Computer Vision HW 3 (1-5) Reid Glaze

1)

convolution 2D:

$$y[m,n] = x[m,n] * h[m,n]$$
$$= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} a[i,j] * h[m-i, n-j]$$

commutative:

$$x[n] * h[n] = h[n] * x[n]$$
$$y[m,n] = h[m,n] * x[m,n]$$
$$= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} h[i,j] * x[m-i, n-j]$$

If $h[m,n]$ is seperable to $(M \times 1)$ and $(1 \times N)$

$$h[m,n] = h_1[m] \cdot h_2[n]$$
$$y[m,n] = h[m,n] * x[m,n]$$
$$= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} h[i,j] * x[m-i, n-j]$$
$$= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} h_1[i] \cdot h_2[j] * x[m-i, n-j]$$

Definition of convolution 1D:

$$y[n] = x[n] * h[n]$$
$$= \sum_{k=-\infty}^{\infty} x[k] * h[n-k]$$

$$y[m,n] = (h_1[m] * h_2[n]) * x[m,n]$$
$$= h_2[n] * (h_1[m] * x[m,n])$$
$$= h_1[m] * (h_2[n] * x[m,n])$$

Convolution is associative,
Convolve with $h_2[n]$ first and with
$h_1[m]$ later

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

(i) convolution 2D

$$y[1,1] = 1(1) + 2(2) + 3(1) + 4(2) + 5(4) + 6(2) + 7(1) + 8(2) + 9(1)$$
$$= 1 + 4 + 3 + 8 + 20 + 12 + 7 + 16 + 9$$
$$= 80$$

(ii) seperable convolution

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$$

$$y[m,1] = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 16 & 20 & 24 \end{bmatrix}$$

horizontal convolution 1D

$$y[1,1] = [16 \quad 20 \quad 24] [1 \quad 2 \quad 1]$$
$$= 16 + 40 + 24$$
$$= 80$$

Q1 b)

Given: $N \cdot N$ image and $(2k+1) \cdot (2k+1)$ kernel

kernel width = multiplications carried out per image

horizontal slides × kernel height × no vertical slides

- total # multiplications of 2D conv
$(2k+1)(N-2k)(2k+1)(N-2k)$
- The # of operations for 1D convolution for $(2k+1)$ kernel $= (2k+1) \times (N-2k) \cdot N$

- Total operations $= 2 \times (2k+1) \times (N-2k) \times N$

- # of operations saved =
total operations for 2D conv −
total operations for 1D conv

$$= (2k+1)^2 (N-2k)^2 - 2 \times (2k+1) \cdot (N-2k) \cdot N$$
$$= (2k+1)(N-2k)\left[(2k+1)(N-2k) - 2N\right]$$
$$= (2k+1)(N-2k)\left[2kN - 4k^2 + N - 2k - 2N\right]$$
$$= (2k+1)(N-2k)\left[2kN - 4k^2 - 2k - N\right]$$
$$= (2k+1)(N-2k)\left[N(2k-1) - 2k(2k+1)\right]$$

2) A gaussian of form:

$$f(x) = A e^{-(x-u)^2/2\sigma^2}$$ where A is a constant, u is the mean and σ^2 is variance.

- The convolution of two gaussians is a gaussian
- Fourier transform $\hat{f}$ of gaussian f is a gaussian.
- The product fg of gaussian f and gaussian g is a gaussian

$$\phi(x,t) = ax^2 + bx + ct^2 + dx + et + f$$

$$t_2 = t - \frac{b}{2c}x$$

$$\therefore Q(x,t) = ax^2 + \left[b\left(\frac{e}{2c}\right) + d\right]x + c\left[t_2 - \frac{b}{2c}x\right]^2 - \left(\frac{b}{2c}\right)^2 + \left(\frac{e}{2c}\right)^2 + f$$

if $f = e^{-(x-u^2)/\sigma^2}$ & $g = e^{-(x-v^2)/\tau^2}$

$$(f*g)(x) = \int e^{-\frac{(t-u)^2}{\sigma^2} - \frac{(x-t-v)^2}{\tau^2}} dt$$

$$= e^{-A(x-r)^2} \int_{-\infty}^{\infty} e^{-B(t-x-s)^2} dt$$

$$\boxed{\therefore (f*g)(x) = c\, e^{-A(x-r)^2}}$$ which is a gaussian as the integral yields a constant c

3) Dimensionality reduction is the process of reducing the number of input variables in the training data. Some features are redundant and make the data harder to work with.

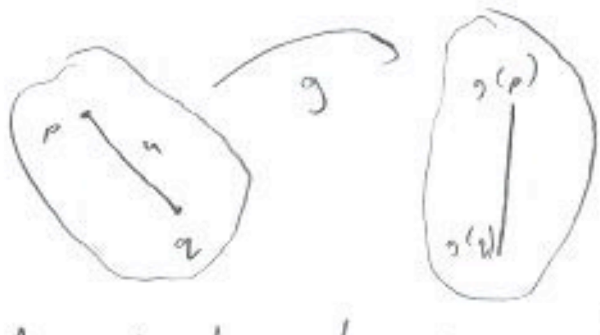
Storage is a big issue when dealing with multiple images. Dimensionality reduction can be used to preserve important data. It can save memory and improve speed of execution for image processing. Dimensionality reduction can used with image transformation.

The disadvantage of using dimensionality reduction could be loss of important features.

4)

Rigid body displacement

Object :- $O \subset R^3$
Map :- $g : O \to R^3$



a) A displacement is a transformation of points

Translation (g) of points induces

an action (g_*) on vectors

$\| g(p) - g(q) \| = \| p - q \|$

g_* acts on vectors

$\therefore$ vector $(p-q)$ can be written

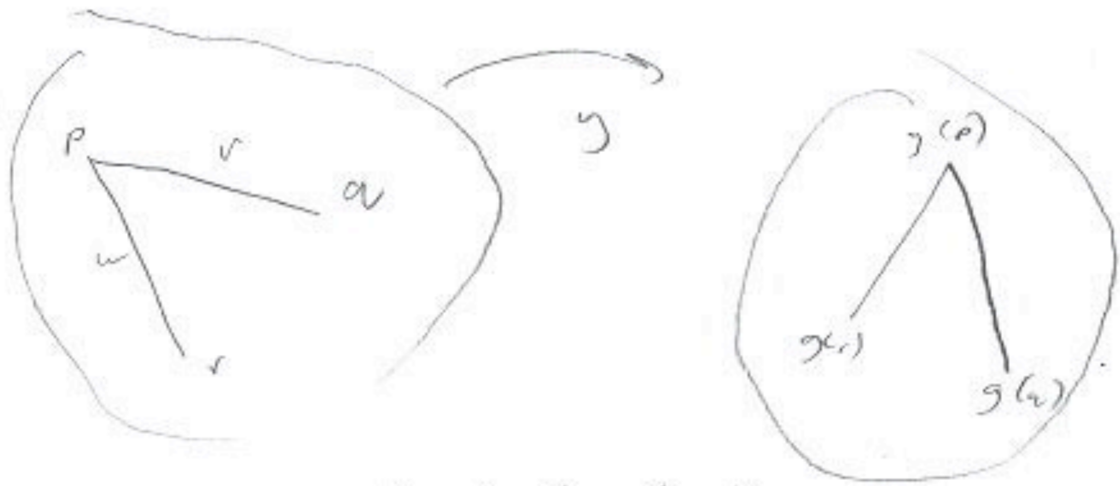
as $g(p) - g(q)$ as well as $g_*(p-q)$

$\therefore \| g_*(p-q) \| = \| p - q \|$

as $(p-q) = v$

$\therefore \| g_* v \| = \| v \|$

b)

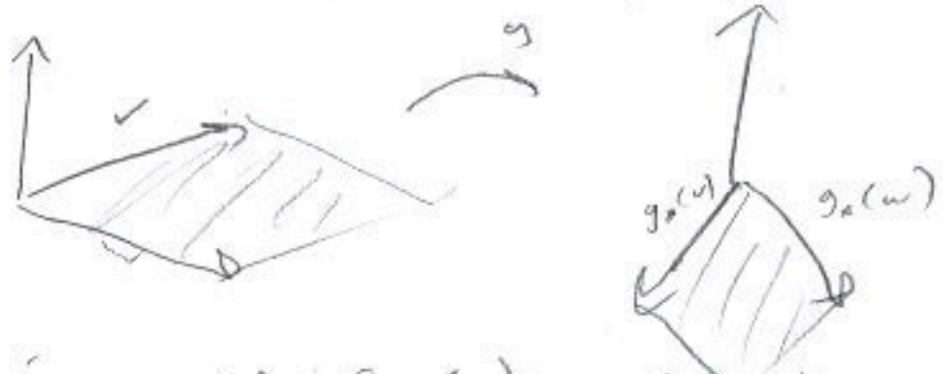


here $\|g^* v\| = \|v\|$

$\|g^* w\| = \|w\|$

Rigid body transformation preserves angles & distances

- The cross product vector is the line perpendicular to the plane containing v and w
- As angles are preserved the cross product of the vectors before and after the transformation are preserved



$\therefore g_*(v) \cdot g_*(w) = g_*(v \cdot w)$

5)

To find efficiency of achieving approx. Gaussian filtering using multiple averagings, determine the specific averaging filter required to approximate Gaussian filtering of desired standard deviation

$$\sigma_{av} = \sqrt{\frac{w^2 - 1}{12}}$$

perform n averaging, the variances of filters sum to equal

$$\sigma_{nav} = \sqrt{\frac{nw^2 - n}{12}}$$

obtain width of averaging filter

$$\sigma_{nav}^2 = \frac{nw^2 - n}{12}$$

$$12\,\sigma_{nav}^2 = n\,(w^2 - 1)$$

$$\frac{12\,\sigma^2_{nav}}{n} = w^2 - 1$$

$$w^2 = \frac{12\,\sigma^2_{nav}}{n} + 1$$

$$w_{ideal} = \sqrt{\frac{12\,\sigma^2_{nav}}{n} + 1}$$

to apply averaging filter, use two sizes of filter. The first filter width w_1 is equal to the first odd integer greater than w_{ideal}

Apply filter width w_1 for m passes
w_u for $(n-m)$ passes

The variances of filters add
st dev =

$$\sigma = \sqrt{\frac{m w_1^2 + (n-m) w_u^2 - n}{12}}$$

$$= \sqrt{\frac{m w_1^2 + (n-m)(w_1+2)^2 - n}{12}}$$

$$\sigma^2 = \frac{m w_1^2 + (n-m)(w_1+2)^2 - n}{12}$$

$$\rightarrow 12\sigma^2 = m w_1^2 + n(w_1+2)^2 - m(w_1+2)^2 - n$$

$$\rightarrow 12\sigma^2 + n - n(w_1+2)^2 = m\left[w_1^2 - (w_1+2)^2\right]$$

$$\rightarrow 12\sigma^2 + n(1 - w_1^2 - 4 - 4w_1) = m\left[w_1^2 - w_1^2 - 4 - 4w_1\right]$$

$$\rightarrow 12\sigma^2 + n(-w_1^2 - 3 - 4w_1) = m\left[-4 - 4w_1\right]$$

$$\rightarrow m = \frac{12\sigma^2 - n w_1^2 - 4n w_1 - 3n}{-4w_1 - 4}$$

To approximate Gaussian filtering..

- calculate w_{ideal} and hence w_l and w_u
- calculate m
- apply mean filtering of width w_l m times
- apply mean filtering of width w_u, $(n-m)$ times

to validate efficiency

$\sigma = 10$ $n = 5$

$w_{ideal} = 16$

$w_l = 15$ $w_u = 17$

$m = 4$ (rounded)

$\sigma = 9.8475$ which is of error

$e = 10 - 9.9475 = 0.0525$

$n = 3 \rightarrow e = 0.1638$ $(10 - 9.8362)$

$n = 5 \rightarrow e = 0.0525$

$n = 10 \rightarrow e = -0.0333$ $(10 - 10.0333)$

in the case of $n = 10$.
the radius of the bell
curve for the given Gaussian.
therefore, it is reccomended
to keep n no more than 6