#\{(i,j) \in \Omega \mid U(i,j) = k\}$ bzw mit dem Kronecker-Delta:

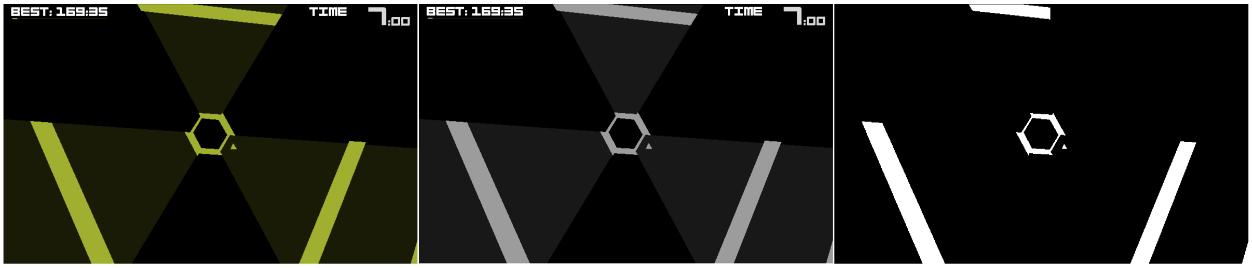
$$\delta_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, H_U(k) = \sum_{i=1}^{N} \sum_{j=1}^{M} \delta_{k,U(i,j)}$$

2.5 Schwarz-Weiß-Segmentierung durch Schwellwertbildung

Sei $s_0 \in \mathbb{R}$ und U ein diskretes graustufen Bild so ist das durch die Segmentierung enstehende Bild $\tilde{U}$ eintragsweise definiert als $\tilde{U} = \begin{cases} 1 & U(i,j) > s_0 \\ 0 & U(i,j) \leq s_0 \end{cases}$ und $\tilde{U} : \mathbb{Z}^2 \to \{0, 1\}$

```
img_np =  np.asarray(sct.grab(monitor_segment))
gray=cv2.cvtColor(img_np, cv2.COLOR_BGR2GRAY)
```

```
binary= cv2.threshold(orginal,thresh, 255, cv2.THRESH_BINARY)[1]
```



2.6 Definition Strukturelement

Ein Strukturelement $B \subset \mathbb{Z}^2$ ist eine nicht-leere Strukturmenge, der zwei dimensionalen diskreten Grundmenge.

Beispiele: 4-Umgebung $\begin{bmatrix} 0 & 1 & 0 \\ 1 & \otimes & 1 \\ 0 & 1 & 0 \end{bmatrix}$ 8-Umgebung $\begin{bmatrix} 1 & 1 & 1 \\ 1 & \otimes & 1 \\ 1 & 1 & 1 \end{bmatrix}$

wobei $\otimes$ den Bezugspunkt darstellt.

2.7 Definition Dilation

Sei B ein Strukturelement und U ein diskretes Bild, so ist

$(U \oplus B)(x) = \begin{cases} 1 & \text{falls } \exists y \in B : U(x+y) = 1 \\ 0 & \text{sonst} \end{cases}$ die Dilation und

$(U \ominus B)(x) = \begin{cases} 1 & \text{falls } \forall y \in B : U(x+y) = 1 \\ 0 & \text{sonst} \end{cases}$ die Erosion

2.8 Lemma: Charakterisierung der Dilation/Erosion durch Suprema und Infima

Sei U ein diskretes Bild und B ein Strukturelement, so ist
$(U \oplus B)(x) = \sup_{y \in B} U(x+y)$ und $(U \ominus B)(x) = \inf_{y \in B} U(x+y)$

Proof. $\sup_{y \in B} U(x+y) \Leftrightarrow \exists y \in B$ mit $U(x,y) = 1 \Leftrightarrow (U \oplus B)(x) = 1$
$\inf_{y \in B} U(x+y) \Leftrightarrow \forall y \in B$: $U(x,y) = 1 \Leftrightarrow also (U \ominus B)(x) = 1$

$\square$

2.9 Bermerkung: Weitere Eigenschaften zu Dilation und Erosion

Die Filter Dilation und Erosion sind dual zu einander,also $-(U \oplus B) = (-U) \ominus B$, sowie translationsinvariant und distributiv.

2.10 Definition Öffnen und Schließen

Sei B ein Strukturelement und B ein diskretes Bild, so ist Öffnen die Zusammensetzung $(U \circ B) = (U \ominus B) \oplus (-B)$ und $(U \bullet B) = (U \oplus B) \ominus (-B)$ Schließen.
Die Operation Schließen ist Extensional bzw $U \leq U \bullet B$ und Öffnen ist antiextensional $U \circ B \leq U$ und beide sind idempotent.

3 Finden von Konturen in binären Bildern

3.1 Definiton Zusammenhängende Punkte bzw Elemente

Zwei Punkte $x_0 = (i,j), x_n = (p,q) \in \mathbb{Z}^2$ eines binären Bildes U hängen zusammen, falls $U(x_0) = U(x_n)$ und $\exists x_k$ Folge von Punkten mit $U(x_k) = U(x_0) = U(x_n) \forall (0 \leq k \leq n)$ und x_k ist in der 4-(8-) Umgebung von x_{k-1}. Hierdurch wird eine Äquivalenzraltion induziert, und die hierdurch induzierten Äquivalenzklassen bilden die:

3.2 Definition Zusammenhängende Komponente

Wobei jede Äquivalenzklasse mit Bildwert 1 eine 1-Komponente bzw mit Bildwert 0 eine 0-Komponente genannt wird.

3.3 Definition Grenzpunkt

Ein Punkt $x_k = (i,j)$ in einem diskreten binären Bild ist ein Grenzpunkt, wenn $\exists (p,q)$ in seiner 4-(8-) Umgebung mit $U(i,j) \neq U(p,q)$ und $U(i,j) = 1 \Rightarrow U(p,q) = 0$.

3.4 Umschließen von Zusammenhängenden Komponenten

Seien S_1, S_2 zusammenhängende Komponenten in einem binären Bild, so umschließt S_2 S_1 ,wenn für jeden Punkt aus $S_{1]}$ in jener 4-(8-) Umgebung ein Punkte aus S_2 liegt.

3.5 Innere und Äußere Grenzen

Eine äußere Grenze in einem diskreten binären Bild ist eine Menge von Grenzpunkten einer 1-Komponente und der 0-Komponente von der sie umschlossen wird.
Umgekehrt ist die innere Grenze die Menge von Grenzpunkten einer 1-Komponente die eine 0-Komponente umschließt.

3.6 Eigenschaft: Innere/Äußere Grenzen sind einzigartig für zusammenhängende Komponenten

Sei S_1 eine zusammenhängende 1-Komponente, die S_2 umschließt, so ist die innere Grenze eindeutig bestimmt. (Die selbe Eigenschaft gilt umgekehrt für die äußere Grenze zwischen der 0-Komponente und der 1-Komponente).

Proof. Sei B die Menge der Grenzpunkte der inneren Grenze von S_1 zu S_2
Annahme:$\exists(i,j) \in S_1$und (i,j) Grenzpunkt zwischen S_2 zu S_1 und $(i,j) \notin B$, die folgt jedoch zu einem Widerspruch, da (i,j) per Konstruktion $\in B$. □

3.7 Vater Grenze

Die vater Grenze einer äußeren Grenze zwischen einer 1-Komponente S_1 und der 0-Komponente S_2 ist, die S_1 umschließt, ist entweder (1) die innere Grenze von S_2 , der S_2 umschließenden 1-Komponente, oder (2) der Bildrahmen, falls S_2 der Hintergrund ist. Umgekehrt ist für eine innere Grenze zwischen einer 0-Komponente und der umschließenden 1-Komponente ihre äußere grenze die Vater Grenze.

3.8 Umschließung von Genzen

Seien B_0, B_n Grenzen in einem diskreten binären Bild, so umschließt B_n die Grenze B_0 , falls $\exists B_0, \ldots, B_n$ Folge von Grenzen mit B_k ist Vater Grenze von $B_{k-1} \forall (1 \leq k \leq n)$

3.9 Eigenschaften

Die Relation der Vater Grenze zusammen mit der Umschließung von Grenzen ist isomorph zu der Zuordnung:
1-Komponente ⇔ äußere Grenze
0-Komponente ⇔ innere Grenze
Hintergrund ⇔ Bildrahmen

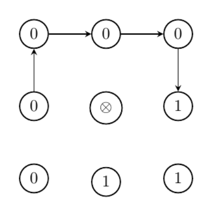
3.10 Algorythmus von Suzuki, Abe

(0)Suche Zeilenweise in einem diskreten binären Bild U nach U(i,j)=1, wenn solch ein paar gefunden wurde gehe zu Schritt (1)

$$(1)\begin{cases} NBD = NBD + 1, (i_2, j_2) = (i, j-1) & \text{falls } U(i,j) = 1 \wedge U(i, j-1) = 0 \\ NBD = NBD + 1, (i_2, j_2) = (i, j+1) & \text{falls } U(i,j) \geq 1 \wedge U(i,j) = LNBD \text{ ,falls } U(i,j) > 1 \\ \text{Gehe zu (4)} & \text{sonst} \end{cases}$$

(3.1) Suche mit dem Uhrezeigersinn startend bei (i_2, j_2) in der Umgebung von (i,j) nach dem ersten Punkt (p,q) mit $U(p,q) \neq 0$.

$$\begin{cases} (i_2, j_2) = (p, q) & \text{falls solch ein Punkt existiert} \\ U(i,j) = -\text{NBD und gehe zu (4)} & sonst \end{cases}$$



(3.2) $(i_2, j_2) = (i_1, j_1)$, $(i_3, j_3) = (i, j)$

(3.3) Suche diesmal **gegen** den Uhrezeigersinn in der Umgebung von (i_3, j_3) beginnend ab (i_2, j_2), falls ein Punkte gefunden wird ist dieser (i_4, j_4)

$$(3.4)\begin{cases} U(i_3, j_3) = \text{-NBD} & \text{falls } U(i_3, j_3 + 1) = 0 ** \\ U(i_3, j_3) = \text{NBD} & \text{falls } U(i_3, j_3 + 1) \neq 0 \end{cases}$$

$$(3.5)\begin{cases} \text{Gehe zu (4)} & \text{falls } (i_4, j_4) = (i, j) \wedge (i_3, j_3) = (i_1, j_1) \\ (i_2, j_2) = (i_3, j_3), (i_3, j_3) = (i_4, j_4), \text{gehe zu (3.3)} & \text{sonst} \end{cases}$$

$$(4)\ \begin{cases} \text{LNBD} = |U(i, j)|, \text{ gehe zu (0)} & U(i, j) \neq 1 \\ \text{Gehe zu (0)} & \end{cases}$$

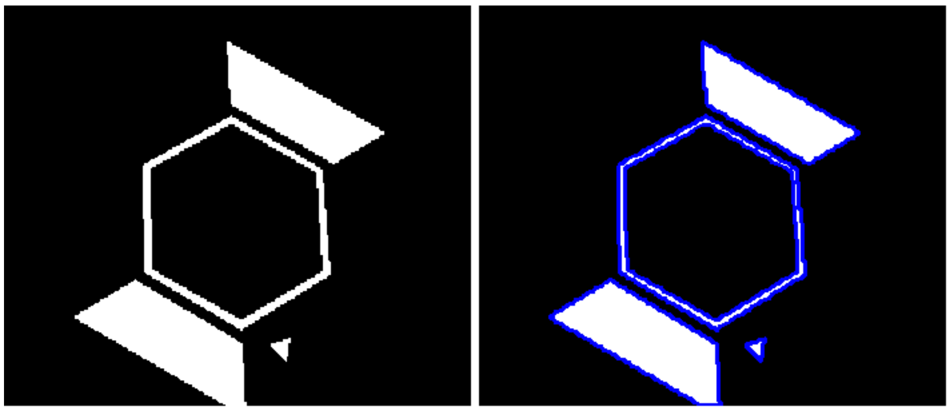


Figure 1: Beispiel des Algorythmus, jeder Grenzpunkt wurde blau gefärbt

3.11 Lemma

Es werden nur innere/äußere Grenzen B zwischen einer 1-Komponente$S_1 \ni (i, j)$ und einer 0-Komponente $S_2 \ni (i, j+1)$ bzw $S_2 \ni (i, j-1)$ gefolgt. (2) Alle Grenzpunkte von B werden gefunden.

Proof. Falls $|S_1| = 1$, so ist seine äußere Grenze nur $(i, j) = S_1 = B$ und wird gefunden, durch die zeilenweise Suche und durch (1) als Grenze erkannt.
(1) Sei $|S_1| > 1$: Sei $(p, q) \in B$ beliebiger Grenzpunkt von B und (h, k) sein Nachfolger $\Rightarrow (h, k)$ liegt in der Umgebung von (p, q) und alle dazwischen betracheteten Bildpunkte in (3.1 oder 3.3) gehören zu S_2 somit ist (h, k) valider Grenzpunkt der Grenze B $\Rightarrow$ (1) (2) Da $|S_1| > 1$ ist $|B| \geq 2$, somit endet das Vorgehen in Schritt 3 bzw genauer (3.3-3.5) erst, wenn der nöchste gefunde Grenzpunkt $(h, k) = (i, j)$ also der Startpunkt der Suche ist. Also stellt B einen Kreis dar
(Es existieren 2 Wege bzw Folgen $B_k, \tilde{B}_k$ von jedem Punkt $(h, l), (m.n)$ mit $B_k \neq \tilde{B}_r \forall k, r$)
Somit ist die Grenze B eindeutig bestimmt und vollständig □

3.12 Lemma

Sei B die Grenze zwischen der 1-Komponente $S_1 \ni (i, j)$ und der 0-Komponente $S_2 \ni (i, j-1)$, falls äußere Grenze bzw $S_2 \ni (i, j+1)$ falls innere Grenze, so wird $(p, q) \in B$ auf einen negativen Wert gesetzt $\Leftrightarrow (p, q+1) \in S_2$

Proof. $|S_1| > 1$: Der wrd eines Grenzpunktes (p, q) wird nur negativ $\Leftrightarrow U(p, q+1) = 0 \wedge (p, q+1) \in S_2$, da in Schritt 3.3 nut 0-Pixel$\in S_2$ betrachtet werden. □

3.13 Lemma

Der erste durch die zeilenweise Suche gefundene 1-Pixel $(i, j) \in S_1$ ist $\in B$ äußere genze zwischen der S_1 umschließenden 0-Komponente $S_2 \ni (i, j-1)$.

Proof. Da (i,j) erster gefundene 1-Pixel $\rightarrow$ vorherige Pixel alles 0-Pxel, also $U(i-1, j-1) = U(i-1, j) = U(i-1, j+1) = U(i, j-1) = 0$, desweiteren kann der Wert $|U(i, j)| \not> 1$ sein, da zuvor noch keine andere genze gefunden wurde. □

4 Eigenschaften einer Kontur

4.1 Gesamtlänge des Streckenzugs

Sei $P_0, \ldots, P_n$ eine Punktfolge , so ist $\sum_{i=1}^{n} d(P_{i-1}, P_i)$ die Gesamtlänge des Streckenzug bzw im Code die ArcLength.

4.2 Lemma: Satz von Green

Sei $R \subseteqq \mathbb{R}^2$ eine kompakte Menge mit glattem Rand $\partial R = C$ und $P, Q : R \rightarrow \mathbb{R}$ stetig mit auf R stetig partiellen Ableitungen $\frac{\partial P}{\partial y}, \frac{\partial Q}{\partial x}$, so ist

$$\oint_C P(x, y)dx + Q(x, y)dy = \iint_R (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy$$

Proof. Sei $C = C_1 \cup C_2$,sodass $C_1 : y = f_1(x) \forall a \leq x \leq b$ und $C_2 : y = f_2(x) \forall b \leq x \leq a$ Integrieren wird nun y zwischen $y = f_1(x)$ und $y = f_2(x)$ erhalten wir:

$$\int_{f_1(x)}^{f_2(x)} \frac{\partial P}{\partial y} dy = P(x, f_2(x)) - P(x, f_1(x)) \text{ integrieren wir nun über (a,b)}$$

$$\int_a^b \int_{f_1(x)}^{f_2(x)} \frac{\partial P}{\partial y} dydx = \int_a^b P(x, f_2(x)) - P(x, f_1(x))dx = \int_a^b P(x, f_2(x))dx - \int_a^b P(x, f_1(x))dx$$

$$= -\int_b^a P(x, f_2(x))dx - \int_a^b P(x, f_1(x))dx = -\int_{C_2} Pdx - \int_{C_1} Pdx = -\oint_C Pdx$$

$$\text{Somit gilt: } \oint_C Pdx = -\int_a^b \int_{f_1(x)}^{f_2(x)} \frac{\partial P}{\partial y} dydx = \iint_R (\frac{\partial P}{\partial y})dxdy$$

Ähnlich zerlegen wir nun C in $C_3 : x = g_1(y) \forall c \leq y \leq d$ und $C_4 : x = g_2(y) \forall d \leq y \leq c$

und erhalten $\displaystyle\int_{g_1(y)}^{g_2(y)} \frac{\partial Q}{\partial x} dx = Q(g_2(y), y) - Q(g_1(y), y)$ integrieren wir über (c,d) erhalten wir:

$$\int_c^d \int_{g_1(y)}^{g_2(y)} \frac{\partial Q}{\partial x} dxdy = \int_c^d Q(g_2(y), y) - Q(g_1(y), y)dy = \int_c^d Q(g_2(y), y)dy - \int_c^d Q(g_1(y), y)dy$$

$$= \int_c^d Q(g_2(y), y)dy + \int_d^c Q(g_1(y), y)dy = \int_{C_4} Qdy + \int_{C_3} Qdy = \oint_C Qdy$$

$$\text{Somit } \oint_C Qdy = \int_c^d \int_{g_1(y)}^{g_2(y)} \frac{\partial Q}{\partial x} dydx = \iint_R (\frac{\partial Q}{\partial x})dxdy$$

$$\text{Insgesamt also } \oint_C Qdy + \oint_C Pdy = \iint_R (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy$$

□

4.3 Lemma: Bestimmung des Flächeninhaltes einer Kontour

Seien $(x_0, y_0), \cdots, (x_n, y_n)$ die Eckpunkte eines Polygons, welche zwischen $(x_k, y_k), (x_{k+1}, y_{k+1})$ stückweise stetig und differenzierbar sind, so ist der Flächeninhalt A gegeben durch:

$$A = \sum_{k=0}^{n} \frac{(x_{k+1} + x_k)(y_{k+1} - y_k)}{2} \text{ mit } x_{n+1} = x_0 \text{ und } y_{n+1} = y_0$$

Proof.

$$\text{Da } \oint_C P(x, y)dx + Q(x, y)dy = \iint_R \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dA \text{ nach Green, ist somit}$$

$A = \iint_R 1 dA$ wobei A der Flächeninhalt der Region R ist. Nun wähle:

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1 \text{ mit } P(x,y) = 0 \text{ und } Q(x,y) = x.$$

Da Wegintegrale über stückweise stetige Funktionen additiv sind gilt:

$$A = \oint_C x dy = \int_{C_0} x dy + \cdots + \int_{C_n} x dy$$

Wähle für $C_k : [0,1] \to \mathbb{R}^2, t \to ((x_{k+1} - x_k)t + x_k, (y_{k+1} - y_k)t + y_k))$
Somit gilt:

$$\int_{C_k} x dy = \int_0^1 ((x_{k+1} - x_k)t + x_k)(y_{k+1} - y_k) dt = (y_{k+1} - y_k)(\int_0^1 (x_{k+1} - x_k) t dt + \int_0^1 x_k dt)$$

$$= (y_{k+1} - y_k)((x_{k+1} - x_k) \int_0^1 t dt + x_k \int_0^1 1 dt) = (y_{k+1} - y_k)(\frac{x_{k+1} - x_k}{2} + x_k) = \frac{(y_{k+1} - y_k)((x_{k+1} + x_k))}{2}$$

Folglich ist

$$A = \sum_{k=0}^{n} \int_{C_k} x dy = \sum_{k=0}^{n} \frac{(x_{k+1} + x_k)(y_{k+1} - y_k)}{2}$$

□

5 Literaturverzeichnis

[1] **Kristian Bredies, Dirk Lorenz**: Mathematische Bildverarbeitung, Einführung in Grundlagen und moderne Theorie, S:53-89 , Vieweg + Teubner, 2011
[2] **Satoshi Suzuki, Keiichi Abe**: Topological Structural Analysis of Digitized Binary Images by Border Following, S:32-46, Elsevier,1985