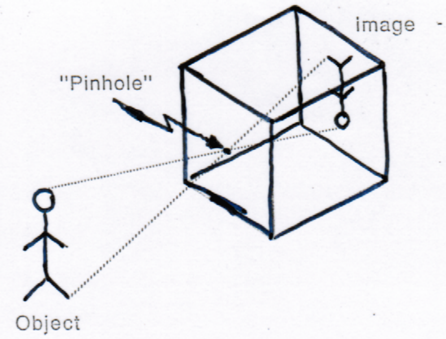
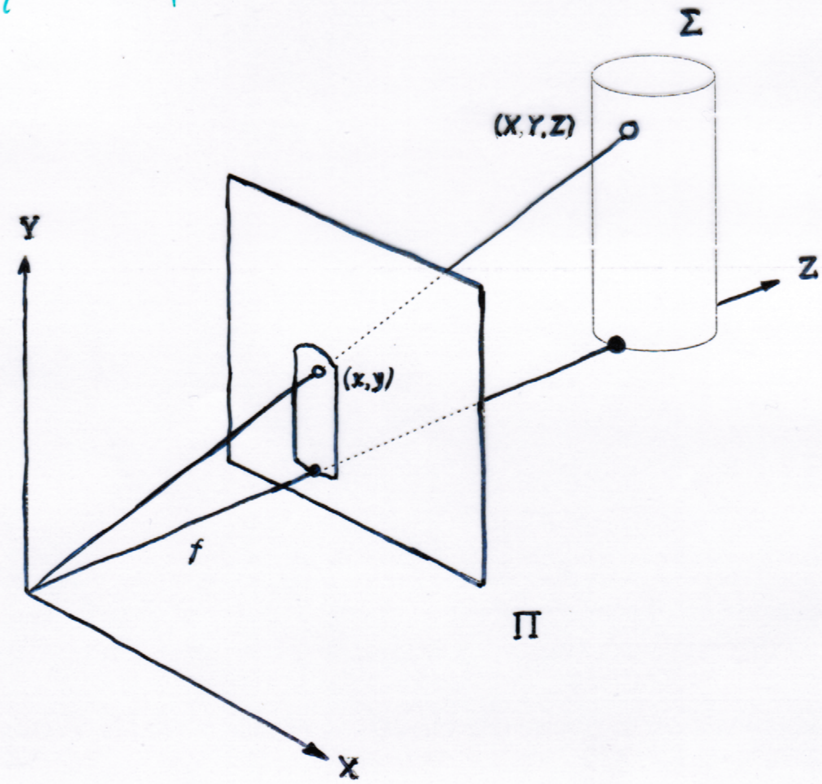


CARACTERÍSTICAS:

— ABERTURA INFINITESIMALMENTE PEQUENA;

— A IMAGEM É FORMADA POR RAIOS DE LUZ EM LINHA RECTA QUE VÃO DO OBJECTO PARA O PLANO DE IMAGEM, ATRAVÉS DA ABERTURA;

— ESTE TIPO DE MAPEAMENTO DAS 3 DIMENSÕES PARA 2 DIMENSÕES É DESIGNADO POR PROJECÇÃO EM PERSPECTIVA;

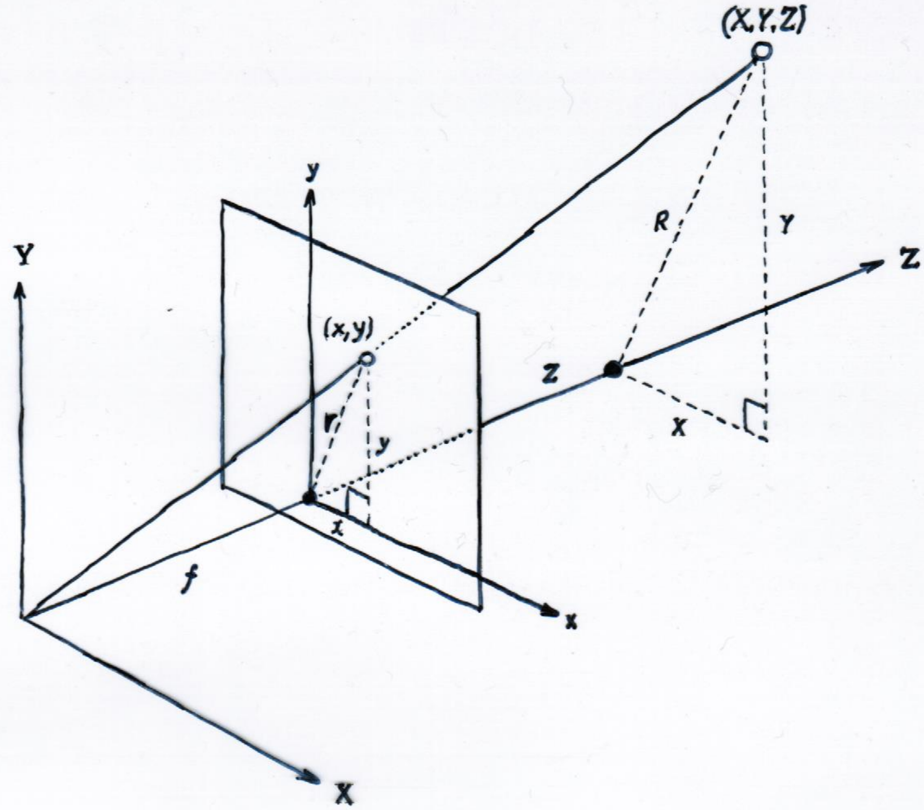


PROJECÇÃO EM PERSPECTIVA

- A PROJECÇÃO EM PERSPECTIVA OU PROJECÇÃO CENTRAL:
  - É A PROJECÇÃO DE UMA ENTIDADE TRI-DIMENSIONAL NUMA SUPERFÍCIE BI-DIMENSIONAL ATRAVÉS DE UM PONTO, O CENTRO DE PROJECÇÃO;

- A PROJECÇÃO EM PERSPECTIVA MODELIZA A GEOMETRIA DA FORMAÇÃO DE IMAGEM NUMA CÂMARA ESTENOPEICA;

- PARA EVITAR A INVERSÃO DA IMAGEM:
  ⇒ VER FIGURA



CÁLCULO DAS COORDENADAS:

– A DISTÂNCIA DO PONTO (X, Y, Z) AO EIXO DOS ZZ É

$$R = \sqrt{X^2 + Y^2}$$

– A DISTÂNCIA DO PONTO PROJECTADO (x, y) À ORIGEM É:

$$r = \sqrt{x^2 + y^2}$$

- O EIXO DOS ZZ, A LINHA DE VISTA DO PONTO (X, Y, Z) E O SEGMENTO DE RECTA DE COMPRIMENTO R (DO PONTO (X, Y, Z) AO EIXO DOS ZZ) FORMAM UM TRIÂNGULO;

- O EIXO DOS ZZ, A LINHA DE VISTA DO PONTO (x, y) E O SEGMENTO DE RECTA DE COMPRIMENTO r (DO PONTO (x, y) AO EIXO DOS ZZ) FORMAM OUTRO TRIÂNGULO SEMELHANTE AO ANTERIOR;

LOGO

$$\frac{f}{Z} = \frac{r}{R}$$

- OS TRIÂNGULOS FORMADOS PELAS COORDENADAS DA CENA X, Y E PELO SEGMENTO R E PELAS COORDENADAS DO PLANO IMAGEM x, y E PELO SEGMENTO r SÃO TAMBÉM TRIÂNGULOS SEMELHANTES

LOGO

$$\frac{x}{X} = \frac{y}{Y} = \frac{r}{R}$$

COMBINANDO AS EQUAÇÕES OBTÉM-SE:

$$\frac{x}{X} = \frac{f}{Z} \wedge \frac{y}{Y} = \frac{f}{Z}$$

$$\Leftrightarrow x = \frac{f}{Z} X \wedge y = \frac{f}{Z} Y$$

CONSIDEROU-SE QUE O CENTRO DE PROJECÇÃO E A ORIGEM COINCIDEM;

- NO CASO GERAL A CÂMARA ESTÁ DESLOCADA E RODADA EM RELAÇÃO AO SISTEMA DE COORDENADAS 3D USADO PARA ESPECIFICAR OS PONTOS NO ESPAÇO;

- TO COMPUTE THE 2-D F.T. ONE CAN USE A 1-D FFT. ALGORITHM:
  - FIRST COMPUTE THE TRANSFORM OF EACH ROW;
  - THEN COMPUTE THE TRANSFORM COLUMN BY COLUMN TAKING AS DATA THE OUTPUT OF THE FIRST COMPUTATION;

- APPLICATIONS OF THE DISCRETE 2D FOURIER TRANSFORM:
  - IMAGE RECONSTRUCTION;
  - IMAGE ENHANCEMENT;
  - IMAGE RESTORATION;
  - THE BOUNDARY OF AN OBJECT CAN BE CODED BY MEANS OF THE APPLICATION OF THE 1-D F.T. (THE COEFFICIENTS OF THE TRANSFORM CAN BE USED AS DESCRIPTORS OF THE BOUNDARY SHAPE).

SPATIAL-DOMAIN TECHNIQUES:

- THE SPATIAL DOMAIN REFERS TO THE AGGREGATE OF PIXELS COMPOSING AN IMAGE;

- SPATIAL-DOMAIN METHODS ARE PROCEDURES THAT OPERATE DIRECTLY ON THESE PIXELS;

- PREPROCESSING FUNCTIONS IN THE SPATIAL DOMAIN MAY BE EXPRESSED AS

$$g(x,y) = h[f(x,y)]$$

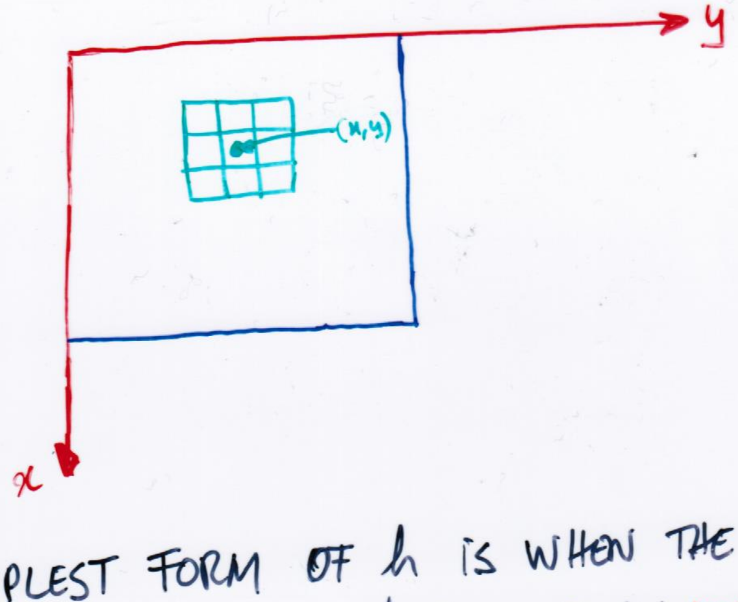
$f(x,y)$ – INPUT IMAGE

$g(x,y)$ – RESULTING IMAGE

h – AN OPERATOR ON f, DEFINED OVER SOME NEIGHBORHOOD OF (x,y).

- h MAY ALSO OPERATE ON A SET OF INPUT IMAGES;

- USUALLY THE NEIGHBORHOOD OF (x,y) IS DEFINED BY MEANS OF A SQUARE OR RECTANGULAR SUBIMAGE AREA CENTERED AT (x,y).



- THE SIMPLEST FORM OF h IS WHEN THE NEIGHBORHOOD IS 1x1. IN THIS CASE h BECOMES AN INTENSITY MAPPING OR TRANSFORMATION.

- ONE OF THE SPATIAL-DOMAIN TECHNIQUES USED MOST FREQUENTLY IS BASED ON THE USE OF THE SO-CALLED convolution masks (or templates, or windows, or filters).

- A MASK IS BASICALLY A SMALL 2D ARRAY;

- LET US SUPPOSE WE ARE TO DETECT ISOLATED PIXELS WHOSE INTENSITIES ARE DIFFERENT FROM THE BACKGROUND; LET US ALSO SUPPOSE THAT THE IMAGE IS OF BASICALLY CONSTANT INTENSITY.

- WE CAN DETECT THESE POINTS BY USING THE MASK:

-1	-1	-1
-1	8	-1
-1	-1	-1

- THE CENTER OF THE MASK IS MOVED AROUND THE IMAGE;

- AT EACH PIXEL POSITION WE MULTIPLY EVERY PIXEL THAT IS CONTAINED WITHIN THE MASK AREA BY THE CORRESPONDING MASK COEFFICIENT; THE RESULTS OF THE NINE MULTIPLICATIONS ARE SUMMED

w_1 $(x-1, y-1)$	w_2 $(x-1, y)$	w_3 $(x-1, y+1)$
w_4 $(x, y-1)$	w_5 (x, y)	w_6 $(x, y+1)$
w_7 $(x+1, y-1)$	w_8 $(x+1, y)$	w_9 $(x+1, y+1)$

- IF $w_1, w_2, w_3, \dots, w_9$ REPRESENT MASK COEFFICIENTS, AND IF WE CONSIDER THE 8-NEIGHBOURS OF (x, y) WE MAY GENERALIZE AS:

$$h[f(x,y)] = w_1 f(x-1, y-1) + w_2 f(x-1, y) + w_3 f(x-1, y+1) + \\ + w_4 f(x, y-1) + w_5 f(x, y) + w_6 f(x, y+1) + \\ + w_7 f(x+1, y-1) + w_8 f(x+1, y) + w_9 f(x+1, y+1)$$

- SMOOTHING

→ SMOOTHING OPERATIONS ARE USED FOR REDUCING NOISE AND OTHER SPURIOUS EFFECTS THAT MAY BE PRESENT IN AN IMAGE AS A RESULT OF SAMPLING, QUANTIZATION, TRANSMISSION, OR DISTURBANCES IN THE ENVIRONMENT DURING IMAGE ACQUISITION.

NEIGHBORHOOD AVERAGING:

$$g(x,y) = \frac{1}{P} \sum_{(n,m) \in S} f(n,m) \quad \forall (x,y) \text{ in } f(x,y)$$

- EACH PIXEL OF $g(x,y)$ IS OBTAINED BY AVERAGING THE INTENSITY VALUES OF THE PIXELS OF f CONTAINED IN A PREDEFINED NEIGHBORHOOD OF (x,y).

MEDIAN FILTERING:

- MEDIAN FILTERING IS A SPECIAL CASE OF RANK FILTERING;
- IN RANK FILTERING THE GRAY LEVEL $g(x,y)$ IS CHOSEN ON THE BASIS OF THE RELATIVE RANK OF PIXELS IN THE NEIGHBORHOOD OF (x,y).
- LET N BE THE NUMBER OF GRAY LEVELS IN A REGION S

$$f_i, \quad i = 1, 2, \ldots, N \qquad f_i \text{ - GRAY LEVEL VALUES}$$

- ORDERING THESE VALUES IN INCREASING VALUE, WE HAVE

$$R = \{f_1, f_2, \ldots, f_N\} \quad \text{WITH } f_i \leq f_{i+1}$$

- THE OUTPUT GRAY LEVEL VALUE $g(x,y)$ IS

$$g(x,y) = RANK_j \, R$$

WHERE $RANK_j$ IS THE GRAY LEVEL AT POSITION, OR RANK, j IN R.

- IF $j=1$ WE GET THE <u>min</u> FILTER, WHERE THE 1st ELEMENT OF R;
- CHOOSING $j=N$ YIELDS THE <u>max</u> FILTER;

- IN THE CASE OF N ODD, $j=(N+1)/2$ YIELDS THE MEDIAN FILTER;

- APPLICATION:
  - LET US ASSUME WE HAVE AN IMAGE CORRUPTED BY IMPULSE OR "SPIKE"-LIKE NOISE.
  - LOW-PASS FILTERING (SMOOTHING) WOULD TEND TO DISTRIBUTE THIS NOISE INTENSITY OVER THE PIXELS SURROUNDING THE NOISE SPIKE.
  - THE MEDIAN FILTER USUALLY REMOVES THIS TYPE OF IMAGE NOISE.

- THE MEDIAN IS A NONLINEAR FILTER AS

  - $med(R_1+R_2) \neq med(R_1) + med(R_2)$
  - $med(KR) = K\, med(R)$
  - $med(K+R) = K + med(R)$

OTHER PROPERTIES OF THE MEDIAN FILTER:

1. THE MEDIAN FILTER REDUCES THE VARIANCE OF THE INTENSITIES IN THE IMAGE.

2. INTENSITY OSCILLATIONS WITH A PERIOD LESS THAN THE WINDOW WIDTH ARE SMOOTHED.

3. MEDIAN FILTERS PRESERVE CERTAIN EDGE SHAPES.

4. IN THE APPLICATION OF A MEDIAN FILTER, NO NEW GRAY LEVEL VALUES ARE GENERATED. THE DYNAMIC RANGE OF A MEDIAN FILTERED IMAGE CANNOT EXCEED THAT OF THE INPUT IMAGE.

IMAGE TEMPORAL AVERAGING:

LET $g(x,y)$ BE A NOISY IMAGE WHICH IS FORMED BY THE ADDITION OF NOISE $n(x,y)$ TO AN UNCORRUPTED IMAGE $f(x,y)$:

$$g(x,y) = f(x,y) + n(x,y)$$

- LET US ASSUME THAT THE NOISE IS UNCORRELATED AND HAS ∅ AVERAGE VALUE.

- THEN

$$\bar{g}(x,y) = \frac{1}{K} \sum_{i=1}^{K} g_i(x,y)$$

$g_i(x,y)$, $i = 1, 2, \ldots, K$ - SET OF NOISY IMAGES

- IT CAN BE PROVED THAT

$$E\{\bar{g}(x,y)\} = f(x,y)$$

$$\sigma^2_{\bar{g}}(x,y) = \frac{1}{K} \sigma^2_n(x,y)$$

WHERE $E\{\bar{g}(x,y)\}$ - IS THE EXPECTED VALUE OF $\bar{g}$

$\sigma^2_{\bar{g}}(x,y)$ - IS THE VARIANCE OF $\bar{g}$

$\sigma^2_n(x,y)$ - IS THE VARIANCE OF n

ALL AT COORDINATES (x,y)

- AS K INCREASES, THE VARIABILITY OF THE PIXEL VALUE DECREASES.

- SINCE $E\{\bar{g}(x,y)\} = f(x,y)$ THIS MEANS THAT $\bar{g}(x,y)$ WILL APPROACH THE UNCORRUPTED IMAGE $f(x,y)$ AS THE NUMBER OF NOISY IMAGES USED IN THE AVERAGING PROCESS INCREASES.

- IMPORTANT: IT WAS ASSUMED THAT ALL NOISY IMAGES ARE REGISTERED SPATIALLY, WITH ONLY THE PIXEL INTENSITIES VARYING.

HOMOMORPHIC FILTERING

- THE CONVERSION OF LIGHT FROM A SCENE BY A SENSOR IS A TRANSFER OF 3D RADIOMETRIC INFORMATION TO 2D PIXEL INTENSITY DATA.

- AN APPROXIMATE MODEL OF THE RADIOMETRIC IMAGE FORMATION IS:

$$f(x,y) = r(x,y)\, e(x,y)$$

WHERE

$f(x,y)$ - IS THE OBSERVED IMAGE

$r(x,y)$ – IS THE SURFACE REFLECTIVITY FUNCTION, SO IT REPRESENTS THE IMAGE VARIATION DUE TO THE SURFACE REFLECTIVITY

$e(x,y)$ – IS THE INCIDENT SURFACE ILLUMINATION, SO IT REPRESENTS THE IMAGE GRAY-LEVEL VARIATION DUE TO THE ILLUMINATION SOURCE.

$$0 < e(x,y) < \infty$$

– USUALLY $e(x,y)$ VARIES SLOWLY THROUGHOUT THE IMAGE; THEREFORE IT HAS LOW-FREQUENCY CHARACTERISTICS;

– THE IMAGE HISTOGRAM IS STRONGLY A FUNCTION OF $e(x,y)$;

$$0 < r(x,y) < 1$$

$\Rightarrow$ $r(x,y)$ CONVEYS SIGNIFICANT INFORMATION CONCERNING THE SURFACE CHARACTERISTICS OF SCENE OBJECTS;

$\rightarrow$ IN PARTICULAR ABRUPT CHANGES IN $r(x,y)$ COULD INDICATE SIGNIFICANT EDGES, OR TRANSITIONS TO DIFFERENT OBJECTS.

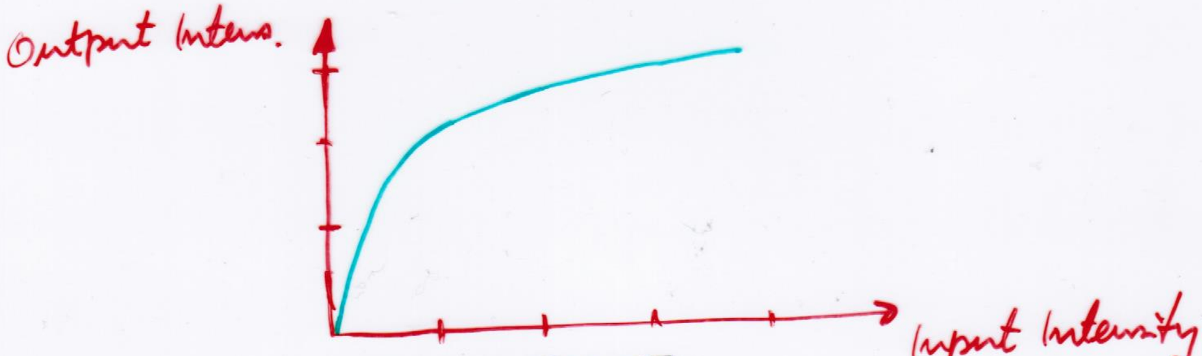
OBJECTIVES OF HOMOMORPHIC FILTERING:

- AN OUTPUT IMAGE REPRESENTATION THAT USES THE DYNAMIC RANGE OF THE GRAY LEVEL VARIABLE;

- ENHANCEMENT OF THE HIGH-FREQUENCY REGIONS OF THE IMAGE;

- THE IMAGE FORMATION MODEL IS MULTIPLICATIVE, RATHER THAN ADDITIVE; THEREFORE DIRECT APPLICATION OF LINEAR FILTERING TECHNIQUES IS QUESTIONABLE; TO OVERCOME THIS DIFFICULTY ONE CAN USE THE TRANSFORMATION

$$z(x,y) = \ln\{f(x,y)\}$$



LOGARITHMIC TRANSF.

$$z(x,y) = \ln\{r(x,y)\} + \ln\{e(x,y)\}$$

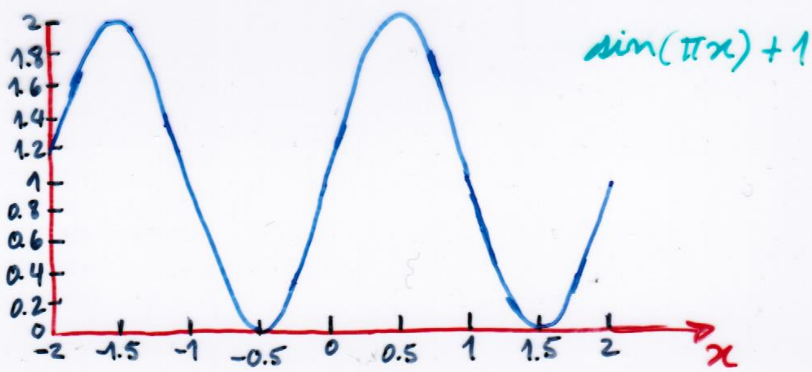
— AFTER THIS TRANSFORMATION THE TRANSFORMED IMAGE IS

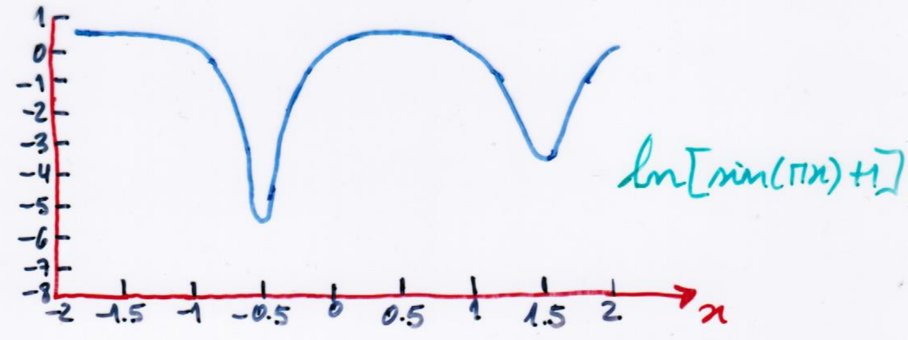
THE SUM OF THE LOGARITHMICALLY TRANSFORMED REFLECTANCE AND ILLUMINATION COMPONENTS.

— FREQUENCY CHARACTERISTICS OF $r(x,y)$ AND $e(x,y)$ UNDER LOGARITHMIC MAPPING:

→ SINCE THE MAPPING IS NONLINEAR, NEW FREQUENCIES ARE CREATED IN $\ln\{r(x,y)\}$ AND $\ln\{e(x,y)\}$;

EXAMPLE:





- THE LOGARITHMICALLY TRANSFORMED VERSION OF THE SINUSOID CONTAINS THE ORIGINAL (FUNDAMENTAL) FREQUENCY, AS WELL AS A NUMBER OF HIGHER FREQUENCY HARMONICS;

- TO ACHIEVE THE ENHANCEMENT OBJECTIVES THE TRANSFORMED IMAGE IS FILTERED WITH A 2-D HIGH-PASS FILTER.

- THE FILTERED AND TRANSFORMED IMAGE $z_f(x,y)$ IS THEN INVERSELY TRANSFORMED BY

$$f_o(x,y) = \exp\{z_f(x,y)\}$$

IMAGE ENHANCEMENT:

- AIM: TO COMPENSATE FOR SEVERAL TYPES OF VARIATIONS, NAMELY, FOR ILLUMINATION VARIATIONS;

IMAGE ENHANCEMENT: POINT OPERATIONS

– LET u BE A GRAY LEVEL;

$$u \in [0, L]$$

$[0, L] \Rightarrow$ SET OF GRAY LEVELS;

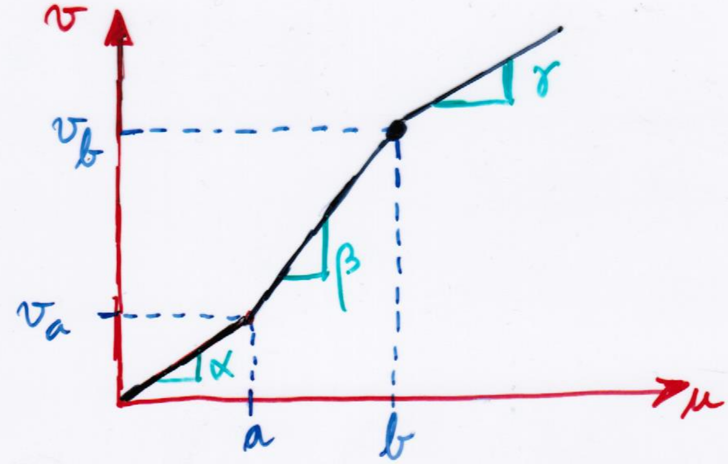
– u IS MAPPED TO A GRAY LEVEL $v \in [0, L]$ ACCORDING TO THE TRANSFORMATION:

$$v = f(u)$$

– CONTRAST STRETCHING:

$$v = \begin{cases} \alpha u & 0 \le u < a \\ \beta(u-a) + v_a & a \le u < b \\ \gamma(u-b) + v_b & b \le u < L \end{cases}$$

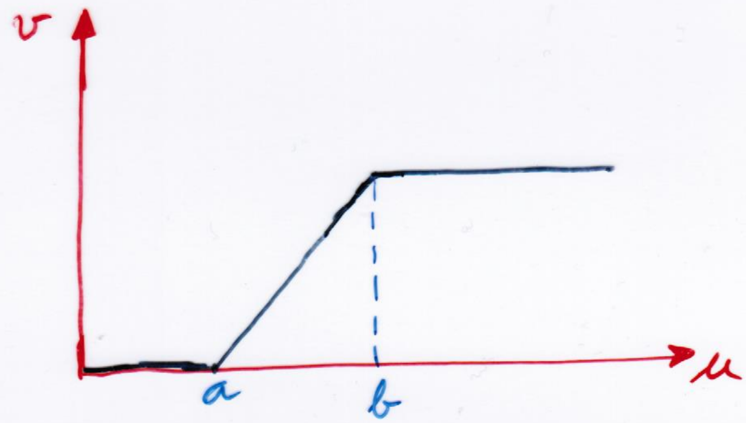
– IN THE REGION WHERE THE CONTRAST IS INCREASED THE SLOPE IS GREATER THAN 1.



- THE VALUES OF a AND b ARE USUALLY DETERMINED BASED ON THE ANALYSIS OF THE HISTOGRAM;

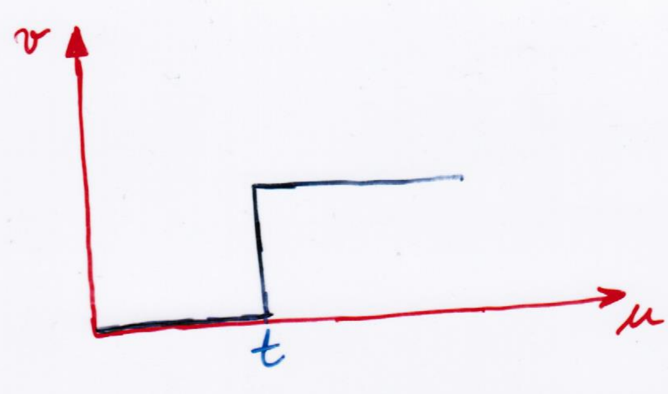
- CLIPPING AND THRESHOLDING

CLIPPING $\Longrightarrow \alpha = \gamma = 0$



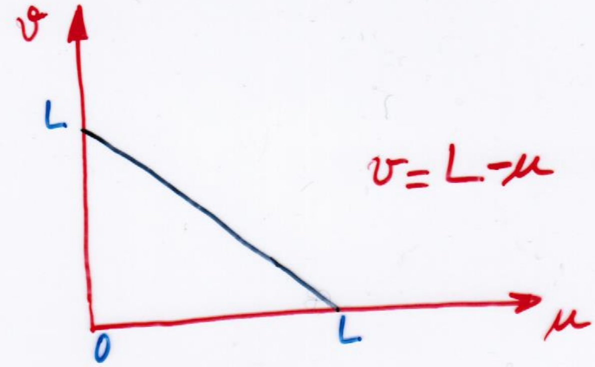
→ USEFUL FOR NOISE REDUCTION WHEN IT IS KNOWN THAT THE INPUT IMAGE GRAY LEVELS RANGE BETWEEN a AND b.

THRESHOLDING $\Longrightarrow (\alpha = \gamma = 0)$ AND $(a = b = t)$



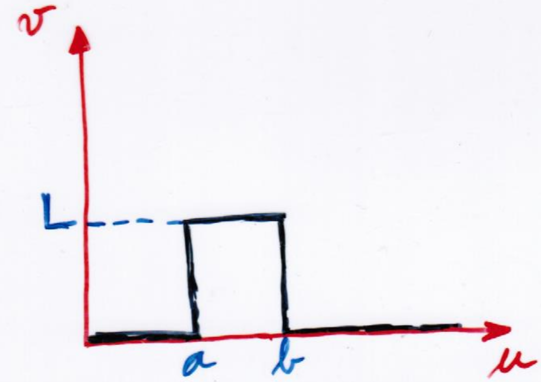
→ THE OUTPUT IMAGE IS BINARY;

- DIGITAL NEGATIVE :



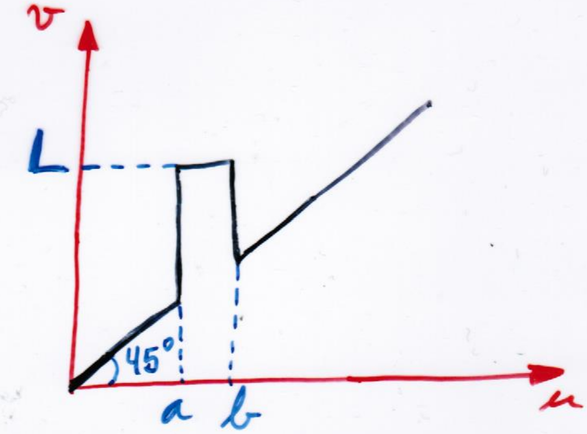
$v = L - \mu$

- SLICING OF THE GRAY LEVELS:



$$v = \begin{cases} L, & a \le \mu \le b \\ 0, & \text{otherwise} \end{cases}$$

NON-BACKGROUND PRESERVING



$$v = \begin{cases} L, & a \le \mu \le b \\ \mu, & \text{otherwise} \end{cases}$$

BACKGROUND PRESERVING

– DYNAMIC RANGE COMPRESSION:

– SOMETIMES THE DYNAMIC RANGE OF THE IMAGE DATA MAY BE VERY LARGE.

– THE DYNAMIC RANGE CAN BE COMPRESSED VIA THE LOGARITHMIC TRANSFORMATION

$$v = c \log_{10}(1 + |u|)$$

c – SCALING CONSTANT

– THIS TRANSFORMATION ENHANCES THE SMALL MAGNITUDE PIXELS COMPARED TO THOSE PIXELS WITH LARGE MAGNITUDES.

– HISTOGRAM MODELING

– THE HISTOGRAM OF AN IMAGE REPRESENTS THE RELATIVE FREQUENCY OF OCCURRENCE OF THE VARIOUS GRAY LEVELS IN THE IMAGE.

– HISTOGRAM-MODELING TECHNIQUES MODIFY AN IMAGE SO THAT ITS HISTOGRAM HAS A DESIRED SHAPE.

– THIS IS A POWERFUL TECHNIQUE FOR IMAGE ENHANCEMENT.

– HISTOGRAM EQUALIZATION

– LET r REPRESENT THE GRAY LEVELS IN AN IMAGE TO BE ENHANCED.

– LET US ASSUME INITIALLY THAT r IS A NORMALIZED, CONTINUOUS VARIABLE LYING IN THE RANGE

$$0 \le r \le 1$$

– FOR ANY r IN THE INTERVAL $[0, 1]$ LET US CONSIDER TRANSFORMATIONS OF THE FORM

$$s = T(r)$$

– IT IS ASSUMED THAT THE TRANSFORMATION FUNCTION T SATISFIES THE CONDITIONS:

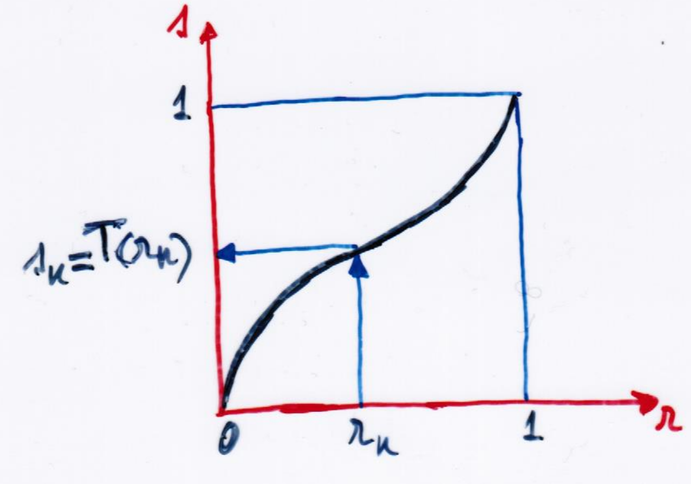
1. $T(r)$ IS SINGLE-VALUED AND MONOTONICALLY INCREASING IN THE INTERVAL $0 \le T(r) \le 1$.

2. $0 \le T(r) (= s) \le 1$ FOR $0 \le r \le 1$.

– CONDITION 1 PRESERVES THE ORDER FROM BLACK TO WHITE IN THE INTENSITY SCALE

– CONDITION 2 GUARANTEES A MAPPING THAT IS CONSISTENT WITH THE ALLOWED 0 TO 1 RANGE OF PIXEL VALUES.

- EXAMPLE OF A TRANSFORMATION SATISFYING THESE CONDITIONS:



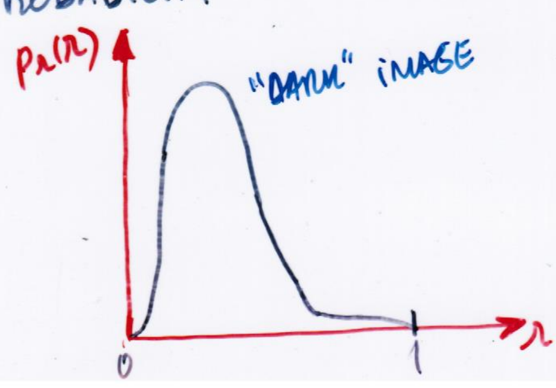
- THE INVERSE TRANSFORMATION FUNCTION FROM s BACK TO r IS DENOTED BY

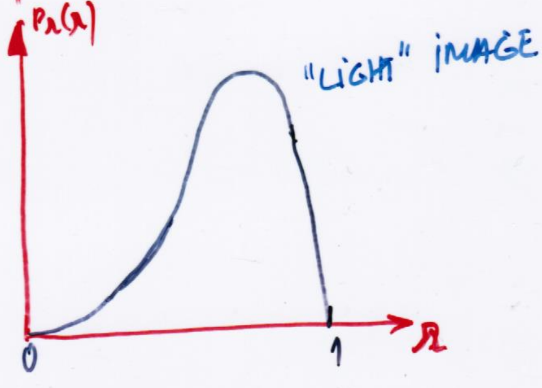
$$r = T^{-1}(s)$$

- IT IS ASSUMED THAT T^{-1} SATISFIES THE TWO CONDITIONS GIVEN ABOVE.

- THE VARIABLES r AND s ARE RANDOM QUANTITIES IN THE INTERVAL $[0,1]$;

- AS SUCH THEY CAN BE CHARACTERIZED BY THEIR PROBABILITY DENSITY FUNCTIONS $p_r(r)$ AND $p_s(s)$.





-LET R BE A RANDOM VARIABLE TAKING VALUES r (67)

WHOSE P.D.F. IS $p_r(r)$ AND WHOSE CUMULATIVE PROBABILITY DISTRIBUTION FUNCTION IS

$$P_R(r)$$

$$P_R(r) = \int_0^r p_r(\omega)\, d\omega$$

-LET S BE THE NEW RANDOM VARIABLE SUCH THAT

$$S = T(R)$$

-WE WANT TO HAVE $T(R)$ SUCH THAT:

-THE RANDOM VARIABLE S HAS A UNIFORM P.D.F. WITH $0 \leq s \leq 1$.

-WE WISH TO MODIFY THE P.D.F., BY CHANGING THE RANDOM VARIABLE.

-AS THE RANDOM VARIABLE S HAS A UNIFORM P.D.F. $p_S(s)$ ONE HAS

$$p_S(s) = 1, \text{ FOR } 0 \leq s \leq 1$$

-THEREFORE

$$P_S(s) = s$$

- WE WISH TO CHANGE FROM THE VARIABLE R TO S SUCH THAT

$P_R(r)$ BECOMES EQUAL TO s WITH $0 \leq s \leq 1$.

- PUTTING

$r = P_R^{-1}(s)$ ONE GETS BY SUBSTITUTING THE VARIABLE IN $P_R(r)$

$$P_R\left(P_R^{-1}(s)\right) = s$$

WHICH IS A UNIFORM DENSITY IN THE INTERVAL OF DEFINITION OF THE TRANSFORMED VARIABLE s.

- THE FUNCTION $T(R)$ IS GIVEN BY:

$$S = T(R) = P_R(r)$$

OR

$$s = T(r) = P_R(r) = \int_0^r p_r(w)\, dw$$

- IN DISCRETE FORM

$$p_r(r_k) = \frac{n_k}{N} \quad , \qquad 0 \leq r_k \leq 1$$

r_k - GRAY LEVELS

$k = 0, 1, \dots, L-1$

N - TOTAL NUMBER OF PIXELS IN THE IMAGE.

– AND THEREFORE

$$s_k = P_R(r_k) = \sum_{i=0}^{k} p_r(r_k)$$

WITH $0 \leq s_k \leq 1$

HISTOGRAM SPECIFICATION:

– IT MAY BE USEFUL TO CHANGE AN IMAGE SO THAT ITS HISTOGRAM IS CHANGED TO A PREDEFINED FORM;

– SUPPOSE THAT A GIVEN IMAGE IS FIRST HISTOGRAM EQUALIZED

$$s = T(r) = \int_0^r p_r(w)\,dw$$

– IF THE DESIRED IMAGE <u>WERE</u> AVAILABLE, ITS LEVELS COULD ALSO BE EQUALIZED BY USING THE TRANSFORMATION FUNCTION

$$v = G(z) = \int_0^z p_z(w)\,dw$$

– THE INVERSE PROCESS, $z = G^{-1}(v)$ WOULD THEN YIELD THE DESIRED LEVELS BACK.

– IN THIS CASE $p_s(s)$ AND $p_v(v)$ WOULD BE IDENTICAL UNIFORM DENSITIES.

– THUS, IF INSTEAD OF USING v IN THE INVERSE PROCESS, WE USE THE INVERSE LEVELS s OBTAINED FROM THE ORIGINAL IMAGE, THE RESULTING LEVELS

$$z = G^{-1}(s)$$

WOULD HAVE THE DESIRED P.D.F. $p_z(z)$.

– ASSUMING THAT $G^{-1}(s)$ IS SINGLE-VALUED, THE PROCEDURE CAN BE SUMMARIZED AS FOLLOWS:

1. EQUALIZE THE LEVELS OF THE ORIGINAL IMAGE;

2. SPECIFY THE DESIRED P.D.F. $p_z(z)$ AND OBTAIN THE TRANSFORMATION FUNCTION $G(z)$ USING:

$$G(z) = \int_0^z p_z(w)\, dw$$

3. APPLY THE INVERSE TRANSFORMATION $z = G^{-1}(s)$ TO THE INTENSITY LEVELS OF THE HISTOGRAM-EQUALIZED IMAGE (STEP 1).

EDGE DETECTION

- EDGE POINTS CAN BE THOUGHT OF AS PIXEL LOCATIONS OF ABRUPT GRAY-LEVEL CHANGE.

- EDGES CHARACTERIZE OBJECT BOUNDARIES AND ARE THEREFORE USEFUL FOR SEGMENTATION, REGISTRATION AND IDENTIFICATION OF OBJECTS IN SCENES.

- THE BOUNDARY REPRESENTATION OF AN OBJECT HAS THE ADVANTAGE OF ITS EASY INTEGRATION INTO A LARGE VARIETY OF OBJECT RECOGNITION ALGORITHMS.

- A LOCAL EDGE IS A SMALL AREA IN THE IMAGE WHERE THE LOCAL GRAY LEVELS ARE CHANGING RAPIDLY IN A SIMPLE (e.g., MONOTONIC) WAY.

- AN EDGE OPERATOR IS A MATHEMATICAL OPERATOR (OR ITS COMPUTATIONAL EQUIVALENT) WITH A SMALL SPATIAL EXTENT DESIGNED TO DETECT THE PRESENCE OF A LOCAL EDGE IN THE IMAGE FUNCTION.

THE MAJORITY OF EDGE OPERATORS CAN BE CLASSIFIED INTO **3** CLASSES:

1. GRADIENT OPERATORS;
2. LAPLACIAN OPERATORS;
3. OPERATORS BASED ON TEMPLATE MATCHING;

GRADIENT OPERATORS

- THE GRADIENT OF AN IMAGE $f(x,y)$ AT LOCATION (x,y) IS DEFINED AS THE 2D VECTOR

$$G[f(x,y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

- THE VECTOR G POINTS IN THE DIRECTION OF MAXIMUM RATE OF CHANGE OF f AT LOCATION (x,y).

- FOR EDGE DETECTION WE ARE USUALLY INTERESTED IN THE MAGNITUDE OF THIS VECTOR,

$$|G| = \sqrt{G_x^2 + G_y^2}$$

- IT IS COMMON PRACTICE TO APPROXIMATE THE GRADIENT MAGNITUDE BY

$$|G| \simeq |G_x| + |G_y|$$

- TO CALCULATE THE PARTIAL DERIVATIVES, WE MAY CONSIDER TWO-POINT APPROXIMATIONS OF THE FORM

$$\frac{\partial f(x,y)}{\partial x} \simeq \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}$$

$$\frac{\partial f(x,y)}{\partial y} \simeq \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y}$$

- THESE OPERATORS CORRESPOND TO CORRELATION OF THE DISCRETIZED IMAGE FUNCTION WITH OPERATORS OF THE FORM

$$\begin{bmatrix} -1 & 1 \end{bmatrix} \quad \text{AND} \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

- TO EXAMINE THE USE OF THESE SIMPLE OPERATORS, LET US CONSIDER A 1-D FUNCTION (FOR SIMPLICITY).

- EXPANDING A 1D FUNCTION $f(x)$ ABOUT x IN A <u>TAYLOR</u> <u>SERIES</u> YIELDS

$$f(x+\Delta x) = f(x) + \Delta x\, f'(x) + \frac{(\Delta x)^2}{2!} f''(x) + \dots$$

— USING THE MEAN VALUE THEOREM OF CALCULUS WE MAY WRITE

$$f(x+\Delta x) - f(x) = \Delta x\, f'(x) + \frac{(\Delta x)^2}{2!} f''(\xi)$$

WHERE

$$\xi \in [x,\ x+\Delta x]$$

— DEFINING THE FORWARD APPROXIMATION D_1 AS

$$D_1(x) = \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

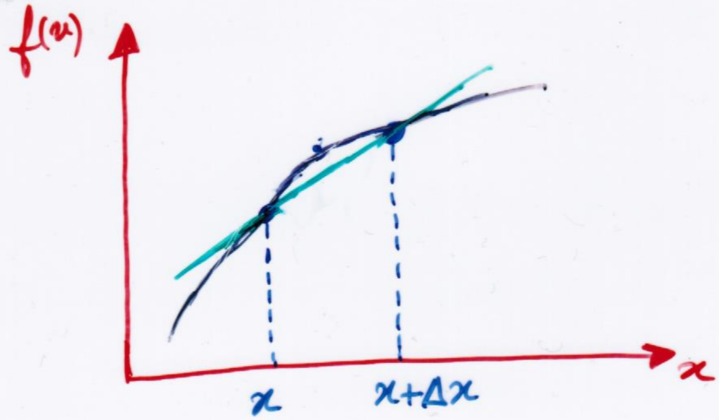
THEN

$$D_1(x) = f'(x) + \frac{\Delta x}{2} f''(\xi)$$

— THE ERROR IN THE DERIVATIVE APPROXIMATION USING D_1 AND A FINITE INTERVAL Δx, EVOLVES AS A FUNCTION OF Δx TO THE FIRST POWER,

$$\text{ERROR}(D_1) \simeq \mathcal{O}(\Delta x)$$

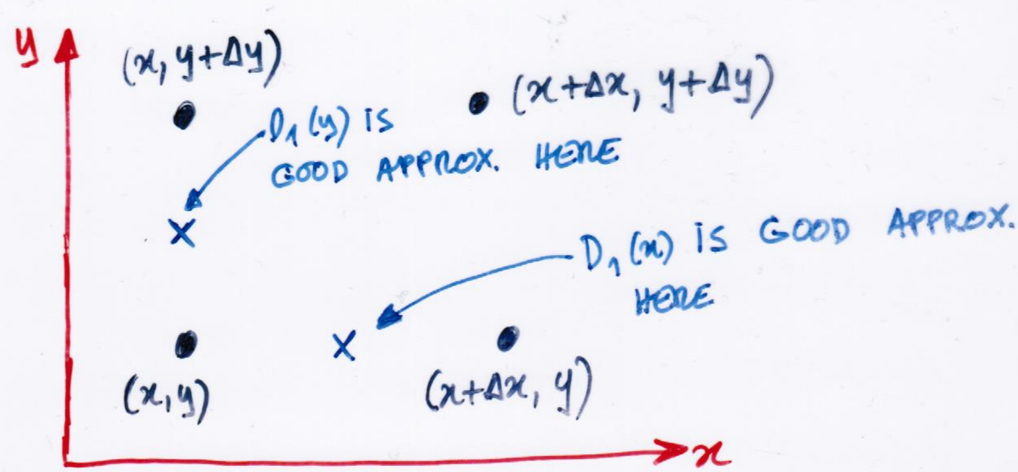
— A DECREASE IN $\underline{\Delta x}$ ONLY LINEARLY AFFECTS THE APPROXIMATION ERROR.



— D_1 IS A BETTER APPROXIMATION TO THE SLOPE OF $f(x)$ AT THE MIDPOINT OF THE INTERVAL $[x, x+\Delta x]$.

— THEREFORE:

— USING THE D_1 OPERATOR FOR $\partial f/\partial x$ AND $\partial f/\partial y$ APPROXIMATIONS RESULTS IN APPROXIMATIONS TO THE GRADIENT NOT AT (x, y) BUT AT DIFFERENT POINTS IN THE (x, y) PLANE.



— AN ALTERNATIVE CENTERED DIFFERENCE APPROXIMATION

$$D_2(x) = \frac{f(x+\Delta x) - f(x-\Delta x)}{2(\Delta x)}$$

HAS SEVERAL CONSEQUENCES:

1. $D_2(x)$ AND $D_2(y)$ DUE TO THEIR CENTERED NATURE, BOTH COMPRISE GOOD ESTIMATES FOR THE RESPECTIVE DERIVATIVES OF $f(x,y)$ AT THE MIDPOINTS OF THE INTERVAL — (x,y)

2. IT MAY BE SHOWN THAT THE APPROXIMATION ERROR EVOLVES AS

$$\text{ERROR}(D_2) \simeq \mathcal{O}((\Delta x)^2)$$

IN FACT:

$$f(x+\Delta x) = f(x) + \Delta x\, f'(x) + \frac{(\Delta x)^2}{2!} f''(x) + \\ + \frac{(\Delta x)^3}{3!} f'''(\xi)$$

AND

$$f(x-\Delta x) = f(x) - \Delta x\, f'(x) + \frac{(\Delta x)^2}{2!} f''(x) - \frac{(\Delta x)^3}{3!} f'''(\xi)$$

THEREFORE

$$f(x+\Delta x) - f(x-\Delta x) =$$

$$= 2\Delta x\, f'(x) + \frac{(\Delta x)^3}{3} f'''(\xi) \Rightarrow$$

$$D_2(x) = f'(x) + \frac{(\Delta x)^2}{6} f'''(\xi)$$

3. $D_2(x)$ AND $D_2(y)$ MAY BE IMPLEMENTED BY CONVOLVING THE IMAGE WITH MASKS

$$\begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \text{ AND } \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

– THIS CONCEPT FORMS THE BASIS FOR EXTENSION TO FAMILIES OF MASKS USED FOR EDGE DETECTION;

– VARIANTS

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \text{ AND } \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

→ THESE ARE "SMOOTHED" OR "AVERAGED" CENTERED DIFFERENCE OPERATORS AND ARE DEFINED AS $D_{2A}(x)$ AND $D_{2A}(y)$.

– SOBEL MASKS:

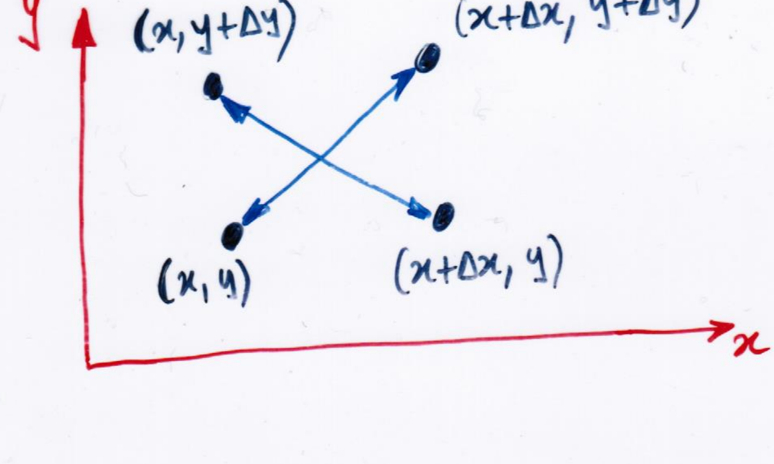
– SOBEL MASKS ARE SMOOTHED OPERATORS WITH A WEIGHTING THAT EMPHASIZES THE CENTRAL PIXEL:

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

– ANOTHER OPERATOR FOR GRADIENT APPROXIMATION THAT HAS BEEN EXTENSIVELY EMPLOYED IS THE ROBERTS OPERATOR AND IS DEFINED BY

$$D_+\{(x,y)\} = f(x+\Delta x, y+\Delta y) - f(x,y)$$

$$D_-\{(x,y)\} = f(x, y+\Delta y) - f(x+\Delta x, y)$$



THIS OPERATOR:

1. IS A VARIANT OF THE D_1 OPERATOR, WITH DERIVATIVES APPROXIMATED ALONG ORTHOGONAL ORIENTATIONS $45°$ AND $135°$ IN THE IMAGE PLANE;

2. THE MIDPOINT OF BOTH INTERVALS USED FOR APPROXIMATION IS THE <u>SAME POINT</u>; THAT IS, IT IS THE POINT $(x+\Delta x/2, y+\Delta y/2)$ LOCATED AT THE CENTER OF THE RECTANGLE DEFINED BY THE 4 POINTS USED.

3. D_+ AND D_- CORRESPOND TO THE CONVOLUTION OF THE IMAGE FUNCTION WITH MASKS:

$$D_+ = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \text{ AND } D_- = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

- SOMETIMES THIS IS CALLED THE ROBERTS "CROSS" OPERATOR;
- THE OUTPUT INTENSITY IN THE EDGE IMAGE IS FORMED VIA

$$f_{edge}(x,y) = \max\{|D_+|, |D_-|\}$$

LAPLACIAN AND SECOND DERIVATIVE OPERATORS

- EDGE DETECTION OPERATORS CAN ALSO BE BASED ON DISCRETE APPROXIMATIONS OF HIGHER ORDER DERIVATIVES;

- LET US CONSIDER THE 1D CASE;

- THE APPROXIMATION OF THE FIRST DERIVATIVE OF A FUNCTION WITH A NON-CENTERED DIFFERENCE FORMULA AND A "BACKWARD" FORMULATION YIELDS

$$D_{-1}(x) = \frac{f(x) - f(x-\Delta x)}{\Delta x}$$

- THE SECOND (FORWARD) DIFFERENCE APPROXIMATION OF THIS QUANTITY YIELDS

$$D_1(D_{-1}(x)) = \frac{1}{\Delta x}\left[\frac{f(x+\Delta x) - f(x)}{\Delta x} - \frac{f(x) - f(x-\Delta x)}{\Delta x}\right]$$

OR

$$D_1^2(x) = \frac{f(x+\Delta x) - 2f(x) + f(x-\Delta x)}{(\Delta x)^2}$$

- WHICH MAY BE REPRESENTED BY A WINDOW OF THE FORM

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

- THE LAPLACIAN OF AN IMAGE FUNCTION $f(x,y)$ IS DEFINED TO BE THE SUM OF THE SECOND SPATIAL DERIVATIVES

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}$$

- THE COMBINATION OF $D_1^2(x)$ AND $D_1^2(y)$ MAY THEN BE USED TO FORM THE DISCRETE APPROXIMATION OF THE LAPLACIAN

$$\nabla^2 f(x,y) = \frac{f(x+\Delta x,y) - 2f(x,y) + f(x-\Delta x,y)}{(\Delta x)^2} +$$

$$+ \frac{f(x,y+\Delta y) - 2f(x,y) + f(x,y-\Delta y)}{(\Delta y)^2}$$

- ASSUMING $\Delta x = \Delta y$ THIS QUANTITY MAY BE REWRITTEN AS:

$$\nabla^2 f(x,y) = \left[\frac{1}{\Delta x}\right]^2 [f(x+\Delta x, y) + f(x-\Delta x, y) + f(x, y+\Delta y) +$$

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

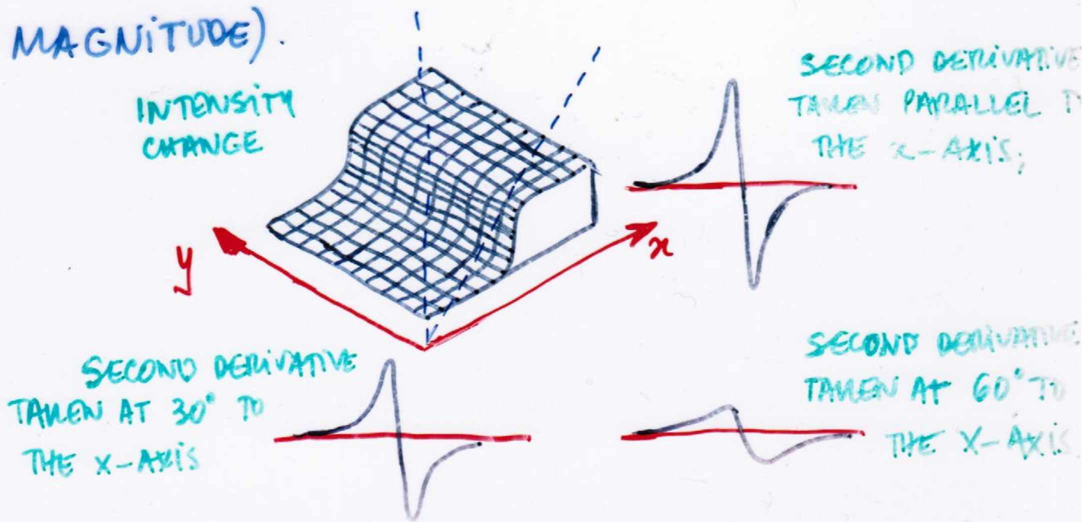
— AND REPRESENTED BY THE 3×3 WINDOW

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

— EDGE DETECTION BASED ON SECOND DERIVATIVES IS BASED ON THE FOLLOWING:

1. THE SECOND DERIVATIVE CROSSES ZERO AT THE SPATIAL LOCATION CORRESPONDING TO THE "MIDPOINT" OF THE EDGE;

2. THIS EFFECT OCCURS IN ALL DIRECTIONS (WITH VARYING MAGNITUDE).



- THE 2nd DERIVATIVE TAKEN ALONG A PARALLEL TO THE Y-AXIS IS IDENTICALLY ZERO;
- PROBLEM WITH THESE OPERATORS: THEIR SENSITIVITY TO NOISE;
- ONE POSSIBLE MINIMIZATION OF THIS PROBLEM: TO SMOOTH THE IMAGE PRIOR TO EDGE ENHANCEMENT;
- ONE SPECIFIC IMPLEMENTATION: THE LAPLACIAN-OF-THE-GAUSSIAN (LOG, $\nabla^2 G$) OPERATOR;

THE LAPLACIAN OF THE GAUSSIAN OPERATOR

- THE LOG OPERATOR SMOOTHS THE IMAGE THROUGH CONVOLUTION WITH A GAUSSIAN-SHAPED KERNEL TO MINIMIZE NOISE;
- FOLLOWING GAUSSIAN SMOOTHING THE LAPLACIAN OPERATOR IS APPLIED;
- THIS OPERATOR WAS FIRST PROPOSED BY MARR & HILDRETH BASED ON NEUROPHYSIOLOGICAL EVIDENCE;

- FOR SEVERAL REASONS THEY ARGUED THAT PRIOR TO THE APPLICATION OF THE LAPLACIAN, THE IMAGE SHOULD BE FILTERED BY A FILTER THAT SHOULD SATISFY TWO CONTRADICTORY REQUIREMENTS:

  1. THE FILTER SHOULD BE A LOW-PASS ONE (FREQUENCY REQUIREMENT);

  2. IT SHOULD ENABLE THE CREATION OF A SEQUENCE OF IMAGES AT VARIOUS SPATIAL RESOLUTIONS OR SCALES; THEREFORE IT SHOULD BE SPATIALLY LOCALIZED, i.e.,
     THE CONTRIBUTIONS FOR EACH POINT IN THE OUTPUT IMAGE SHOULD COME FROM AN AVERAGE OF THE NEIGHBOR PIXELS;

- THE EDGES SHOULD BE EXTRACTED IN ALL THE IMAGES (THEREFORE AT VARIOUS LEVELS OF RESOLUTION);
  - → AN EDGE AT A LOWER RESOLUTION SHOULD REMAIN AS THE RESOLUTION IS INCREASED;
  - → AS WE GO FROM LOW RESOLUTION TO HIGHER RESOLUTIONS NEW ZERO CROSSINGS MAY APPEAR, BUT EXISTING ONES NEVER DISAPPEAR;

- TO SATISFY THE ABOVE MENTIONED REQUIREMENTS WE KNOW FROM THE FOURIER TRANSFORM THEORY (BRACEWELL...) THAT:

$$\Delta x . \Delta \omega \geq \frac{1}{4}\pi$$

(EQUIVALENT TO THE UNCERTAINTY PRINCIPLE OF QUANTUM MECHANICS!)

- QUESTION: WHICH FUNCTION MINIMIZES THIS RELATION-SHIP?
  - GABOR FUNCTIONS (SINUSOIDAL FUNCTIONS MODELLED BY GAUSSIANS);

    → PROBLEM: THEY ARE NOT DIRECTIONNALY ISOTROPIC?

- NEW QUESTION:
  - WHICH FUNCTION MINIMIZES THIS RELATIONSHIP BEING OMNIDIRECTIONAL (WITH SPATIAL ISOTROPY)?

  - ANSWER: THE GAUSSIAN FUNCTION AS THE F.T. OF A GAUSSIAN IS STILL A GAUSSIAN!

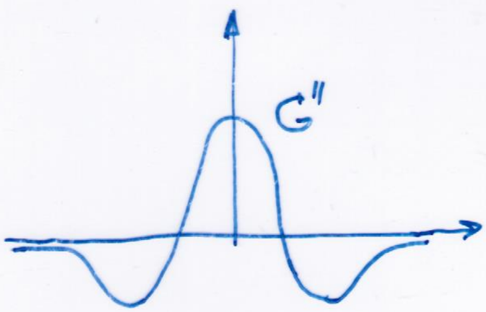
$$G(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\pi\sigma^2}\right);$$

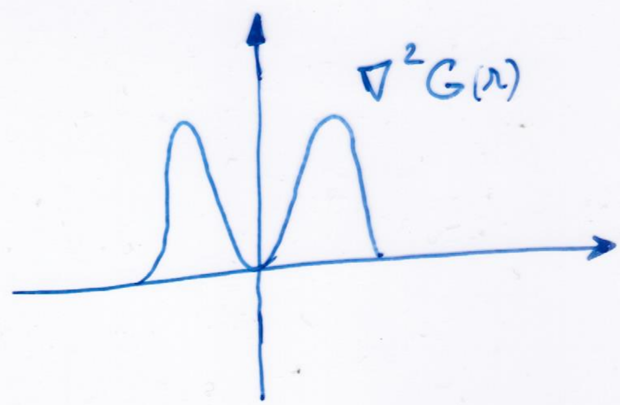
– IN 1-D:

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

WHOSE F.T. IS:

$$\tilde{G}(\omega) = \exp\left(-\frac{1}{2}\sigma^2\omega^2\right)$$





– IT CAN BE SHOWN THAT:

$$G''(x) = -\frac{1}{\sigma^3(2\pi)^{1/2}}\left(1-\frac{x^2}{\sigma^2}\right)\exp\left(-\frac{x^2}{2\sigma^2}\right)$$

AND:

$$\nabla^2 G(r) = -\frac{1}{\pi\sigma^4}\left[1-\frac{r^2}{2\sigma^2}\right]\exp\left(-\frac{r^2}{2\sigma^2}\right)$$

$$r^2 = x^2+y^2$$

– SO :

1° SMOOTH THE IMAGE BY USING A GAUSSIAN

$$G(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\pi\sigma^2}\right)$$

$$r^2 = x^2+y^2$$

2° APPLY THE LAPLACIAN TO THE FILTERED IMAGE:

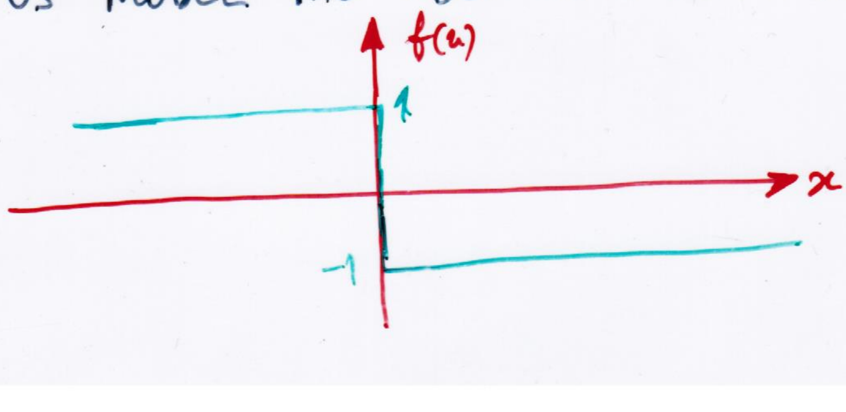
$$\nabla^2\left(G(x,y) * f(x,y)\right) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \iint_{-\infty}^{+\infty} G(x-\xi, y-\eta)\, f(\xi,\eta)\, d\xi\, d\eta =$$

$$= \underbrace{\left\{\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) G(x,y)\right\}}_{\text{"MEXICAN-HAT"} \equiv \nabla^2 G} * f(x,y)$$

EXAMPLE:

– LET US CONSIDER A 1D EXAMPLE OF THIS APPROACH:

– LET US MODEL THE EDGE AS A NEGATIVE STEP



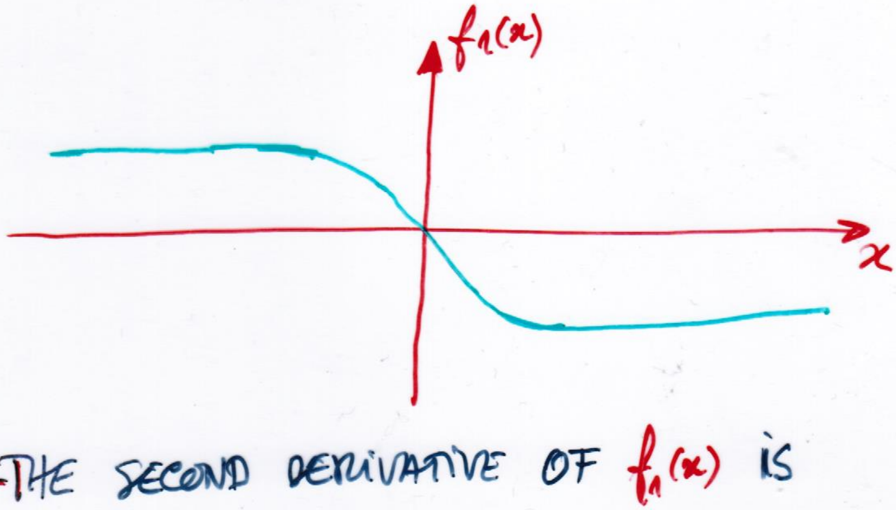
$$f(x) = -2u(x) + 1$$

– LET US NOW CONVOLVE THIS FUNCTION WITH A 1D GAUSSIAN SMOOTHING FUNCTION

$$g(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp(-x^2/2\sigma^2)$$

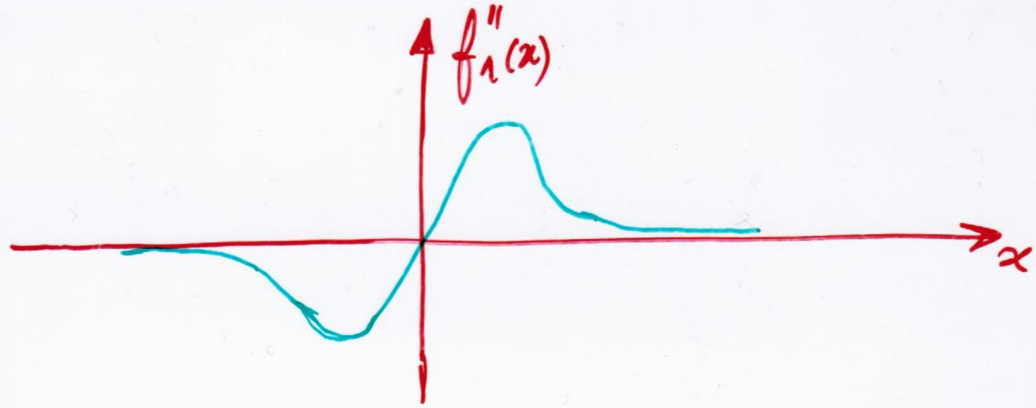
– AS A RESULT WE WILL HAVE A 1D BLURRED IMAGE $f_1(x)$

$$f_1(x) = 1 - 2\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} \exp(-\tau^2/2\sigma^2)\, d\tau$$



– THE SECOND DERIVATIVE OF $f_1(x)$ IS

$$f_1''(x) = \frac{2x}{\sigma^3\sqrt{2\pi}} \exp(-x^2/2\sigma^2)$$



— IN THE 2D CASE $G(x,y)$ IS CIRCULARLY SYMMETRIC AND THE SMOOTHING EFFECT MAY BE CONTROLLED THROUGH σ.

— IT IS IMPORTANT TO NOTE THAT $\underline{G(x,y)}$ MAY NOT BE CAST AS FIXED SIZE WINDOW; THE SPATIAL EXTENT OF THE OPERATOR VARIES WITH $\underline{\sigma}$.

— IN TERMS OF THE IMPLEMENTATION THE DISCRETE CONVOLUTION IS TRUNCATED WHEN THE VALUES OF THE SMOOTHING KERNEL BECOME NUMERICALLY INSIGNIFICANT.

— ALTHOUGH THE SPATIAL EXTENT OF THE $\underline{LOG}$ $(\nabla^2 G(x))$ IS THEORETICALLY INFINITE, THE NUMERICALLY SIGNIFICANT EXTENT OF THE OPERATOR IS BOUNDED.

NOTE: MORE THAN 99% OF THE AREA UNDER THE GAUSSIAN CURVE LIES WITHIN A DISTANCE OF $\pm 3\sigma$ FROM THE ORIGIN.

— FROM

$$\nabla^2 G(r) = \frac{1}{\pi\sigma^4}\left(\frac{r^2}{2\sigma^2} - 1\right) \exp\left(-\frac{r^2}{2\sigma^2}\right)$$

$$r^2 = x^2 + y^2$$

IT CAN BE SHOWN THAT THE NEGATIVE PORTION OF THE LOG OPERATOR IS WITHIN A DISTANCE OF $\underline{\sigma\sqrt{2}}$ FROM THE ORIGIN; THUS, THE DISTANCE BETWEEN ZERO CROSSINGS OF THE LOG OPERATOR IS $\Delta = 2\sqrt{2}\,\sigma \simeq 3\sigma$. THEREFORE FOR EDGE DETECTION ONE SHOULD USE MASK DIMENSIONS IN THE RANGE OF 3Δ TO 4Δ.

MASK SIZE OF THE LOG OPERATOR AS A FUNCTION OF σ

σ	$6\sqrt{2}\sigma$	MASK SIZE
0.5	4.24	5x5
1	8.48	9x9
2	17.0	17x17
3	25.4	26x26
4	33.9	34x34
5	42.4	43x43

— IT CAN BE PROVED THAT THE LOG OPERATOR CAN BE APPROXIMATED BY THE DIFFERENCE OF TWO IMAGES, EACH FILTERED WITH THE GAUSSIAN OPERATOR BUT WITH DIFFERENT σ VALUES;

ADVANTAGES OF THE ZERO CROSSING METHOD:

1. SIMPLICITY: THE CONTOUR POINT EXTRACTION USES RELATIVELY SIMPLE OPERATIONS;

2. THIN CONTOURS: SINCE ZERO CROSSING POINTS ARE DEFINED AS THE TRANSITIONS BETWEEN POSITIVE AND NEGATIVE REGIONS, THEY FORM ONE POINT WIDE CONNECTED CHAINS IN THE DISCRETE IMAGE:

DISADVANTAGE:

— THE ISOTROPY OF CONTOUR POINT DETECTION CAN BECOME A DISADVANTAGE;

— THE LAPLACIAN COMBINES SECOND DERIVATIVES IN x AND y; RESULTS ARE GOOD IN REGIONS WITH A SINGLE CONTOUR WITH A DEFINITE ORIENTATION, BUT UNSATISFACTORY IN REGIONS OF INTERSECTION OF MULTIPLE CONTOURS WITH DIFFERENT ORIENTATIONS;

- CORNERS ARE DEFORMED AND INTERSECTIONS OF 3 OR MORE CONTOURS ARE DISCONNECTED.

- OTHER EDGE DETECTORS:

  - CANNY'S; DEFINED USING OPTIMALITY CRITERIA FOR THE DETECTION OF EDGES; (F.I.R. IMPLEMENTATION);

  - DERICHE'S; RETAINED CANNY'S OPTIMALITY CRITERIA, BUT EXTENDED THE RESULT TO THE RECURSIVE CASE (I.I.R. IMPLEMENTATION);

EDGE DETECTION VIA TEMPLATE MATCHING

- ONE TRIES TO MATCH THE IMAGE AGAINST A PREDEFINED MODEL OF AN EDGE;

  - EXAMPLE: HUECKEL OPERATOR;

OR

- ONE MATCHES THE IMAGE AGAINST A SET OF TEMPLATES REPRESENTING POSSIBLE EDGES;

  - EXAMPLE: FREI-CHEN TEMPLATES;

FREI-CHEN TEMPLATES:

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

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

$$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \quad \begin{bmatrix} -2 & 1 & -2 \\ 1 & 4 & 1 \\ -2 & 1 & -2 \end{bmatrix}$$

– SEGMENTATION

– SEGMENTATION IS THE PROCESS THAT SUBDIVIDES A SCENE INTO ITS CONSTITUENT PARTS OR OBJECTS.

– AT THIS STAGE OF PROCESSING OBJECTS ARE EXTRACTED FROM A SCENE FOR SUBSEQUENT RECOGNITION AND ANALYSIS.

SEGMENTATION ALGORITHMS ARE GENERALLY BASED ON ONE OF TWO BASIC PRINCIPLES:

- DISCONTINUITY;
  AND
- SIMILARITY;

- THE PRINCIPAL APPROACH IN THE FIRST CATEGORY IS BASED ON EDGE DETECTION;

- THE PRINCIPAL APPROACHES IN THE SECOND CATEGORY ARE BASED ON THRESHOLDING AND REGION GROWING;

- EDGE LINKING AND BOUNDARY DETECTION

- IDEALLY EDGE DETECTION TECHNIQUES SHOULD YIELD ONLY PIXELS LYING ON THE BOUNDARY BETWEEN OBJECTS AND THE BACKGROUND.

- IN PRACTICE: THIS SET OF PIXELS SELDOM CHARACTERIZES A BOUNDARY COMPLETELY BECAUSE OF NOISE, BREAKS IN THE BOUNDARY DUE TO NON-UNIFORM ILLUMINATION, AND OTHER EFFECTS THAT INTRODUCE SPURIOUS INTENSITY DISCONTINUITIES.

SEGMENTATION ALGORITHMS ARE GENERALLY BASED ON ONE OF TWO BASIC PRINCIPLES:

- DISCONTINUITY;
  AND
- SIMILARITY;

- THE PRINCIPAL APPROACH IN THE FIRST CATEGORY IS BASED ON EDGE DETECTION;

- THE PRINCIPAL APPROACHES IN THE SECOND CATEGORY ARE BASED ON THRESHOLDING AND REGION GROWING;

- EDGE LINKING AND BOUNDARY DETECTION

- IDEALLY EDGE DETECTION TECHNIQUES SHOULD YIELD ONLY PIXELS LYING ON THE BOUNDARY BETWEEN OBJECTS AND THE BACKGROUND.

- IN PRACTICE: THIS SET OF PIXELS SELDOM CHARACTERIZES A BOUNDARY COMPLETELY BECAUSE OF NOISE, BREAKS IN THE BOUNDARY DUE TO NON-UNIFORM ILLUMINATION, AND OTHER EFFECTS THAT INTRODUCE SPURIOUS INTENSITY DISCONTINUITIES.

– THEREFORE:

– EDGE DETECTION ALGORITHMS ARE TYPICALLY FOLLOWED BY LINKING AND OTHER BOUNDARY DETECTION PROCEDURES DESIGNED TO ASSEMBLE EDGE PIXELS INTO A MEANINGFUL SET OF OBJECT BOUNDARIES.

– SOME EXAMPLES OF TECHNIQUES SUITED FOR THIS PURPOSE:

– LOCAL ANALYSIS:

– IN AN IMAGE THAT HAS UNDERGONE AN EDGE DETECTION PROCESS THE CHARACTERISTICS OF PIXELS IN A SMALL NEIGHBORHOOD (3x3 OR 5x5) ARE ANALYZED.

ABOUT EVERY POINT (x, y);

– ALL PIXELS THAT ARE SIMILAR ARE LINKED, THUS FORMING A BOUNDARY OF PIXELS THAT SHARE SOME COMMON PROPERTIES.

– THERE ARE TWO PRINCIPAL PROPERTIES USED FOR ESTABLISHING SIMILARITY OF EDGE PIXELS:

1. THE STRENGTH OF THE RESPONSE OF THE GRADIENT OPERATOR USED TO PRODUCE THE EDGE PIXEL;

IF $G[f(x,y)]$ IS THE MAGNITUDE OF THE GRADIENT AT (x,y) THEN WE SAY THAT A PIXEL WITH COORDINATES (x', y') IN THE PREDEFINED NEIGHBORHOOD OF (x,y) IS SIMILAR IN MAGNITUDE TO THE PIXEL AT (x,y) IF

$$|G[f(x,y)] - G[f(x', y')]| \leq T$$

T - THRESHOLD

2. THE DIRECTION OF THE GRADIENT;

- THE DIRECTION OF THE GRADIENT MAY BE ESTABLISHED FROM THE ANGLE OF THE GRADIENT VECTOR, THAT IS

$$\theta = \tan^{-1}\left[\frac{G_y}{G_x}\right]$$

WHERE θ IS THE ANGLE ALONG WHICH THE RATE OF CHANGE HAS THE GREATEST MAGNITUDE.

THEN WE SAY THAT AN EDGE PIXEL AT (x', y') IN THE PREDEFINED NEIGHBORHOOD OF (x,y) HAS AN ANGLE SIMILAR TO THE PIXEL AT (x,y) IF

$$|\theta - \theta'| < A$$

— IN ORDER TO DISTINGUISH THE LINKED EDGE PIXELS FROM THE OTHER PIXELS ONE MAY ASSIGN A DIFFERENT GRAY LEVEL TO EACH ~~SET~~ SET OF LINKED EDGE PIXELS.

— GLOBAL ANALYSIS VIA HOUGH TRANSFORM

— WE ARE GOING TO CONSIDER THE LINKING OF BOUNDARY POINTS BY DETERMINING WHETHER OR NOT THEY LIE ON A CURVE OF SPECIFIED SHAPE.

— LET US SUPPOSE THAT, GIVEN n POINTS IN THE xy PLANE OF AN IMAGE, WE WISH TO FIND SUBSETS THAT LIE ON STRAIGHT LINES.

— ONE POSSIBLE SOLUTION:

— FIRST FIND ALL LINES DETERMINED BY EVERY PAIR OF POINTS;

— THEN FIND ALL SUBSETS OF POINTS THAT ARE CLOSE TO PARTICULAR LINES;

— PROBLEM: WE HAVE ~~TO~~ FIND $\binom{n}{2}$ LINES

$$\binom{n}{2} = \frac{n!}{(n-2)!\,2!} = \frac{n(n-1)}{2} \simeq \underline{\underline{n^2}}$$

AND THEN PERFORMING

$$n \frac{n(n-1)}{2} \simeq n^3$$

COMPARISONS OF EVERY POINT TO ALL LINES.

- THIS IS COMPUTATIONALLY PROHIBITIVE;
- SOLUTION: THE SO-CALLED HOUGH TRANSFORM.
- LET US CONSIDER A POINT WITH COORDINATES (x_i, y_i).
- GENERAL EQUATION OF A STRAIGHT LINE THAT PASSES THROUGH (x_i, y_i)

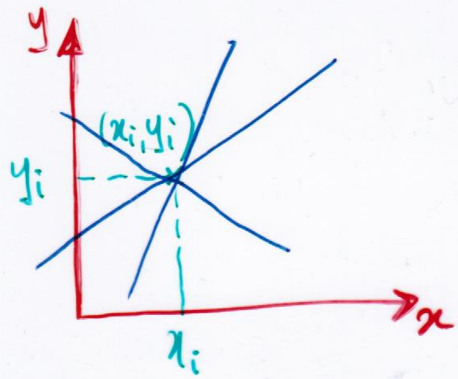
$$y_i = a x_i + b$$

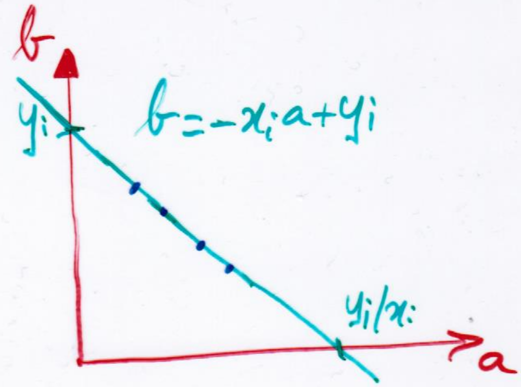
- FOR VARYING VALUES OF a AND b THERE IS AN INFINITE NUMBER OF LINES THAT PASS THROUGH (x_i, y_i).
- IF WE WRITE THIS EQUATION AS

$$b = -x_i a + y_i$$

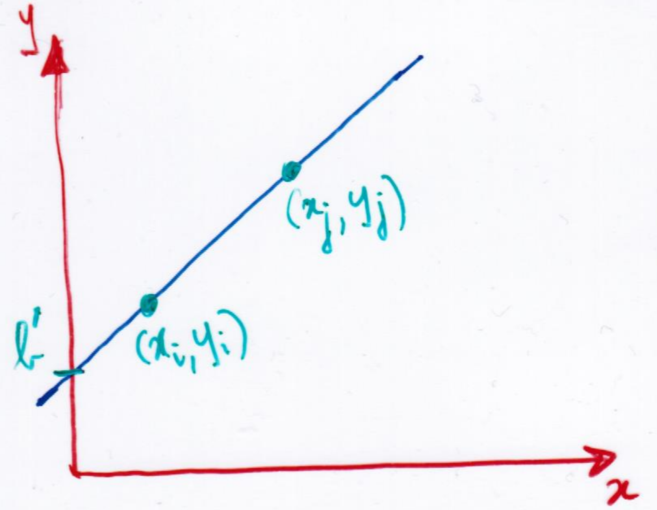
- AND CONSIDER THE ab PLANE (ALSO CALLED PARAMETER SPACE)

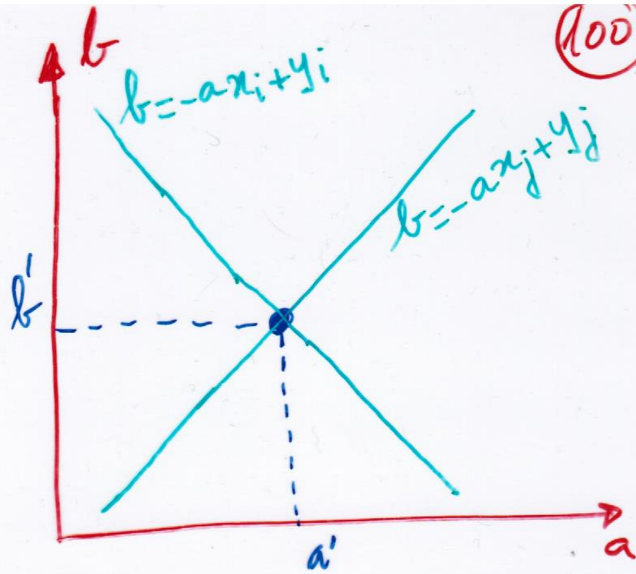
- THEN WE HAVE THE EQUATION OF A SINGLE LINE FOR A FIXED PAIR (x_i, y_i)



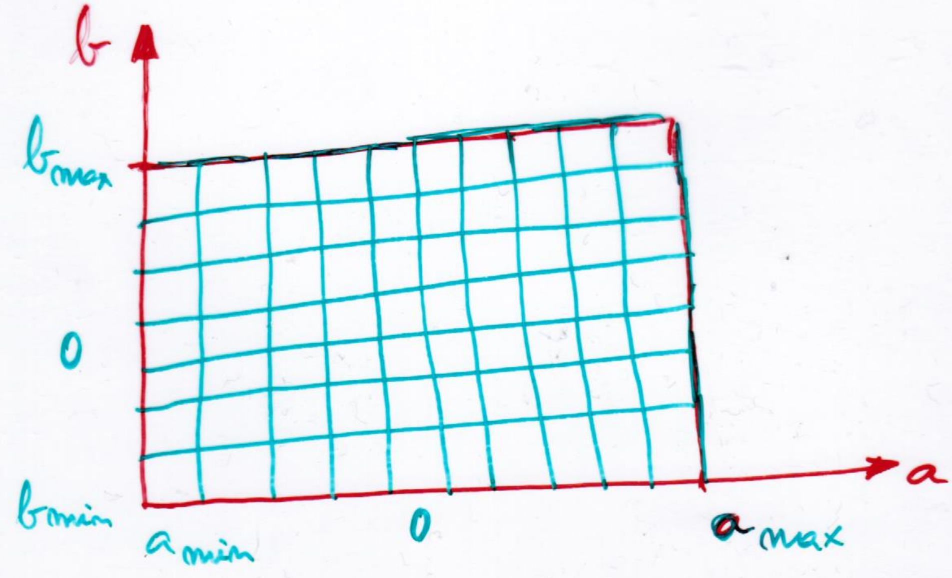


- IF WE CONSIDER A SECOND POINT (x_j, y_j), ANOTHER LINE IN PARAMETER SPACE WILL BE ASSOCIATED WITH IT.
- THIS LINE WILL INTERSECT THE LINE ASSOCIATED WITH (x_i, y_i) AT (a', b') WHERE a' IS THE SLOPE AND b' THE INTERCEPT OF THE LINE CONTAINING BOTH (x_i, y_i) AND (x_j, y_j) IN THE xy PLANE.
- ALL POINTS CONTAINED ON THIS LINE (DEFINED BY (x_i, y_i) AND (x_j, y_i)) WILL HAVE LINES IN PARAMETER SPACE WHICH INTERCEPT AT (a', b')





- THE COMPUTATIONAL INTEREST OF THE HOUGH TRANSFORM ARISES FROM SUBDIVIDING THE PARAMETER SPACE INTO SO-CALLED ACCUMULATOR CELLS.



- WHERE (a_{max}, a_{min}) AND (b_{max}, b_{min}) ARE THE EXPECTED RANGES OF SLOPE AND INTERCEPT VALUES.

- ACCUMULATOR CELL $A(i,j)$ CORRESPONDS TO THE SQUARE ASSOCIATED WITH PARAMETER SPACE COORDINATES (a_i, b_j).

- INITIALLY THESE CELLS ARE SET TO ZERO.

- THEN, FOR EVERY PIXEL (x_k, y_k) IN THE IMAGE PLANE, WE LET THE PARAMETER a EQUAL EACH OF THE ALLOWED SUBDIVISION VALUES ON THE a AXIS AND SOLVE FOR THE CORRESPONDING b USING EQUATION

$$b = -a x_k + y_k$$

- THE RESULTING b's ARE THEN ROUNDED OFF TO THE NEAREST ALLOWED VALUE IN THE b AXIS.

- IF A CHOICE OF a_p RESULTS IN SOLUTION b_p WE LET

$$A(p,q) = A(p,q) + 1$$

- AT THE END OF THIS PROCEDURE, A VALUE OF M IN CELL $A(i,j)$ CORRESPONDS TO M POINTS IN THE xy PLANE LYING ON THE LINE

$$y = a x_i + b_j$$

- THE ACCURACY OF THE COLINEARITY OF THESE POINTS IS ESTABLISHED BY THE NUMBER OF SUB-DIVISIONS IN THE ab PLANE.

- NOTE THAT IF WE SUBDIVIDE THE a AXIS INTO K INCREMENTS, THEN FOR EVERY POINT (x_k, y_k) WE OBTAIN K VALUES OF b CORRESPONDING TO THE K POSSIBLE VALUES OF a.

- SINCE THERE ARE n IMAGE POINTS, THIS INVOLVES nK COMPUTATIONS.

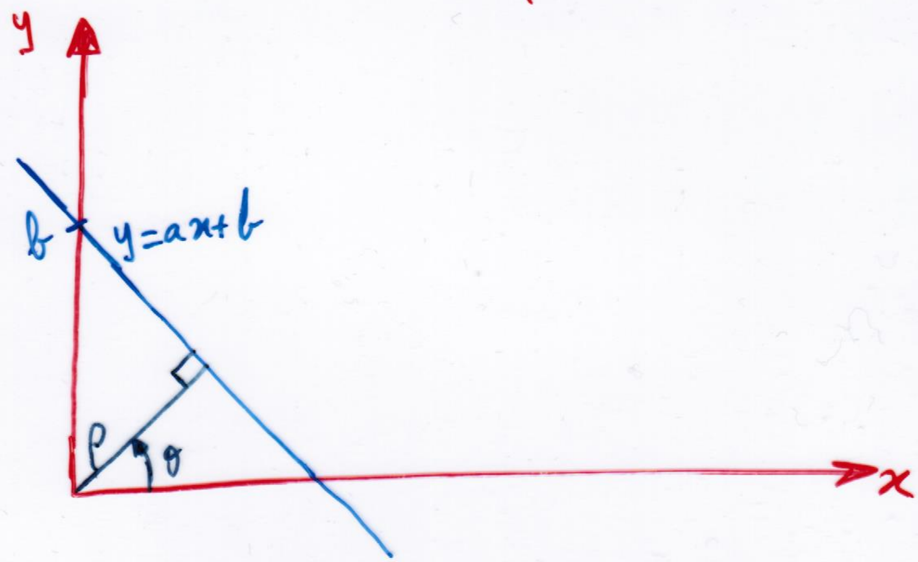
- THIS PROCEDURE IS THEREFORE LINEAR IN n, AND THE PRODUCT nK DOES NOT APPROACH n^2 UNLESS K APPROACHES OR EXCEEDS n.

- PROBLEM WITH USING EQUATION $y = ax + b$ TO REPRESENT A LINE:

  → BOTH THE SLOPE AND INTERCEPT APPROACH INFINITY AS THE LINE APPROACHES A VERTICAL POSITION;

- TO GET AROUND THIS DIFFICULTY:
  - USE THE SO-CALLED "NORMAL REPRESENTATION" OF A LINE GIVEN BY

$$x \cos\theta + y \sin\theta = \rho$$



- FROM THE FIGURE:

$$y = ax + b \Rightarrow tg\,\theta = \frac{\sin\theta}{\cos\theta} = -\frac{1}{a} \Leftrightarrow a = -\frac{\cos\theta}{\sin\theta}$$

$$b \cos(90-\theta) = \rho \Leftrightarrow b \sin\theta = \rho \Leftrightarrow b = \frac{\rho}{\sin\theta}$$

SUBSTITUTING IN $\underline{y = ax + b}$:

$$y = -\frac{\cos\theta}{\sin\theta} x + \frac{\rho}{\sin\theta} \Leftrightarrow \underline{x \cos\theta + y \sin\theta = \rho}$$

- THE USE OF THIS REPRESENTATION IN CONSTRUCTING A TABLE OF ACCUMULATORS IS IDENTICAL TO THE METHOD JUST DISUSSED;

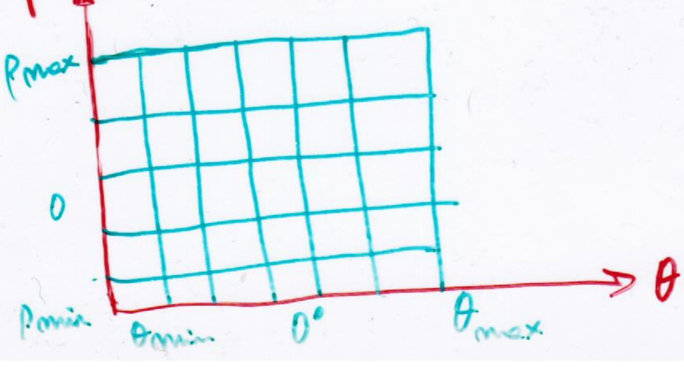
- THE ONLY DIFFERENCE IS THAT, INSTEAD OF STRAIGHT LINES, WE NOW HAVE SINUSOIDAL CURVES IN THE $\theta\rho$ PLANE.

- AS BEFORE M COLINEAR POINTS LYING ON A LINE

$$x \cos\theta_i + y \sin\theta_i = \rho_j$$

WILL YIELD M SINUSOIDAL CURVES WHICH INTERCEPT AT (θ_i, ρ_j) IN THE PARAMETER SPACE.

- IF WE USE THE METHOD OF INCREMENTING θ AND SOLVING FOR THE CORRESPONDING ρ, THE PROCEDURE WILL YIELD <u>M</u> ENTRIES IN ACCUMULATOR $A(i,j)$ ASSOCIATED WITH THE CELL DETERMINED BY (θ_i, ρ_j).



- THE HOUGH TRANSFORM IS APPLICABLE TO ANY FUNCTION OF THE FORM

$$g(\bar{x}, \bar{c}) = 0$$

WHERE $\bar{x}$ IS A VECTOR OF COORDINATES AND $\bar{c}$ IS A VECTOR OF COEFFICIENTS.

- THE LOCUS OF POINTS LYING ON THE CIRCLE

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2$$

CAN BE DETECTED BY USING THIS APPROACH.

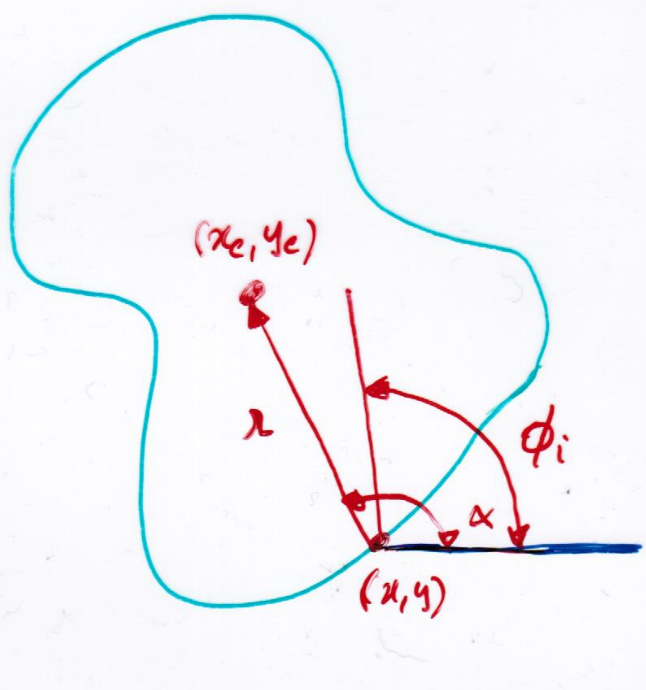
- HOWEVER IN THIS CASE THERE ARE 3 PARAMETERS $\underline{c_1, c_2, \text{ AND } c_3}$ WHICH RESULT IN A 3D PARAMETER SPACE WITH CUBELIKE CELLS AND ACCUMULATORS OF THE FORM $A(i, j, k)$.

- THE PROCEDURE IS TO INCREMENT $\underline{c_1}$ AND $\underline{c_2}$ AND SOLVE FOR $\underline{c_3}$, AND THEN UPDATE THE ACCUMULATOR CORRESPONDING TO THE CELL ASSOCIATED WITH THE TRIPLE (c_1, c_2, c_3).

- THE COMPLEXITY OF THE HOUGH TRANSFORM IS STRONGLY DEPENDENT ON THE NUMBER OF COORDINATES AND COEFFICIENTS IN A GIVEN REPRESENTATION.

- THE HOUGH TRANSFORM MAY BE GENERALIZED TO DETECT CURVES WITH NO ANALYTIC REPRESENTATION.

- SUPPOSE THAT THE OBJECT APPEARS IN THE IMAGE WITH KNOWN SHAPE, ORIENTATION AND SCALE.
  - IF ORIENTATION AND SCALE ARE UNKNOWN, THEY CAN BE HANDLED IN THE SAME WAY THAT ANY ADDITIONAL PARAMETERS;



- PICK A REFERENCE POINT IN THE SILHOUETTE AND DRAW A LINE TO THE BOUNDARY.

- AT THE BOUNDARY POINT COMPUTE THE GRADIENT DIRECTION AND STORE THE REFERENCE POINT AS A FUNCTION OF THIS DIRECTION.

– THUS IT IS POSSIBLE TO PRECOMPUTE THE LOCATION OF THE REFERENCE POINT FROM BOUNDARY POINTS GIVEN THE GRADIENT ANGLE.

– THE SET OF ALL SUCH LOCATIONS, INDEXED BY GRADIENT ANGLE, COMPRISES A TABLE TERMED THE "R-TABLE".

– THE BASIC STRATEGY OF THE HOUGH TRANSFORM IS TO COMPUTE THE POSSIBLE LOCI OF REFERENCE POINTS IN PARAMETER SPACE FROM EDGE POINT DATA IN IMAGE SPACE AND INCREMENT THE PARAMETER POINTS IN AN ACCUMULATOR ARRAY.

"THE R-TABLE"

ANGLE MEASURED FROM FIGURE BOUNDARY TO REFERENCE POINT	SET OF RADII $\{\bar{r}^k\}$ WHERE $\bar{r} = (r, \alpha)$
ϕ_1	$\bar{r}^1$
ϕ_2	$\bar{r}^2$
$\vdots$	$\vdots$
ϕ_m	$\bar{r}^m$

- THE REFERENCE POINT COORDINATES (x_c, y_c) ARE THE ONLY PARAMETERS (ASSUMING THAT ROTATION AND SCALING HAVE BEEN FIXED).

- THUS AN EDGE POINT (x, y) WITH GRADIENT ORIENTATION ϕ CONSTRAINS THE POSSIBLE REFERENCE POINTS TO BE AT

$$\{x + r_1(\phi)\cos[\alpha_1(\phi)],\ y + r_1(\phi)\sin[\alpha_1(\phi)]\}$$

ALGORITHM (GENERALIZED HOUGH)

STEP 1 – MAKE A TABLE FOR THE SHAPE TO BE LOCATED.

STEP 2 – FORM AN ACCUMULATOR ARRAY OF POSSIBLE REFERENCE POINTS $A(x_{c\,min} : x_{c\,max}, y_{c\,min} : y_{c\,max})$ INITIALIZED TO ZERO.

STEP 3 – FOR EACH EDGE POINT DO THE FOLLOWING:

3.1 – COMPUTE $\phi(x, y)$;

3.2 – CALCULATE THE POSSIBLE CENTERS, THAT IS FOR EACH TABLE ENTRY FOR ϕ, COMPUTE

$$x_e := x + r(\phi) \cos[\alpha(\phi)]$$
$$y_e := y + r(\phi) \sin[\alpha(\phi)]$$

3.3 - INCREMENT THE ACCUMULATOR ARRAY

$$A(x_e, y_e) := A(x_e, y_e) + 1$$

STEP 4. - POSSIBLE LOCATIONS FOR THE SHAPE ARE GIVEN BY MAXIMA IN ARRAY A.

- IF WE WERE TO TAKE INTO ACCOUNT THE PARAMETERS OF SCALE AND ROTATION S AND θ, WE SHOULD HAVE TO EXPAND THE ACCUMULATOR ARRAY AND DO MORE WORK IN THE INCREMENTATION STEP.

  - IN STEP 2 THE ACCUMULATOR ARRAY IS CHANGED TO

$$(x_{e\,min} : x_{e\,max},\ y_{e\,min} : y_{e\,max},\ S_{min} : S_{max},\ \theta_{min} : \theta_{max})$$

  - AND STEP 3.2 IS CHANGED TO

FOR EACH TABLE ENTRY FOR $\underline{\phi}$ DO

FOR EACH $\underline{S}$ AND $\underline{\theta}$

$$\begin{cases} x_e := x + r(\phi) S \cos[\alpha(\phi) + \theta] \\ y_e := y + r(\phi) S \sin[\alpha(\phi) + \theta] \end{cases}$$

AND STEP 3.3 IS NOW

$$A(x_c, y_c, S, \theta) := A(x_c, y_c, S, \theta) + 1$$

REGION-ORIENTED SEGMENTATION

– THE GOAL OF SEGMENTATION IS TO PARTITION AN IMAGE INTO REGIONS.

– LET R REPRESENT THE ENTIRE IMAGE REGION. WE MAY VIEW SEGMENTATION AS A PROCESS THAT PARTITIONS R INTO n SUBREGIONS $R_1, R_2, \dots, R_n$ SUCH THAT

1. $\bigcup_{i=1}^{n} R_i = R$
2. R_i IS A CONNECTED REGION, $i = 1, 2, \dots, n$
3. $R_i \cap R_j = \phi$ FOR ALL i AND j, $i \neq j$.
4. $P(R_i) = \text{TRUE}$ FOR $i = 1, 2, \dots, n$.
5. $P(R_i \cup R_j) = \text{FALSE}$ FOR $i \neq j$.

WHERE $P(R_i)$ IS A LOGICAL PREDICATE DEFINED OVER THE POINTS IN SET R_i.

- PODE VERIFICAR-SE A EQUIVALÊNCIA FAZENDO AS SUBSTITUIÇÕES:

$$(x_i^h, y_i^h, z_i^h, w_i^h) = (\alpha x_i, \alpha y_i, \alpha z_i, \alpha) \text{ E}$$

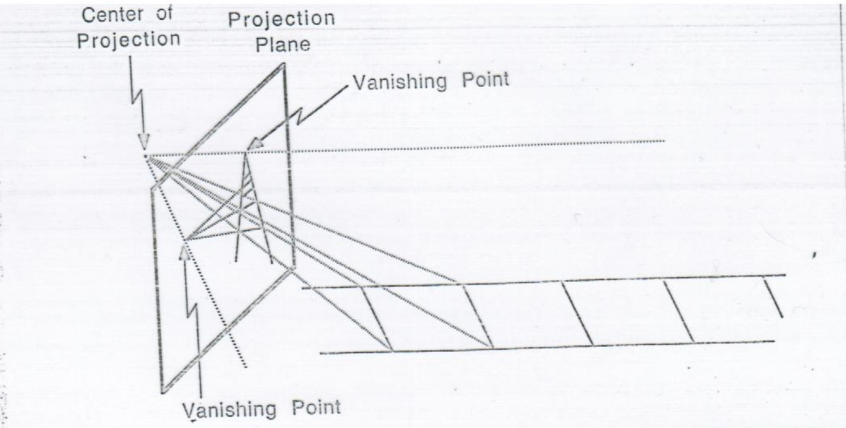
$$(x_0^h, y_0^h, z_0^h, w_0^h) = (\beta x_0, \beta y_0, \beta z_0, \beta)$$

- UMA TRANSFORMAÇÃO É LINEAR SE E SÓ SE PUDER SER EXPRESSA SOB A FORMA DE MULTIPLICAÇÃO DE MATRIZES;

- PONTO DE FUGA E LINHA DE FUGA

- PONTO DE FUGA DE UMA LINHA RECTA EM PROJECÇÃO EM PERSPECTIVA É O PONTO NA IMAGEM PARA ALÉM DO QUAL A PROJECÇÃO DA LINHA RECTA NÃO PODE SER PROLONGADA.

- OU SEJA: SE A LINHA RECTA FOSSE INFINITAMENTE COMPRIDA NO ESPAÇO A LINHA PARECERIA "DESAPARECER" NO SEU PONTO DE FUGA NA IMAGEM;



— O PONTO DE FUGA DE UMA LINHA RECTA DEPENDE APENAS DA ORIENTAÇÃO DA LINHA E NÃO DA SUA POSIÇÃO NO ESPAÇO;

— PORTANTO:

COM EXCEPÇÃO DO CASO EM QUE LINHAS PARALELAS SÃO PARALELAS AO PLANO DA IMAGEM, LINHAS PARALELAS NO ESPAÇO, PROJECTAM-SE, EM PERSPECTIVA, EM LINHAS RECTAS QUE, PROLONGADAS, SE INTERSECTAM NUM PONTO (PONTO DE FUGA);

— O PONTO DE FUGA CORRESPONDE A PONTOS NO INFINITO NA DIRECÇÃO DAS LINHAS RECTAS PARALELAS NO ESPAÇO;

— O PONTO DE FUGA É DADO PELA INTERSECÇÃO NA IMAGEM DE UMA LINHA RECTA PARALELA QUE PASSA PELO CENTRO DE PROJECÇÃO;

- O PONTO DE FUGA DE QUALQUER LINHA RECTA CONTIDA NUM PLANO (E DE QUALQUER LINHA RECTA PARALELA A ESSE MESMO PLANO) ESTÁ SOBRE UMA LINHA RECTA PARTICULAR NO PLANO DA IMAGEM;

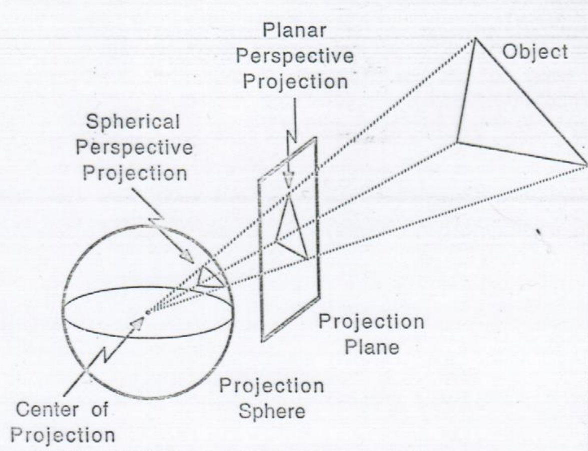
- TRATA-SE DA LINHA DE FUGA DO PLANO, LOCALIZADA ONDE UM PLANO PARALELO QUE PASSA PELO CENTRO DE PROJECÇÃO INTERSECTA O PLANO DA IMAGEM;

- PROJECÇÃO EM PERSPECTIVA PLANAR E ESFÉRICA

PLANAR: IMAGEM FORMADA NA INTERSECÇÃO DO CONE DE RAIOS DE PROJECÇÃO COM UM PLANO;

NA PROJECÇÃO EM PERSPECTIVA PLANA A IMAGEM DEPENDE NÃO APENAS DO CENTRO DE PROJECÇÃO MAS TAMBÉM DA ORIENTAÇÃO E POSIÇÃO DA SUPERFÍCIE ONDE SE FORMA A IMAGEM (UM PLANO);

- PARA SUPRIMIR TAL DEPENDÊNCIA:



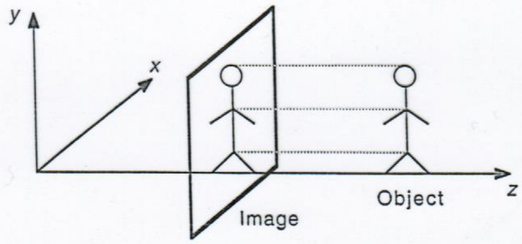
– PROJECTA-SE A CENA NUMA ESFERA DE RAIO UNITÁRIO CENTRADA NO CENTRO DE PROJECÇÃO;

PROJECÇÃO EM PERSPECTIVA ESFÉRICA:

PROJECÇÃO AO LONGO DE UM CONE DE RAIOS DE PROJECÇÃO NUMA SUPERFÍCIE ESFÉRICA, NÃO NECESSÀRIAMENTE DE RAIO UNITÁRIO;

– SOB <u>PERSPECTIVA ESFÉRICA</u> LINHAS RECTAS SÃO MAPEADAS EM ARCOS DE GRANDES CÍRCULOS – i.e., SÃO MAPEADAS EM ARCOS DE CÍRCULO CENTRADOS NO CENTRO DA ESFERA;

– UMA IMAGEM FORMADA SOB PERSPECTIVA PLANAR DEFINE A CORRESPONDENTE IMAGEM SOB PROJECÇÃO ESFÉRICA;



PROJECÇÃO ORTOGRÁFICA

– É A PROJECÇÃO DE UMA ENTIDADE TRI-DIMENSIONAL NUM PLANO POR UM CONJUNTO DE RAIOS PARALELOS PERPENDICULARES AO PLANO.

$$\begin{cases} X = x \\ Y = y \end{cases}$$

– A PROJECÇÃO ORTOGRÁFICA (SOB DETERMINADAS CONDIÇÕES) APROXIMA A PROJECÇÃO EM PERSPECTIVA, A MENOS DE UM FACTOR DE ESCALA UNIFORME.

– A PROJECÇÃO ORTOGRÁFICA É UMA TRANSFORMAÇÃO LINEAR NO ESPAÇO TRI-DIMENSIONAL.

– NA PROJECÇÃO EM PERSPECTIVA:

– À MEDIDA QUE O OBJECTO É AFASTADO DO CENTRO DE PROJECÇÃO, AO LONGO DO EIXO DOS ZZ, O TAMANHO DA IMAGEM DIMINUI.

- O FACTOR DE AMPLIAÇÃO NAS EQUAÇÕES DA PROJECÇÃO EM PERSPECTIVA

$$\frac{f}{Z_0}$$

TORNA-SE MENOS SENSÍVEL A Z_0.

- OU SEJA,

$\frac{f}{(Z_0+\Delta Z_0)}$ É MAIS "CORRECTAMENTE" APROXIMADO POR $\frac{f}{Z_0}$ À MEDIDA EM QUE $\frac{Z_0}{\Delta Z_0}$ TENDE A CRESCER.

- A APROXIMAÇÃO DE $\frac{f}{Z_0+\Delta Z_0}$ POR $\frac{f}{Z_0}$ PODE LEVAR-NOS A CRER QUE, SEMPRE QUE A PROFUNDIDADE MÉDIA DE UM OBJECTO FOR GRANDE COMPARADA COM A GAMA DE PROFUNDIDADES DO OBJECTO, A PROJECÇÃO EM PERSPECTIVA PODE SER APROXIMADA PELA ORTOGRÁFICA A MENOS DE UM FACTOR DE ESCALA $\frac{f}{Z_0}$; MAS

ISSO É INCORRECTO!

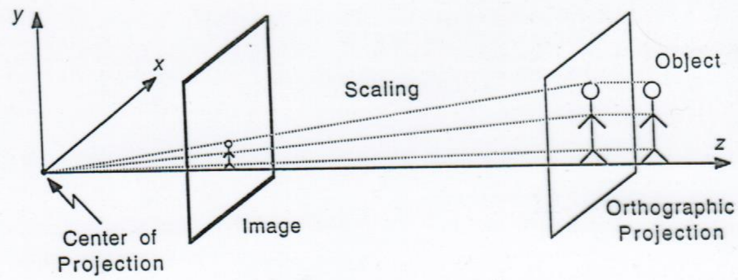
- HÁ DUAS CONDIÇÕES NECESSÁRIAS E SUFICIENTES PARA QUE A PROJECÇÃO EM PERSPECTIVA POSSA SER APROXIMADA PELA PROJECÇÃO ORTOGRÁFICA, A MENOS DE UM FACTOR DE ESCALA UNIFORME:

1. O OBJECTO DEVE ESTAR JUNTO DO EIXO ÓPTICO; O EIXO ÓPTICO É A LINHA QUE PASSA PELO CENTRO DE PROJECÇÃO E É PERPENDICULAR AO PLANO DA IMAGEM;

2. AS DIMENSÕES DO OBJECTO DEVEM SER PEQUENAS;

- APROXIMAÇÃO DA PROJECÇÃO EM PERSPECTIVA NUM PROCESSO DE DOIS PASSOS:

1. PROJECÇÃO ORTOGRÁFICA NUM PLANO PRÓXIMO PARALELO AO PLANO DA IMAGEM;

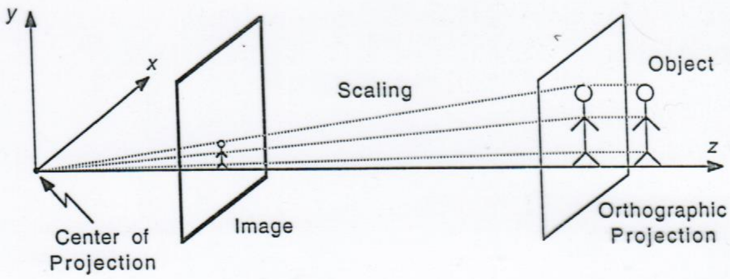


2. PROJECÇÃO EM PERSPECTIVA NO PLANO DA IMAGEM (≡ A "ESCALAMENTO" UNIFORME);

— PROJECÇÃO PARALELA

— É UMA GENERALIZAÇÃO DA PROJECÇÃO ORTOGRÁFICA EM QUE UM OBJECTO É PROJECTADO NO PLANO DA IMAGEM POR UM CONJUNTO DE RAIOS PARALELOS QUE NÃO SÃO NECESSARIAMENTE ORTOGONAIS AO PLANO DA IMAGEM;

— A PROJECÇÃO PARALELA É UMA TRANSFORMAÇÃO LINEAR NO ESPAÇO 3D.

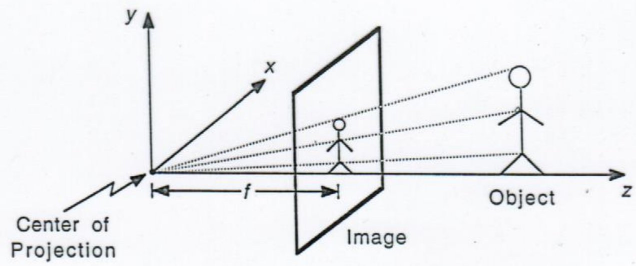


2. PROJECÇÃO EM PERSPECTIVA NO PLANO DA IMAGEM (≡ A "ESCALAMENTO" UNIFORME);

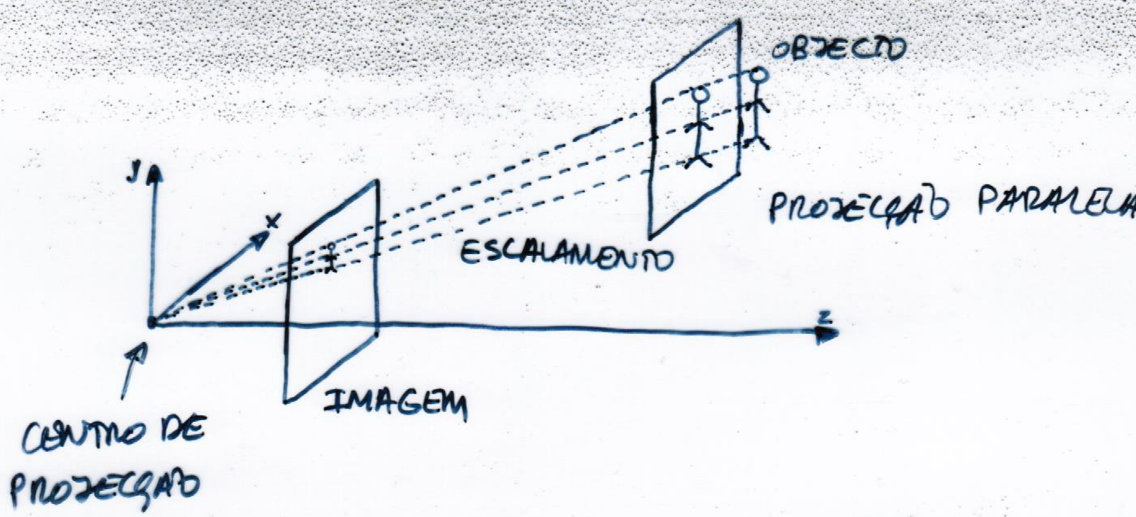
— PROJECÇÃO PARALELA

— É UMA GENERALIZAÇÃO DA PROJECÇÃO ORTOGRÁFICA EM QUE UM OBJECTO É PROJECTADO NO PLANO DA IMAGEM POR UM CONJUNTO DE RAIOS PARALELOS QUE NÃO SÃO NECESSARIAMENTE ORTOGONAIS AO PLANO DA IMAGEM;

— A PROJECÇÃO PARALELA É UMA TRANSFORMAÇÃO LINEAR NO ESPAÇO 3D.



— A PROJECÇÃO EM PERSPECTIVA PODE SER APROXIMADA PELA PROJECÇÃO PARALELA A MENOS DE UM FACTOR DE ESCALA SEMPRE QUE:



— AS DIMENSÕES DO OBJECTO FOREM PARALELAS COMPARADAS COM A DISTÂNCIA MÉDIA DO OBJECTO AO CENTRO DE PROJECÇÃO;

— A DIRECÇÃO DA PROJECÇÃO PARALELA É DEFINIDA PELA "DIRECÇÃO MÉDIA" DA PROJECÇÃO EM PERSPECTIVA;

– QUANDO O OBJECTO PARA ALÉM DE PEQUENO ESTÁ PRÓXIMO DO EIXO ÓPTICO, A PROJECÇÃO EM PARALELO PODE SER FEITA AO LONGO DO EIXO ÓPTICO, E ASSIM OBTÉM-SE A PROJECÇÃO ORTOGRÁFICA;

– MESMO QUANDO AS DIMENSÕES DA CENA NÃO SÃO PEQUENAS QUANDO COMPARADAS COM A DISTÂNCIA MÉDIA DA CENA AO CENTRO DE PROJECÇÃO, PODE SER POSSÍVEL DIVIDIR A CENA EM SUB-CENAS MAIS PEQUENAS, DE TAL MODO QUE AS DIMENSÕES DE CADA UMA DELAS SEJAM PEQUENAS QUANDO COMPARADAS COM A SUA DISTÂNCIA MÉDIA AO CENTRO DE PROJECÇÃO.

NESTAS CONDIÇÕES:

– PODE APROXIMAR-SE A PROJECÇÃO EM PERSPECTIVA DE TODA A CENA POR UM CONJUNTO DE PROJECÇÕES PARALELAS, CADA UMA DAS QUAIS COM A SUA PRÓPRIA DIRECÇÃO DE PROJECÇÃO E ESCALA ⟹

PROJECÇÃO EM PARAPERSPECTIVA

FORMAÇÃO DE IMAGENS COM LENTES

- À MEDIDA QUE A ABERTURA DA CÂMARA ESTENOPEICA É REDUZIDA SERIA DE ESPERAR UMA IMAGEM MAIS DEFINIDA;

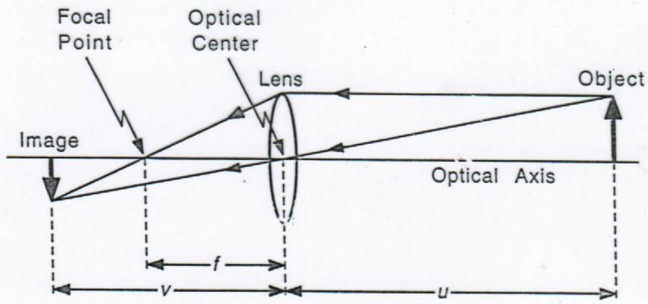
- NO ENTANTO ISSO SÓ SE VERIFICA ATÉ UM CERTO PONTO; A PARTIR DE CERTA DIMENSÃO É NECESSÁRIO CONSIDERAR O FENÓMENO DA DIFRACÇÃO;

DIFRACÇÃO - CURVAMENTO DOS RAIOS DE LUZ EM TORNO DAS ARESTAS DOS OBJECTOS OPACOS;

EM GERAL - QUANTO MENOR FOR O DIÂMETRO DA ABERTURA RELATIVAMENTE AO COMPRIMENTO DE ONDA DA LUZ INCIDENTE MAIS PRONUNCIADA É A DIFRACÇÃO;

OUTRO PROBLEMA COM A REDUÇÃO DA ABERTURA:

- REDUÇÃO NA INTENSIDADE DA IMAGEM;



Thin-Lens Equation: $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

LENTES PORQUE:

– NOS PERMITEM REPRODUZIR A GEOMETRIA "PINHOLE" SEM RECORRER A ABERTURAS DEMASIADO PEQUENAS;

→ IDEALMENTE UMA LENTE CAPTA TODA A LUZ RADIADA POR UM PONTO-OBJECTO E FOCA TODA A RADIAÇÃO NUM ÚNICO PONTO-IMAGEM;

→ CONSIDERAMOS QUE:

– A LENTE É DELGADA;

– O SEU EIXO ÓPTICO É PERPENDICULAR À SUPERFÍCIE ONDE A IMAGEM SE FORMA, QUE É UM PLANO;

— EQUAÇÃO DAS LENTES DELGADAS:

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \quad \left(\frac{1}{u} + \frac{1}{v} = \frac{1}{f_1} + \frac{1}{f_2} \right)$$

— PONTOS-OBJECTO LOCALIZADOS NO INFINITO SÃO "FOCADOS" NO PONTO FOCAL OU FOCO;

— O PLANO PERPENDICULAR AO EIXO ÓPTICO NO FOCO É O PLANO FOCAL;

— VAMOS CONSIDERAR QUE:

→ OS OBJECTOS SÃO FOCADOS NO PLANO FOCAL;

→ O PLANO IMAGEM É COINCIDENTE COM O PLANO FOCAL;

— A ABERTURA É CONSIDERADA CIRCULAR E NO PLANO DAS LENTES;

- O CENTRO DE PROJECÇÃO EFECTIVO DA PROJECÇÃO EM PERSPECTIVA, NO CASO DE UMA LENTE DELGADA ESTÁ COLOCADO NO CENTRO ÓPTICO DA LENTE;

- CENTRO ÓPTICO DA LENTE:
  - É O PONTO CENTRAL DA LENTE NO SEU EIXO ÓPTICO, ATRAVÉS DO QUAL OS RAIOS DE LUZ PASSAM SEM SER DEFLECTIDOS;

  - O CAMPO DE VISÃO OU VISUAL DE UM DISPOSITIVO DE AQUISIÇÃO DE IMAGENS DESCREVE O CONE DE DIRECÇÕES DE VISÃO DO DISPOSITIVO;

  - O CONE INCLUI TODAS AS DIRECÇÕES DOS RAIOS DE LUZ QUE CHEGAM AO PLANO DE IMAGEM DEPOIS DE PASSAREM ATRAVÉS DO CENTRO EFECTIVO DA PROJECÇÃO DA LENTE;

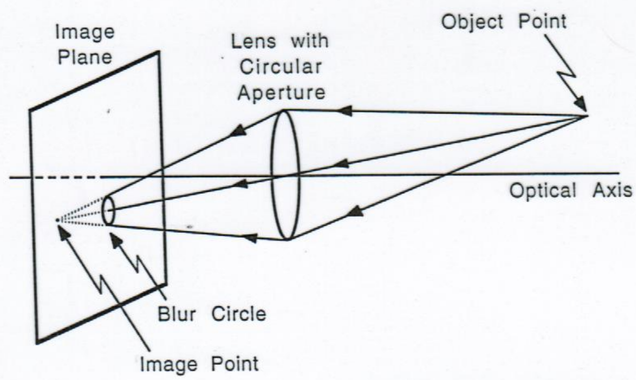
- PARA UM DETERMINADO TAMANHO DE IMAGEM O CAMPO VISUAL É INVERSAMENTE PROPORCIONAL À AMPLIFICAÇÃO DA LENTE;

- LENTES GRANDE-ANGULARES TÊM PEQUENAS DISTÂNCIAS FOCAIS E GRANDES CAMPOS VISUAIS;
- LENTES TELEFOTO TÊM GRANDES DISTÂNCIAS FOCAIS;

- NA PRÁTICA:
  - A PROJECÇÃO EM PERSPECTIVA PODE SER APROXIMADA PELA ORTOGRÁFICA (A MENOS DE UM FACTOR DE ESCALA) QUANDO SE USA UMA LENTE TELEFOTO PARA SE ADQUIRIR A IMAGEM DE UMA CENA DISTANTE QUE TEM UMA GAMA DE PROFUNDIDADES RELATIVAMENTE PEQUENA;
  - ESTA APROXIMAÇÃO NÃO PODE SER APLICADA NO CASO DE SE USAR UMA GRANDE ANGULAR;

- FOCAGEM E DESFOCAGEM
  - COMO RESULTA DA EQUAÇÃO DAS LENTES DELGADAS, PARA UMA DETERMINADA POSIÇÃO DO PLANO IMAGEM, SÓ NUM PLANO-OBJECTO É QUE SÃO PROJECTADOS EM FOCO;

PONTOS



— OS PONTOS QUE NÃO ESTEJAM NO PLANO QUE É FOCADO TÊM COMO IMAGENS CÍRCULOS QUE SÃO DESIGNADOS POR CÍRCULOS DE CONFUSÃO;

— CADA CÍRCULO DE CONFUSÃO É FORMADO PELA INTERSECÇÃO DO CONE DE RAIOS DE LUZ COM O PLANO-IMAGEM;

— O DIÂMETRO DO CÍRCULO É PROPORCIONAL AO DIÂMETRO DA ABERTURA;

— ASSIM À MEDIDA QUE A ABERTURA DIMINUI, A GAMA DE PROFUNDIDADES AO LONGO DA QUAL O MUNDO ESTÁ APROXIMADAMENTE FOCADO (A PROFUNDIDADE DE CAMPO) AUMENTA;