CV Repetitorium - Test 1

Image encoding and compression

- Wie viel Speicherplatz benötigt man für die Speicherung des Bildinhaltes bei einem RGB Farbbild der Größe 1024x768, wenn pro Farbkanal 4096 Werte kodiert werden sollen?

$$4096 \text{ values} \rightarrow 2^{12} = 4096$$
$$8 \text{ Bit} = 1 \text{ Byte}$$

$$\text{Number of pixels: } 1024 \cdot 768 = 786432 \text{ pixels}$$
$$3 \text{ channels, } 3 \cdot 12 \text{ Bit/Pixel}$$
$$786432 \cdot 3 \cdot 12 = 28311552 \text{ Bit} = 3538944 \text{ Byte} = 3538.9 \text{ kByte}$$

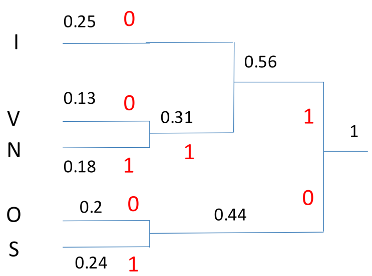
Codierung

Huffman Code

- Given a source alphabet X ={I, O, N, S, V} and their respective probability.

x_i	I	O	N	S	V
p_i	0.25	0.2	0.18	0.24	0.13

- Determine the Huffman coding with the code alphabet C = {0, 1}, assigning '1' to the more likely branch.



Symbol	Code
I	10
V	110
N	111
O	00
S	01

DCT Basic Functions

- Bei einer JPEG-Komprimierung entsteht das unten links angegebene Ergebnis nach der Anwendung der Diskreten Cosinus-Transformation (DCT) auf einen 8x8 Block. Geben Sie das Ergebnis der Quantisierung mittels der angegebenen Quantisierungsmatrix für die fett umrandeten Bildpunkte an.

Ergebnis der DCT:

$\boxed{93}$	-12	29	-1	-6	0	-1	-2
-11	-1	4	-13	5	2	-5	1
12	-14	-1	-2	-2	5	2	-3
59	2	-34	5	6	-2	4	-4
$\boxed{-87}$	-15	24	-3	-5	1	0	2
-42	30	11	-10	0	-2	0	4
25	1	2	-3	-2	1	0	0
59	-33	7	-6	5	3	-3	$\boxed{1}$

Quantisierungsmatrix:

$\boxed{16}$	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
$\boxed{18}$	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	$\boxed{99}$

$\frac{93}{16} \approx 6$ $\quad \frac{-87}{18} \approx -5$ $\quad \frac{1}{99} \approx 0$

Ergebnis:

$\boxed{6}$							
$\boxed{-5}$							
							$\boxed{0}$

Man dividiert die DCT Were durch die Quantisierungsmatrix und rundet.

Point Operations

- Angenommen, ein 10-Bit-Grauwertbild I(u,v) weist einen minimalen Intensitätswert von 50 und einen maximalen Intensitätswert von 250 auf. Wie lautet in diesem Fall die affine (lineare) Punktoperation, die den Kontrast des Bildes auf den gesamten Intensitätsbereich verstärkt?

$$p_{\max} = 2^{10} - 1 = 1023 \qquad q_{\max} = 250$$
$$p_{\min} = 0 \qquad q_{\min} = 50$$

$$I'(u,v) = (p_{\max} - p_{\min}) \frac{I(u,v) - q_{\min}}{q_{\max} - q_{\min}} + p_{\min}$$

Einsetzen in Formel: $I'(u,v) = 1023 \frac{I(u,v) - 50}{200}$

oder andere Art um das zu lösen:

$$f(x) = kx + d$$

$$f(50) \stackrel{!}{=} 0 \qquad 0 = 50k + d$$
$$f(250) \stackrel{!}{=} 1023 \qquad 1023 = 250k + d$$

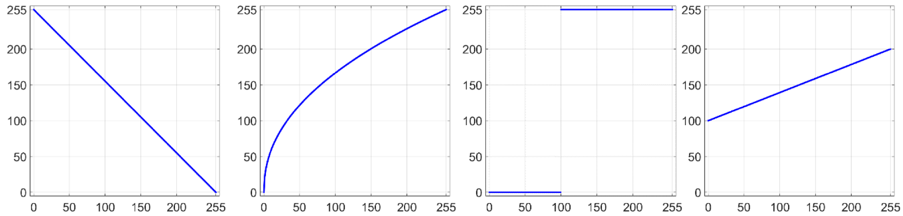
$$k = 5.12, d = -255.75$$

Andere Prüfungsfragen zu Point Operations

sind hier noch die Punktoperatioen verschiedenen Ergebnisbildern zuordnen

- Weisen Sie den unten gezeigten Abbildungsfunktionen 1-4 (transfer functions) die korrekte Punktoperation A-F zu.

A: Gamma-Korrektur **B**: Bildinvertierung. **C**: lin. Kontrasterhöhung
D: lin. Kontrastreduktion **E**: Schwellwertoperation **F**: Identitätsfunktion



B A E D

Local Operations

- Gegeben ist folgendes 8-Bit-Grauwertbild:
- Berechnen Sie für das fett markierte Pixel mit dem Wert 145 in der Mitte des Bildes das Ergebnis folgender Bildoperationen in Fließkommazahlen:
- Invertierung: 255-145 = 110
- 3x3 Mittelwertfilter: (20+20+20+125+145+50+0+70+70)/9 = 57,77
- 3x3 Median-Filter: Sortieren: 0, 20, 20, 20, 50, 70, 70, 125, 145 -> 50
- Faltung mit Laplace-Filter H: 20·1 + 125·1 + 50·1 + 70·1+145·(-4) = -315

10	15	5	5	0
20	20	20	20	10
100	125	**145**	50	50
0	0	70	70	60
5	5	20	30	50

$$H = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

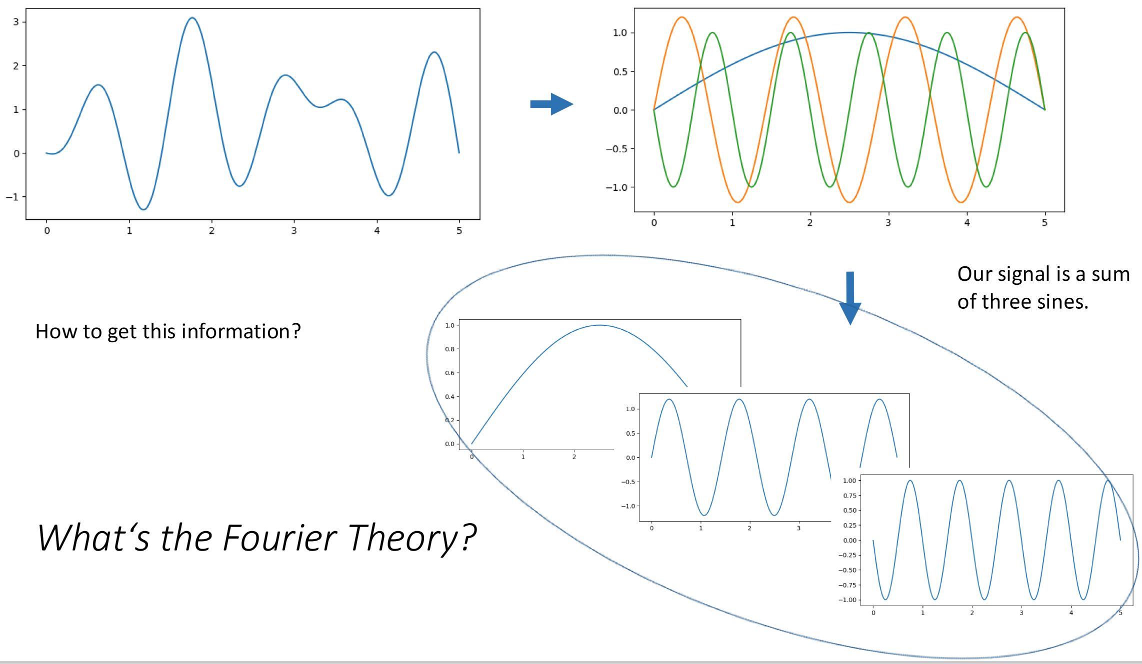
Edge Filtering

- Approximation of second derivative.

$$L = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad \text{or} \quad L = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

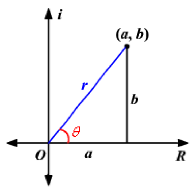
- Due to **high pass** characteristics (2nd order filter = 2nd derivative), Laplace filter is very **sensitive to noise**.
- Usually combined with Gaussian filter that reduces noise before Laplace filter is applied.

Globale Operation



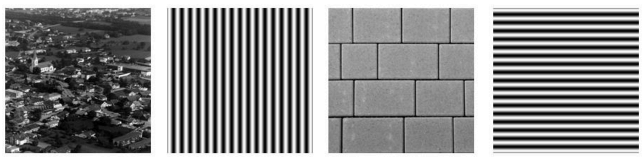
Das kommt nicht:

Fourier Transform

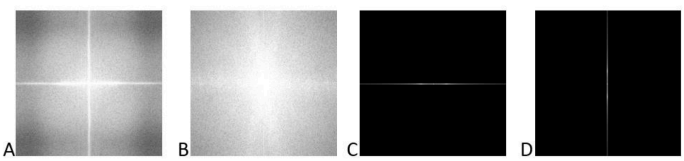


- Fouriertransformation: $F(\omega) = \int_{-\infty}^{\infty} f(t) \cdot \mathrm{e}^{-i\omega t}\mathrm{dt}$
- Property of transforms:
  - convert function from image domain to frequency domain without loss of information
- Usually, we are only interested in the magnitude: $\|F(\omega)\|$
  - $F(\omega)$ is complex: $a + bi$ -> $r = \sqrt{a^2 + b^2}$, $\varphi = \arctan\frac{b}{a}$
- Note: not every function has a fourier transform
  - e.g. if integral does not converge

Images A-D show the logarithmic Fourier spectrum of an image. Assign the input images I1 to I4 to the correct spectrum from A to D.



DFT(I_1)= B DFT(I_2)= C DFT(I_3)= A DFT(I_4)= D



A B C D

Hough Space

We parametrize our lines via $r = x\cos(\theta) + y\sin(\theta)$

$$\theta \in [-90°, 90°] \text{ vs. } k, d \in [-\infty, \infty]$$

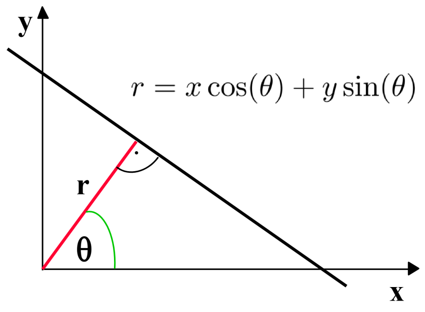
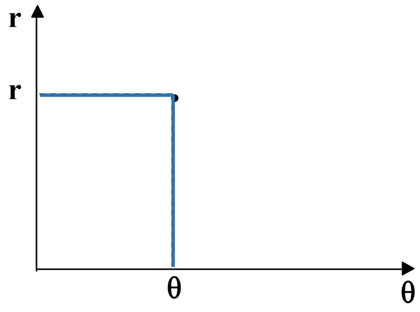


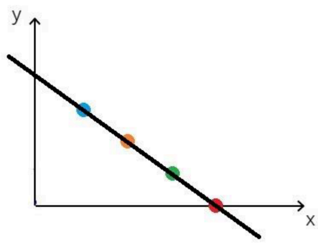
Image Space



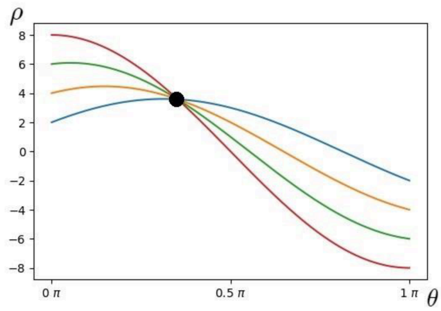
Parameter Space

- All possible lines of $P(x, y)$ are defined by:

$$r = x\cos(\theta) + y\sin(\theta)$$



Points which form a line



Bunch of sinusoids intersecting at one point

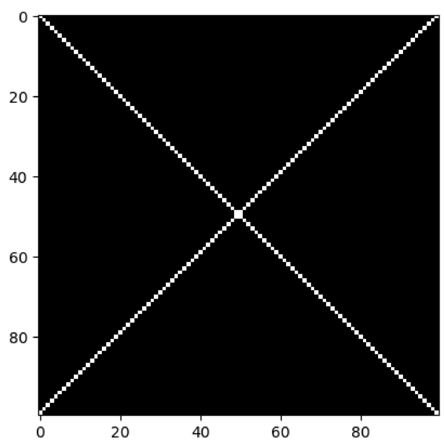
Beispiel

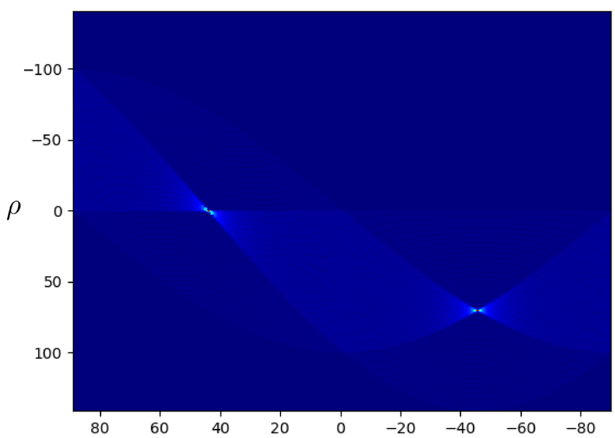
$\theta_2 = -\theta_1$

$\theta_2 - \theta_1 = 90°$

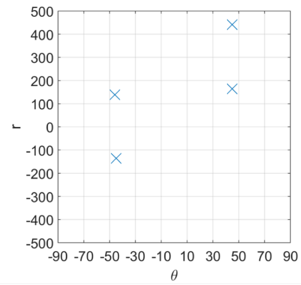
Orthogonal Lines!

$\rho_2 = \sqrt{(50^2 + 50^2)} \approx 70, \theta_2 = -45°$

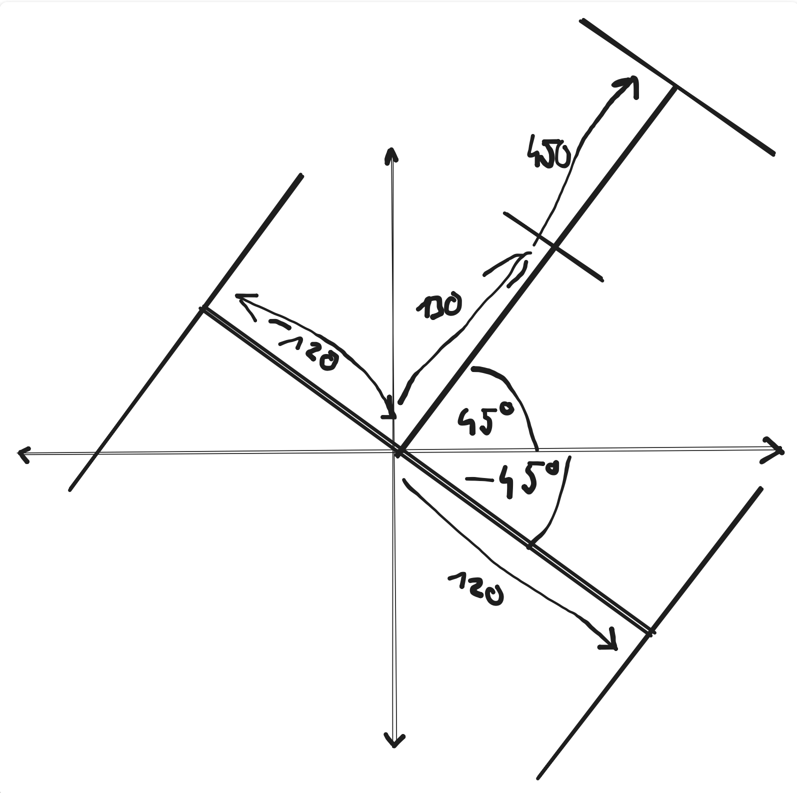




In the Hough transformation for detecting lines, these are represented in Hessian normal form: r = x cos(θ)+y sin(θ). In the diagram below, 4 detected lines in Hough space (accumulator array) are marked with an 'X'. Which of the following statements are true or false?



All 4 lines are parallel to each other	☐ True	☒ False
At least one line runs horizontally	☐ True	☒ False
At least one line runs vertically	☐ True	☒ False
The start and end points of the lines cannot be determined from the Hough space	☒ True	☐ False
Lines are represented by local maxima in Hough space	☒ True	☐ False



1. Falsch weil dann müsste der Winkel bei allen gleich sein
2. Falsch, weil keine bei 90 Grad sind
3. Falsch, weil keine den Winkel bei 0 haben
4. Richtig
5. Richtig

Interest Points

Given a 5x5 image section to which the Moravec corner detector is to be applied. Calculate the changes in the intensities E for the point marked with an asterisk * (3,3) and the 4 displacements (1,0), (1,1), (0,1) and (-1,1). Use a window size of 3x3 and the sum of squared differences (SSD). Furthermore, determine the 'interest value' from the changes in intensity.

y					
1	70	60	70	60	60
2	80	80	90	80	80
3	80	90	**100***	100	100
4	80	90	100	100	100
5	70	80	100	100	110
x	1	2	3	4	5

$E(1,0)=$ 400

$E(1,1)=$ 1200

$E(0,1)=$ 700

$E(-1,1)=$ 1400

Interest value: 400

***E*(1,0):** SSD der Pixelwerte von ☐ und ⬚ :
$(80-90)^2+(90-80)^2+(90-100)^2+(90-100)^2 = \underline{400}$

Interest value: Minimum der 4 Werte

- Why is an orientation, the 'dominant' gradient direction of the image region, assigned to each interest point in SIFT?

-> **Invariance to rotation**

- With SIFT, a feature vector is calculated using gradient histograms in 4x4 windows. A gradient histogram has 8 bins and the feature vector therefore has a total of **128** elements.

Hier Erklärung dafür in den Slides: SIFT