Q.1

(1)

$$\begin{array}{cc|cc} 21 & 6 & 15 & 5 \\ -9 & 3 & 6 & 3 \\ \hline 3 & -3 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{array}$$

first pass $T_0 = 16$ based on larger value 21

$$\begin{array}{cccc} 21 & 6 & -9 & 3 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ sp & zr & zr & zr \end{array}$$

$\downarrow$ $1.5\,T_0$

$$\begin{array}{cc|cc} 24 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \xrightarrow{\text{Correction}} 21-24 < 0 \rightarrow 24 - \frac{T_0}{4} \rightarrow \begin{array}{cc|cc} 20 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

bitstream after first pass: 11/00/00/00/~~0~~ 9 bits in total.

Second pass $T_1 = \frac{T_0}{2} = 8$

$$\begin{array}{cc|cc} \cancel{21} & 6 & 15 & 5 \\ -9 & 3 & 6 & 3 \\ \hline 3 & -3 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{array}$$

$$\begin{array}{ccccccccccc} 6 & -9 & 3 & 15 & 5 & 6 & 3 & 3 & -3 & 1 & 0 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ iz & sn & zr & sp & zr & zr & zr & zr & zr & zr & zr \end{array}$$

$\downarrow$ $1.5\,T_1$

$$\begin{array}{cc|cc} 20 & 0 & 12 & 0 \\ -12 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \xrightarrow{\text{correction}}$$

$21 - 20 > 0 \ (+2)$

$-9 - (-12) > 0 \ (+2)$

$15 - 12 > 0 \ (+2)$

$$\begin{array}{cc|cc} 22 & 0 & 14 & 0 \\ -10 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

bitstream after 2nd pass

11/00/00/00/~~0~~/10/01/00/11/00/00/00/00/00/00/00/~~00~~

30 bits

(2.) To obtain the original image let

$$F_1 = \begin{array}{cc|cc} 21 & 6 & 15 & 5 \\ -9 & 3 & 6 & 3 \\ \hline 3 & -3 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{array}$$

F_2

Then obtain: $A_2^T F_2 A_2 = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}21 & 6\\ -9 & 3\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \begin{bmatrix}21 & 3\\ 33 & 27\end{bmatrix}$

and finally: $A_1^T F_1 A_1 = \frac{1}{2}\begin{bmatrix}1 & 0 & 1 & 0\\ 1 & 0 & -1 & 0\\ 0 & 1 & 0 & 1\\ 0 & 1 & 0 & -1\end{bmatrix}\left[\begin{array}{cc|cc}21 & 3 & 15 & 5\\ 33 & 27 & 6 & 3\\ \hline 3 & -7 & 0 & -1\\ 1 & 0 & 0 & 0\end{array}\right]\begin{bmatrix}1 & 1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 1 & -1 & 0 & 0\\ 0 & 0 & 1 & -1\end{bmatrix} = (\ldots)$

To obtain the reconstructed image let $F_1 = \left[\begin{array}{cc|cc}20 & 0 & 12 & 0\\ -12 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right]$ based on 30 bit availability

and repeat the process.

Q.2

k	0	1	2	3	4	5	6	7
R(k)	1	0.5	0.3	0.1	0.8	0.3	0.2	0.1

1. By inspecting the values of $R(k)$ we can recognize a period of 4. (distance of two max values 1 and 0.8)
   Therefore the pitch = 4

2. Using linear prediction we find that

$$\begin{bmatrix}R(0) & R(1)\\ R(1) & R(0)\end{bmatrix}\begin{bmatrix}a_1\\ a_2\end{bmatrix} = \begin{bmatrix}R(1)\\ R(2)\end{bmatrix} \Rightarrow \begin{bmatrix}a_1\\ a_2\end{bmatrix} = \begin{bmatrix}R(0) & R(1)\\ R(1) & R(0)\end{bmatrix}^{-1}\begin{bmatrix}R(1)\\ R(2)\end{bmatrix}$$

$$\begin{bmatrix}a_1\\ a_2\end{bmatrix} = \begin{bmatrix}1 & 0.5\\ 0.5 & 1\end{bmatrix}^{-1}\begin{bmatrix}0.3\\ 0.5\end{bmatrix} = \ldots$$

3. without LPC 8×8 bits/value $= 64$ bits
   with LPC: (Pitch, a_1, a_2) $\times$ 8 bits/value $= 3 \cdot 8 = 24$ bits.

4. $e(n) \to \boxed{(a_1; a_2)} \to$ speech

where $e(n) = \delta(n) + \delta(n-4) + \delta(n-8)$

Then

$$x(n) = a_1 x(n-1) + a_2 x(n-2) + e(n)$$

and

$$x(0) = a_1 x(-1) + a_2 x(n-2) + e(n) = a_1 0 + a_2 \cdot 0 + 1 = 1$$

$$x(1) = a_1 x(0) + a_2 x(1) = a_1 x(0) + a_2 \cdot 0 = a_1$$

$$x(2) = a_1 x(1) + a_2 x(0) = a_1 a_1 + a_2 (1) = a_1^2 + a_2$$

$$x(3) = a_1 x(2) + a_2 x(1) = a_1(a_1^2 + a_2) + a_2 \cdot a_1$$

$$x(4) = a_1 x(3) + a_2 x(0) + 1 = \ldots$$

Q.3

1. They both using a linear predictor but the underlying model is different

2. With 20 frames/sec the motion between two consecutive frames will be larger than two consecutive frames in 30 frames/sec

So the search area should be larger $p > 16$.

3. Since the resolution is decreased p should be decreased as well thus $p < 16$.

4. It means that the calculation of the 2D transform for an image can be substituted by 2-1D transforms over rows and columns separately → easier calculation

5. The image is not random

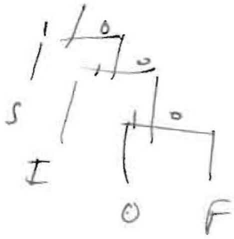
6. IPBBBPBBBB

7. DCT maps N time values → N frequency values
MDCT maps $2N$ time values → ~~N~~ frequency values

8. 1. Spatial
2. frequency
3. perceptual
4. time

9. SISIFOS

	S	I	F	O
p	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	$\frac{1}{7}$



S	1
I	01
O	001
F	000

rate: $1 \times \frac{3}{7} + 2 \cdot \frac{2}{7} + 3\frac{1}{7} + 3\frac{1}{7} = \frac{13}{7}$ bits/sample $\geq$ entropy

10 Same as in 8.

Q.4

1. $4DCT = T f T^T = \begin{pmatrix} g(0,0) & g(0,1) & \cdots & g(0,3) \\ g(3,0) & & & g(3,3) \end{pmatrix}$

Then $\begin{pmatrix} \text{round}\left(\frac{g(0,0)}{10}\right) & \cdots & \\ & & \text{round}\left(\frac{g(3,3)}{100}\right) \end{pmatrix} \cdots$

2. Repeat $T * g T^T = \cdots$

3. $g(x,y)$ is smoother than $f(x,y)$ due to the linear filtering operation with $h(x,y)$. Therefore the DCT of $g(x,y)$ has more energy compaction than the DCT of $f(x,y)$ therefore JPEG of $g(x,y)$ is better (more efficient)

4. Yes the more the filtering the smoother the image the better the JPEG operation.