



Institutt for datateknikk
og informasjonsvitenskap

**KONTINUASJONSEKSAMEN I EMNE TDT4195
BILDETEKNIKK
LØRDAG 21. AUGUST 2010
KL. 09.00 – 13.00**

LØSNINGSFORSLAG

OPPGAVE 1 Bildebehandling – Grunnleggende begreper

a) Forskjeller mellom det menneskelige øyet og et kamera:

1. Fokuseringsmetode:

The eye adjusts the power of the lens to focus an image on the retina. The distance between the lens and the retina is not adjusted - apart from the small change that might result from the change of curvature of the lens. The ciliary muscles apply a force to the elastic lens that adjusts its curvature.

A camera adjusts the focusing by changing the distance between the camera and the sensitive plane. There are no changes in the curvature of individual lenses. A telephoto lens moves its component lenses relative to each other to change the power of the composite.

2. Tilpasning som respons på lysintensitet:

The eye has an iris diaphragm that adjusts the aperture of the lens. A very similar mechanism is found in a camera.

Both the eye and the camera have a secondary method of tolerating different light intensities.

The eye has two types of light sensitive cells: cones, effective with larger incident light energies and rods that respond in lower light intensities. The two types of cell have an overlapping working range.

The camera has passed through several generations with different technologies. The current CCD camera has elements on a chip. These have an adjustable integration

time that can be used to adapt the elements to different light intensities. The greatest limitation on the camera's output dynamic range is the convention of truncating each pixel value to 8 bits.

3. Respons på farge:

Colour is the common name for a combination of various light frequencies in the human visible range.

The human eye has three types of cone receptor. These have a maximum response in the red, green and blue components of the visible spectrum. The different cones have different pigments that filter the incoming light energy so as to tune the response.

Other animals have different spectral sensitivities. Dogs and cats have greater sensitivity in red and blue than humans. Insects generally respond in ultra-violet.

A lower cost colour CCD camera has a similar arrangement. A Bayer mask is placed over the CCD chip. The mask is arranged as a series of four colour filters: 1 each for red and blue, and two green.

A higher cost camera has 3 CCD chips for red, green and blue. The splitting of the incident energy is achieved by a dichroic beam splitter prism. Whereas the Bayer mask may absorb 2/3 of the incident energy, the prism passes almost all the incoming energy to the CCD chip.

The CCD chip responds to electromagnetic energy in bands outside the visible range for humans. Additional filters must be used to screen out, for example, thermal energy.

4. Synsfelt (field of view):

The human eye is a nearly spherical structure. The retina extends over a very large part of the inner surface. The lens is far from perfect but the incident light can be interpreted by the brain even though the focusing away from the optic axis is imprecise.

The result of the design is that there is image recognition out to at least 70 degrees off the axis. At larger angles, the eye/brain response becomes a sensitivity to movement.

The mounting of the eyes in the skull allow a rapid rotation to centre a new point of interest at the intersection of the optical axis with the retina.

Evolution has adjusted the distribution of cells over the retina so as to exploit the physics of image formation. The yellow spot is an area with a dense population of cone cells giving high resolution colour vision at the intersection of the optic axis with the retina. Further off axis the density of cone cells decreases relative to the rods. This tends towards monochrome vision with good light sensitivity where the focusing is not precise. Towards the periphery, the rods give effective sensitivity to changes in incident energy due to motion.

The camera has a sensitive element in the form of a plane. The field of view is ultimately set by the size of the plane area. Accurate focusing may not be available over the entire sensor depending on lens parameters. The sensitive plane has a uniform covering of light sensitive elements.

The general design of cameras encourages their use where the field of view is only a few degrees.

b) Svar:

1. An effective criterion is the need to have 4-connected representation of all of the structures in the image. If each pixel is a square of side h , then the smallest width of the image on the pixel grid should be $\sqrt{2} \cdot h$. These values can be projected back onto the real world.

In this example, 2048 pixels span 100 cms. Thus $h = 100/2048$ cms. in the real world.

The narrowest possible width is $\sqrt{2} \cdot h = 1.414 * 100/2048 = 0.069$ cms.

The narrowest width is given as 5mm = 0.5 cms. Since this is nearly an order greater than the width, it is certain to be represented by 4-connected pixels and will be adequately sampled.

2. The second region of interest has a width of 0.5 mm. This is slightly below the 0.69 mm. criterion. This region will not be represented by 4-connected pixels.

- c) The diagram should show: $f(x, y)$ transformed into $F(u, v)$, the real and imaginary parts of $F(u, v)$ being multiplied by $G(u, v)$ to give a new Fourier transform, $Q(u, v)$. $Q(u, v)$ should be inverse transformed to give $q(x, y)$.

OPPGAVE 2 Bildebehandling – Kantbaserte metoder

- a) Emphasis of transitions along the x (horizontal) direction

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Emphasis of transitions along the y (vertical) direction

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

- b) The Sobel masks, taken together, simulate the grad function.

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} \vec{i} + \frac{\partial f(x, y)}{\partial y} \vec{j}$$

- c) An edge is regarded as a vector quantity because it has a magnitude and a direction.
- d) The grad function has a direction that is the greatest change of intensity. The normal concept of an edge is orthogonal to the grad direction.
- e) The tendency to increase noise by differencing/differentiation can be reduced by smoothing the data beforehand. This technique is used in difference of Gaussian, Canny and Laplacian of Gaussian.

OPPGAVE 3 Bildebehandling – Filtrering i det romlige domenet

- a) Motion is directional and the blur it produces is effectively a weighting of pixels adjacent in the direction of motion. A convolution kernel that is designed to reduce motion blur will have a reflection symmetry in a line that is in the direction of motion. There will also be a reflection symmetry in the kernel in a line orthogonal to the direction of motion through the middle of the kernel.

- b) original = high frequency + low frequency

high frequency = original - low frequency

high frequency boost = original + k* high frequency
 = original + k*(original - low frequency)
 = (1+k)*original - k*low frequency

Thus a high frequency boost image can be computed from the original and a blurred version of the original.

- c) The method depends on convolving with a smoothing filter that adjusts to smooth in a direction orthogonal to the line of greatest slope.

steps:

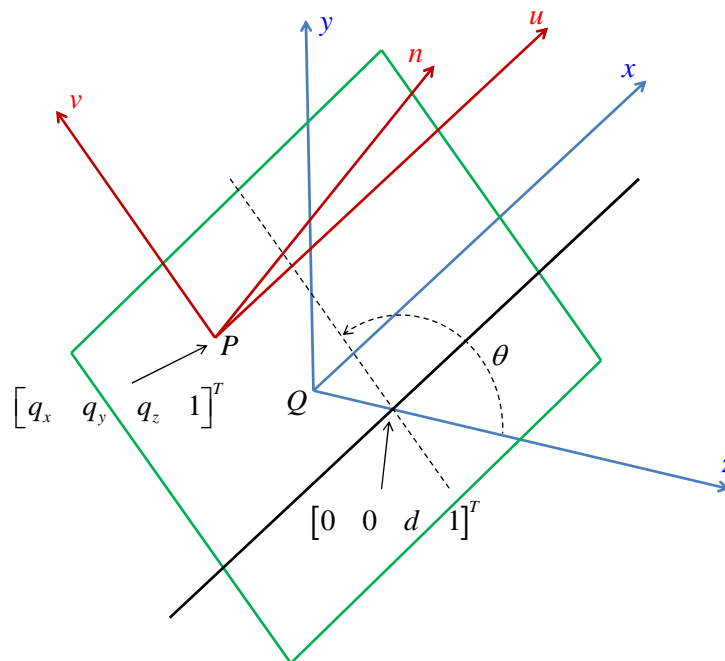
1. form an edge enhanced image eg. using Sobel masks pixel by pixel:
2. refer to the edge enhanced image to estimate the direction of the line of greatest slope
3. create a smoothing kernel that is orientated along the orthogonal direction
4. apply the kernel at that pixel position.

OPPGAVE 4 Grafikk – Perspektivprosjeksjon

En metode for å løse denne oppgaven, kan være:

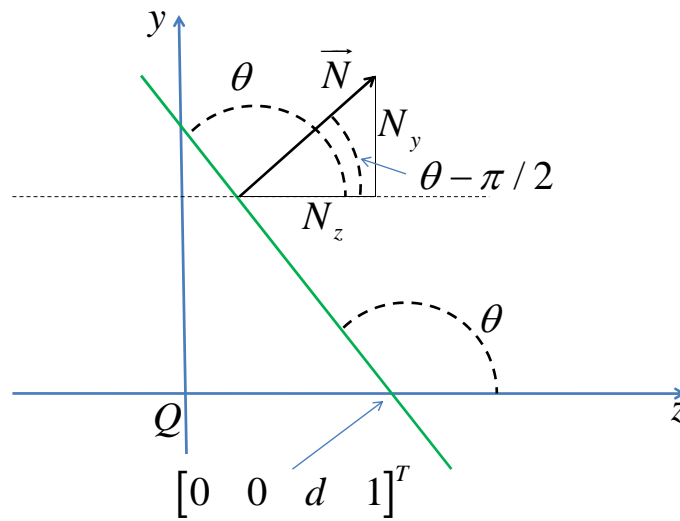
1. Transformere objektet fra verdenskoordinatsystemet til et kamerakoordinatsystem
2. Avbilde (projisere) objektet i kamerakoordinatsystemet
3. Transformere bildet (det avbildete objektet) tilbake til verdenskoordinatsystemet

Vi kan definere kamerakoordinatsystemet med akser u , v og n slik at projeksjonssenteret er i origo og at projeksjonsplanet står vinkelrett på n -aksen. Basis i verdenskoordinatsystemet velger vi som $[e_x, e_y, e_z, Q]^T$ og i kamerakoordinatsystemet som $[e_u, e_v, e_n, P]^T$.



Figur 1

Vi vil først finne matrisen for avbildning i kamerakoordinatsystemet. Til det trengs avstanden mellom projeksjonssenteret og avbildningsplanet. En enkel måte til å bestemme denne størrelsen er å anvende likningen for planet på implisitt form. Vi trenger da en normalisert normal til planet. Forutsetningene for oppgaven gir at x -komponenten av normalen er 0.



Figur 2

Vi får:

$$N_x = 0$$

$$N_y = \sin(\theta - \pi/2) = -\cos\theta$$

$$N_z = \cos(\theta - \pi/2) = \sin\theta$$

Den implisitte likningen for planet er:

$$N_x \cdot x + N_y \cdot y + N_z \cdot z + D = 0$$

Det står igjen å bestemme D . Planet skjærer z -aksen i punktet $[0 \ 0 \ d \ 1]^T$ og vi får:

$$N_x \cdot 0 + N_y \cdot 0 + N_z \cdot d + D = 0$$

$$D = -N_z \cdot d$$

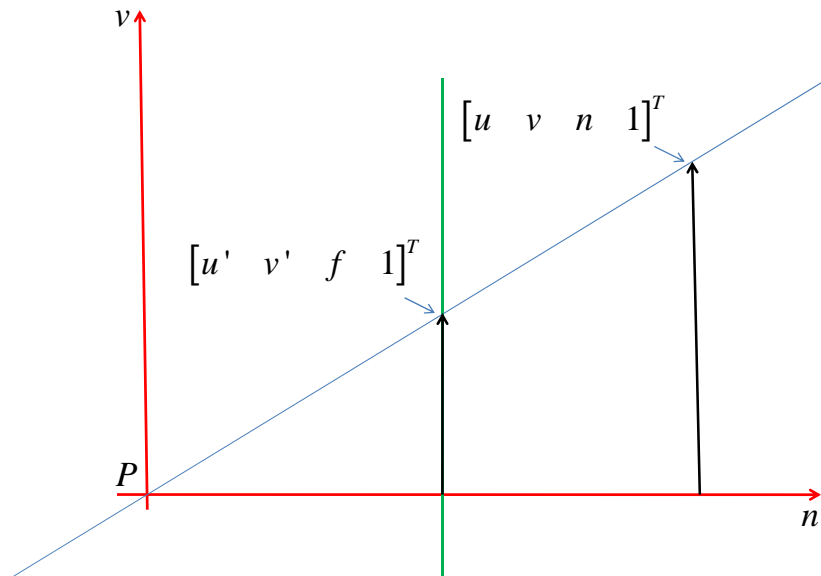
Med alt dette innsatt blir den implisitte likningen for planet:

$$-y \cdot \cos\theta + z \cdot \sin\theta - d \cdot \sin\theta = 0$$

$$\underline{(z - d) \cdot \sin\theta - y \cdot \cos\theta = 0}$$

Avstanden til et punkt utenfor planet fås ved å sette punktets koordinater inn i denne likningen. For å få rett fortegn på avstanden, må likningen beholdes slik den er (ikke omordnes slik at fortegnene endres). Avstanden mellom projeksjonssenter og avbildningsplanet blir derfor:

$$(1) \quad \underline{f = (q_z - d) \cdot \sin \theta - q_y \cdot \cos \theta}$$



Figur 3

Av figur 3 får vi:

$$\frac{v'}{v} = \frac{f}{n} \Rightarrow v' = \frac{v}{n/f}$$

Tilsvarende:

$$u' = \frac{u}{n/f}$$

Dessuten:

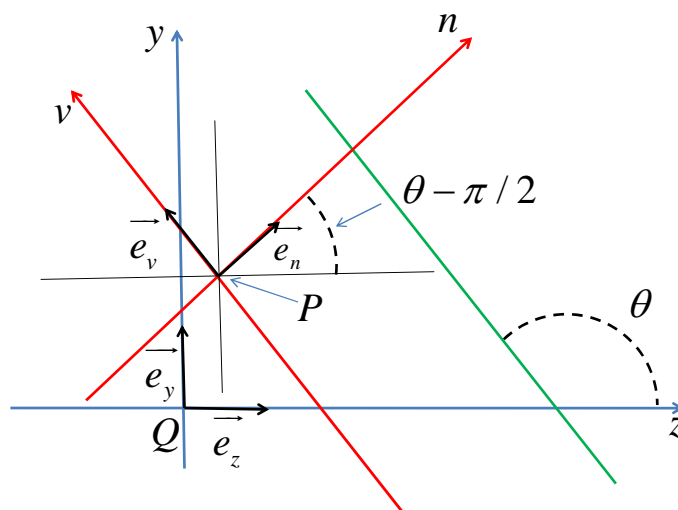
$$n' = f$$

Dette kan stilles sammen på matriseform:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ n \\ 1 \end{bmatrix} = \begin{bmatrix} u \\ v \\ n \\ n/f \end{bmatrix}$$

Den søkte matrisen for avbildning i kamerakoordinatsystemet er:

$$(2) \quad M_{avb} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix}$$



Figur 4

Figur 4 gir:

$$(3) \quad \begin{aligned} \vec{e}_n &= \vec{e}_y \sin(\theta - \pi/2) + \vec{e}_z \cos(\theta - \pi/2) = -\vec{e}_y \cos \theta + \vec{e}_z \sin \theta \\ \vec{e}_v &= \vec{e}_y \cos(\theta - \pi/2) - \vec{e}_z \sin(\theta - \pi/2) = \vec{e}_y \sin \theta + \vec{e}_z \cos \theta \end{aligned}$$

Dessuten har vi:

$$(4) \quad \begin{aligned} \vec{e}_u &= \vec{e}_x \\ Q &= -q_x \cdot \vec{e}_x - q_y \cdot \vec{e}_y - q_z \cdot \vec{e}_z - P \end{aligned}$$

Metoden vi skal bruke for å konvertere representasjoner i verdenskoordinatsystemet til representasjoner i kamerakoordinatsystemet, krever uttrykk for basis i verdenskoordinatsystemet ved basis i kamerakoordinatsystemet.

$$\begin{aligned}
 \vec{e}_x &= \gamma_{11}\vec{e}_u + \gamma_{12}\vec{e}_v + \gamma_{13}\vec{e}_n \\
 \vec{e}_y &= \gamma_{21}\vec{e}_u + \gamma_{22}\vec{e}_v + \gamma_{23}\vec{e}_n \\
 \vec{e}_z &= \gamma_{31}\vec{e}_u + \gamma_{32}\vec{e}_v + \gamma_{33}\vec{e}_n \\
 Q &= \gamma_{41}\vec{e}_u + \gamma_{42}\vec{e}_v + \gamma_{43}\vec{e}_n + P
 \end{aligned}
 \tag{5}$$

Dette kan skrives på matriseform:

$$\begin{bmatrix} e_x \\ e_y \\ e_z \\ Q \end{bmatrix} = M \cdot \begin{bmatrix} e_u \\ e_v \\ e_n \\ P \end{bmatrix}$$

(6)

$$M = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & 0 \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & 0 \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & 0 \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & 1 \end{bmatrix}$$

Et objektpunkt har representasjonen $a = [x_p \quad y_p \quad z_p \quad 1]^T$ i verdenskoordinatsystemet. I

kamerakoordinatsystemet har det samme punktet representasjonen $b = [u_p \quad v_p \quad n_p \quad 1]^T$. Vi får:

$$a^T \begin{bmatrix} e_x \\ e_y \\ e_z \\ Q \end{bmatrix} = a^T \cdot M \cdot \begin{bmatrix} e_u \\ e_v \\ e_n \\ P \end{bmatrix} = b^T \cdot \begin{bmatrix} e_u \\ e_v \\ e_n \\ P \end{bmatrix}$$

som gir:

$$a^T \cdot M = b^T$$

(7)

$$b = M^T \cdot a$$

Matrisen som konverterer representasjoner i verdenskoordinatsystemet til representasjoner i kamerakoordinatsystemet er altså M^T . Fra (4) får vi:

$$(8) \quad M_{w \rightarrow c} = M^T = \begin{bmatrix} \gamma_{11} & \gamma_{21} & \gamma_{31} & \gamma_{41} \\ \gamma_{12} & \gamma_{22} & \gamma_{32} & \gamma_{42} \\ \gamma_{13} & \gamma_{23} & \gamma_{33} & \gamma_{43} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Likningen (3) – (4) gir:

$$(9) \quad \begin{aligned} \vec{e}_x &= \vec{e}_u \\ \vec{e}_y &= \vec{e}_v \sin \theta - \vec{e}_n \cos \theta \\ \vec{e}_z &= \vec{e}_v \cos \theta + \vec{e}_n \sin \theta \\ P &= q_x \cdot \vec{e}_u + (q_y \sin \theta + q_z \cos \theta) \cdot \vec{e}_v - (q_y \cos \theta - q_z \sin \theta) \cdot \vec{e}_n + Q \end{aligned}$$

Dette gir konverteringsmatrisen:

$$(10) \quad M_{w \rightarrow c} = \begin{bmatrix} 1 & 0 & 0 & q_x \\ 0 & \sin \theta & \cos \theta & q_y \sin \theta + q_z \cos \theta \\ 0 & -\cos \theta & \sin \theta & -q_y \cos \theta + q_z \sin \theta \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Vi har også bruk for matrisen som konverterer fra kamerakoordinatsystemer til verdenskoordinatsystemet. Denne matrisen kan settes opp på samme måte med utgangspunkt i likningene (3) – (4):

$$(11) \quad M_{c \rightarrow w} = \begin{bmatrix} 1 & 0 & 0 & -q_x \\ 0 & \sin \theta & -\cos \theta & -q_y \\ 0 & \cos \theta & \sin \theta & -q_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Den søkte matrisen blir:

$$(12) \quad \underline{\underline{M_{resultat} = M_{c \rightarrow w} \cdot M_{avb} \cdot M_{w \rightarrow c}}}$$

OPPGAVE 5 Grafikk – Phongs refleksjonsmodell

a) Phongs refleksjonsmodell er slik formulert i forelesningene:

$$I_{\lambda} = \sum_i \left\{ \frac{1}{a + bd_i + cd_i^2} \left[k_{\lambda d} I_{i\lambda d} (l_i \cdot n) + k_{\lambda s} I_{i\lambda s} (h_i \cdot n)^{\alpha} \right] + k_{\lambda a} I_{i\lambda a} \right\} + k_{\lambda a} I_{\lambda a}$$

Lysstyrkene $I_{i\lambda d}$ og $I_{i\lambda s}$ mellom hakeparentesene tilhører de spesifikke lyskildene som kan modelleres. Det er faktoren:

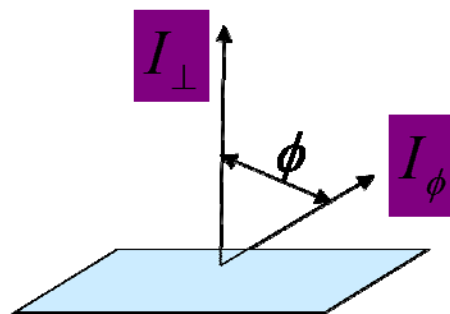
$$\frac{1}{a + bd_i + cd_i^2}$$

som anvendes til modellering av lyskildenes spesifikke egenskaper. d er avstanden fra lyskilden til objektpunktet. For punktluskilder er a og b begge lik 0. For å hindre alt for stor lysstyrke ved små avstander, kan imidlertid a gis en verdi litt større enn 0. For lyskilder med utstrekning sette også b til en verdi større enn 0. Dette gjør at lysstyrken avtar langsommere med avstanden enn tilfellet er for punktluskilder.

b) For en perfekt matt flate med lysspredning i alle retninger, gjelder Lamberts cosinuslov:

$$I_{\phi} = I_{\perp} \cos(\phi)$$

der I_{\perp} er lysstyrken vinkelrett ut fra flaten mens I_{ϕ} er lysstyrken i en retning som danner vinkelen ϕ med normalen.



Figur 5

Arealet av flaten er dA . Sett under vinkelen ϕ er den tilsynelatende størrelsen av flaten:

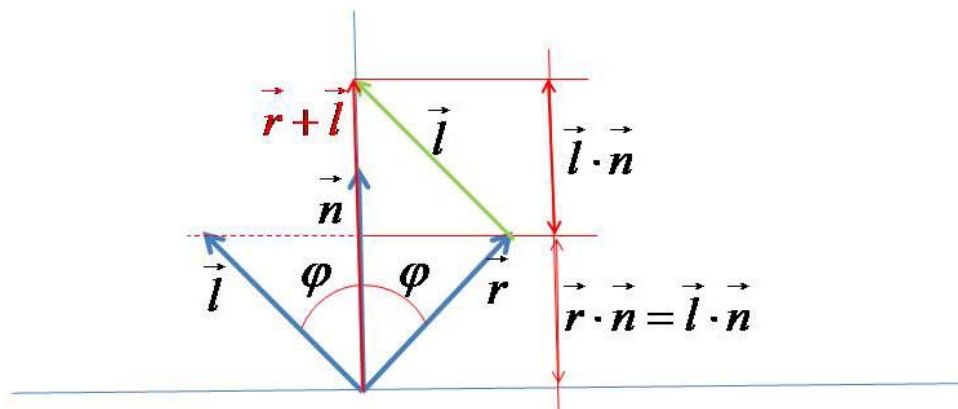
$$dA \cdot \cos \phi$$

Lystettheten for flaten blir da:

$$\frac{I_{\phi}}{\text{tilsynelatende areal}} = \frac{I_{\perp} \cos \phi}{dA \cdot \cos \phi} = \frac{I_{\perp}}{dA} = \text{konstant}$$

Lystettheten er altså uavhengig av betraktningsretningen.

c)



Figur 6

Av figuren ser vi:

$$\vec{r} + \vec{l} = (\vec{r} \cdot \vec{n} + \vec{l} \cdot \vec{n})\vec{n} = 2(\vec{l} \cdot \vec{n})\vec{n}$$

$$\Downarrow$$

$$\underline{\underline{r = 2(\vec{l} \cdot \vec{n})\vec{n} - \vec{l}}}$$

OPPGAVE 6 Grafikk – Geometriske transformasjoner

- a) Resultatet av å anvende matrisen M på punktet P er representasjonen av punktet P i system 2.

Siden vektoren \vec{u} , \vec{v} og \vec{n} er ortonormale, blir resultatet når en anvender matrisen M på hver av dem:

$$M \cdot \vec{u} = \begin{bmatrix} u_{2x} & u_{2y} & u_{2z} & 0 \\ v_{2x} & v_{2y} & v_{2z} & 0 \\ n_{2x} & n_{2y} & n_{2z} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_{2x} \\ u_{2y} \\ u_{2z} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$M \cdot \vec{v} = \begin{bmatrix} u_{2x} & u_{2y} & u_{2z} & 0 \\ v_{2x} & v_{2y} & v_{2z} & 0 \\ n_{2x} & n_{2y} & n_{2z} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} v_{2x} \\ v_{2y} \\ v_{2z} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$M \cdot \vec{n} = \begin{bmatrix} u_{2x} & u_{2y} & u_{2z} & 0 \\ v_{2x} & v_{2y} & v_{2z} & 0 \\ n_{2x} & n_{2y} & n_{2z} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} n_{2x} \\ n_{2y} \\ n_{2z} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Dette betyr at matrisen roterer aksene i system 2 slik at de faller sammen med aksene i system 1. Dersom vi lar alle objekter, punktet P inklusive roter med aksene i system 2, vil vi få objektene representasjoner i system 2.

- b) En rotasjonsmatrise har formen:

$$M_{rot} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

I tillegg kreves at matrisen er ortogon, det vil si:

$$M^{-1} = M^T$$

En matrise for uniform skalering har formen:

$$M_{uskal} = \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Vi undersøker de to mulige produktene av de to matrisene:

$$M_{rot} \cdot M_{uskal} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} sa_{11} & sa_{12} & sa_{13} & 0 \\ sa_{21} & sa_{22} & sa_{23} & 0 \\ sa_{31} & sa_{32} & sa_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$M_{uskal} \cdot M_{rot} = \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} sa_{11} & sa_{12} & sa_{13} & 0 \\ sa_{21} & sa_{22} & sa_{23} & 0 \\ sa_{31} & sa_{32} & sa_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Vi ser at de to produktene er like og konklusjonen blir dermed at rotasjon og uniform skalering kommuterer.

c) En affin transformasjon er en transformasjon som bevarer parallellitet. Alle de tre basistransformasjonene:

- rotasjon
- skalering
- translasjon

er affine. Den generelle formen for matrisen for en affin transformasjon er:

$$M_{affin} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Perspektivprojeksjon er eksempel på en transformasjon som ikke er affin. Her bevarer generelt parallelle linjer ikke parallelliteten under transformasjonen.

d) Den generelle matrisen for affine transformasjoner, se deloppgave c), har 12 koeffisienter. til å bestemme disse koeffisientene trenger vi 12 likninger som ikke er lineært avhengige. Det er tilstrekkelig å kjenne representasjonen til 4 tredimensjonale punkt før, P , og etter, P' , transformasjonen. Hvert punkt gir tre likninger.

Likningssystemer blir:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} = \begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ P_4' \end{bmatrix}$$

$$P = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} \quad P' = \begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ P_4' \end{bmatrix}$$

Det er et krav at disse likningene er lineært uavhengige. Det blir de dersom punktene ikke er koplanare.