1. .z!.! ! Zc! Zx! C! C.. C! C! C. C! Zx. C! Cx. Zx. C. Z!. C! Xc!! X!!!..! ! X!!! C!. C x c!!!!! Sx!! X! C!!! Zc x!! C. X. Cx!. X **Ccx I x c. ! X. Xcx. Cxx zc c c. I cz. X. C xzcc c c. C. A z c. The. A. C! A. Cz. Z. Xx. Z. C. X. X. Z!! X!. C. And c! Cc c. . Xx. Cc. . C cx z. X! C. ! C c. C. X cc. C. X. if. .!!!. C.. . C. X. Cx. Cx c. C! . X c!. X. C. X c. .. A. A. C! X. C. C c. Z x.. Xx x c c x c Xx. !. Xccz.! X!. !. I x.. Cc.. C. C**

2. **!!.Czc x. XxXXXxc.XcXXxcXXza. C c. C. .. A. A. A. A! X c! C!. C. C! !, c c! !! Z., c. C. C. C!, x c! C. ! !! C! C!!!!. X. Z. . X. C! .!!!!. X!!!! X z cccx!!!! Zc!. C!.! !!!! ! X c.!! ! C,!! Z!!! Z! Cc! ! C. X!! A.! A x.!! C. C!. !.!,!! X! C! C. C!, c. C!. C. C! X! !!! cXfinition: Transform coding** is a type of data compression for "natural" data like audio signals or photographic images. The transformation is typically lossless on its own but is used to enable better (more targeted) quantization, which then results in a lower quality copy of the original input (lossy compression).
3. **Purpose:** We apply transform coding to one image, or the prediction error image, with the goal to minimize the number of bits needed for transmitting or storing this image. Transform coding can be viewed as being part of the redundancy reduction and also the irrelevance reduction. In this way, it is an addition to the predictive coding (e.g. with motion compensation), or can also be used as an alternative. But unlike the predictive coding, for transform coding, we can also use psycho-optical effects, because we can take advantage of the different sensitivities of our eye to different spatial frequencies, as signified by the Contrast Sensitivity Function of the eye.

4. **Algorithm**: The basic principle of applying transforms to our images or prediction error images is, to first subdivide our image into blocks (for instance size 8x8 pixels), and then apply our transform to each block. Usually, we use a so-called separable transform (meaning it is defined only for 1 dimension, like a DFT or a DCT), which is then applied first to one dimension (for instance the rows). The result is then used to apply the transform to the other dimension, like the columns.

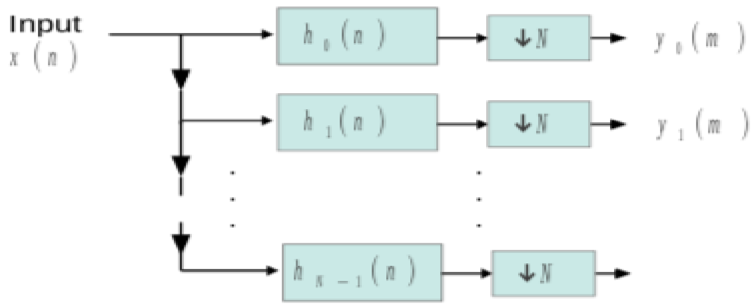
1. Difference between predictive coding and transform coding

- **Predictive Coding**: It is used for lossless compression. Lossless compression takes advantage of statistical dependencies in the signal (correlations, predictability). Hence the compression ratio depends on the signal. In natural images, there are many even surfaces, meaning that there is a high correlaton between neighbouring pixels. A simple method to use this correlaton between neighbouring pixels is predictve coding.

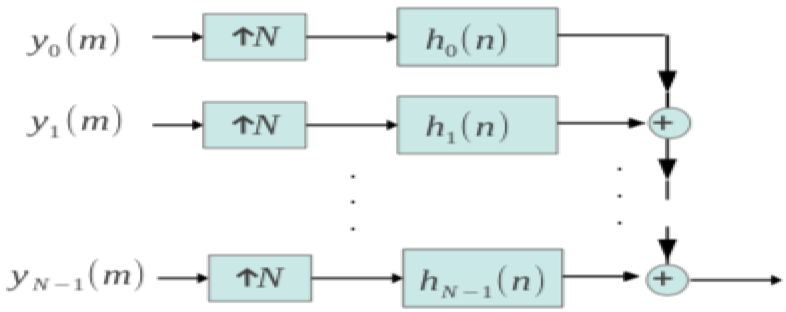
- **Transform coding:** It can be viewed as being part of the redundancy reduction and also the irrelevance reduction. In this way, it is an addition to the predictive coding, or can also be used as an alternative. But unlike the predictive coding, for transform coding, we can also use psycho-optical effects, because we can take advantage of the different sensitivities of our eye to different spatial frequencies, as signified by the Contrast Sensitivity Function of the eye.

2. What is critical sampling? Explain Critical sampled filter banks

- A critical sample filter bank is a filter bank, where the downsampling factor is identical to the number of subbands.
- **Analysis Filter:**
  - $y = T^T . x. T$
  - we first, filter our subband signals with our analysis filters and then downsample them by N

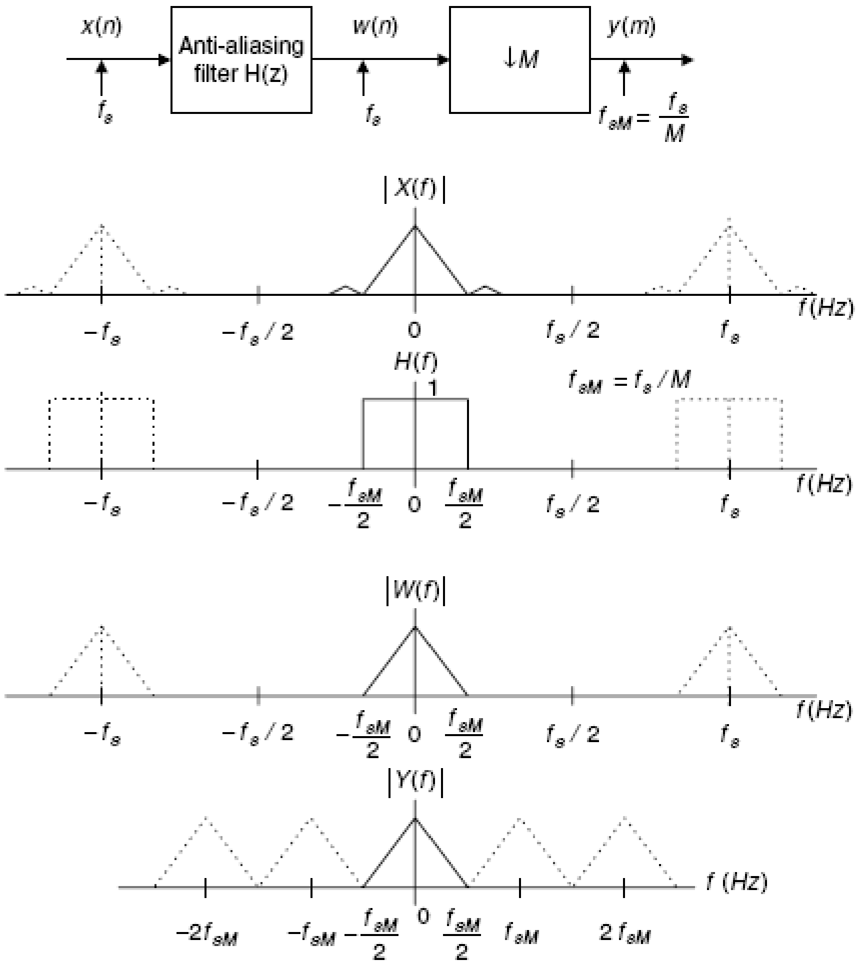


- **Synthesis Filter:**
  - $x = T^{-1}.y. T^{-T}$
  - we first upsample our subband signals by N and then filter them with our synthesis filters.



3. Downsampling in the frequency domain

- Our signal frequency must be really smaller than half of our new sampling frequency. This is called the Sampling theorem or Nyquist Theorem.



4. Why 8*8 block used in DCT?

- Any matrices of sizes greater than 8 X 8 are harder to do mathematical operations (like transforms etc..) or not supported by hardware or take longer time.
- Any matrices of sizes less than 8 X 8 dont have enough information to continue along with the pipeline. It results in bad quality of the compressed image.

5. How is downsampling and upsampling done without getting motion artifacts?

- To avoid artefacts, we suitably **lowpass** our image **before downsampling**, and also **lowpass filter it after upsampling** as a sort of interpolation.

6. Lossless vs. Lossy coding

- Lossless Compression (redundancy reduction)
    - Lossless compression takes advantage of statistical dependencies in the signal (correlations, predictability).
    - Hence the compression ratio depends on the signal.
    - If the captured image or video is to be further processed in a studio or investigated like an x-ray image, it is best to apply lossless

compression, since the assumptions of lossy compression are not true here.

    ○ An example is predictive coding.

● Lossy Compression (redundancy and irrelevance reduction)

    ○ **Lossy** compression takes advantage of the properties of the receiver, here the human eye.

    ○ This already starts with the capturing and recording process, by having only a limited amount of pixels (sampling the image at the location of the pixels), and by only having a limited number of pictures (frames) per second.

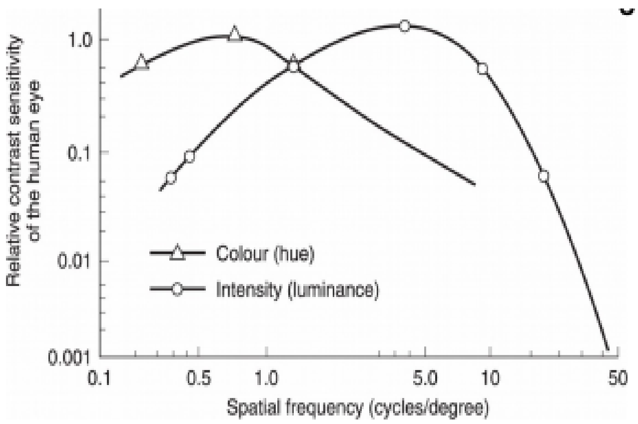
    ○ This uses the fact that the eye has only a limited spatial resolution (as shown by its Contrast Sensitivity Function), and a limited temporal resolution (hence the 25 pictures per second).

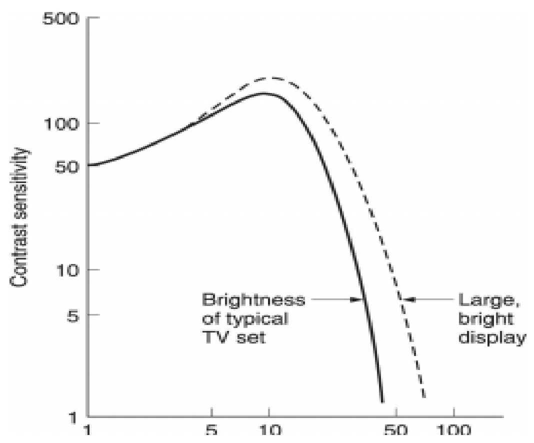
7. 3 properties of the eye + they can be used in video coding
    a. CSF (limited spatial resolution, limited temporal resolution)
    b. More sensitive to luminance
    c. the receptive field of eye
8. CSF

● Contrast Sensitivity Function measures the sensitivity of the eye for different spatial frequencies.

● It describes the capacity of the HVS to recognize the difference in Luminance and chrominance as a function of contrast and spatial frequency

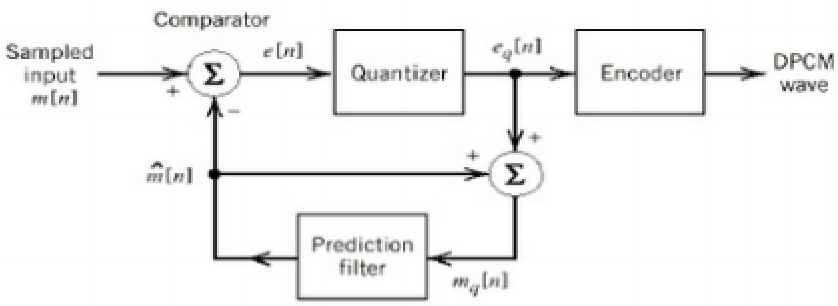
● We can take advantage of this property of eye and decompose our image into subbands with different spatial frequencies and quantize them with an accuracy which is suitable to the sensitivity for the eye at the spatial frequency of that subband.



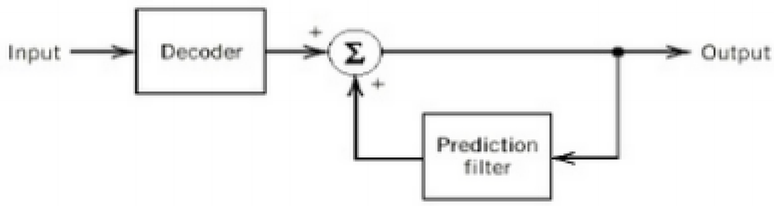


9. Explain DPCM

- In Differential Pulse Code Modulation, prediction is based on quantized samples.
- The actual pixel m(n) is subtracted from the prediction $\hat{m}$(n) and the difference, the prediction error is transmitted to the receiver.
    - e(n)=m(n)-$\hat{m}$(n)
- The prediction error will typically be small due to spatial correlation and compression can be acheived by representing common, small prediction errors with short binary codes and large, less common errors with longer codes.
- The reconstruction equation can be written as
    - $\hat{m}_q$(n) = $e_q(n)$ +$\hat{m}$(n)
- And finally, the reconstruction error equation which is computed the same way as the quantization error
    - $\hat{m}_q$(n) - m(n) = $e_q(n) + e(n) = q$



(a)



(b)

10. DFT formula

$$X(k_1,k_2)=\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} x(n_1,n_2)\cdot e^{-j\frac{\pi}{N_1}\cdot k_1 n_1}\cdot e^{-j\frac{\pi}{N_2}\cdot k_2 n_2}$$

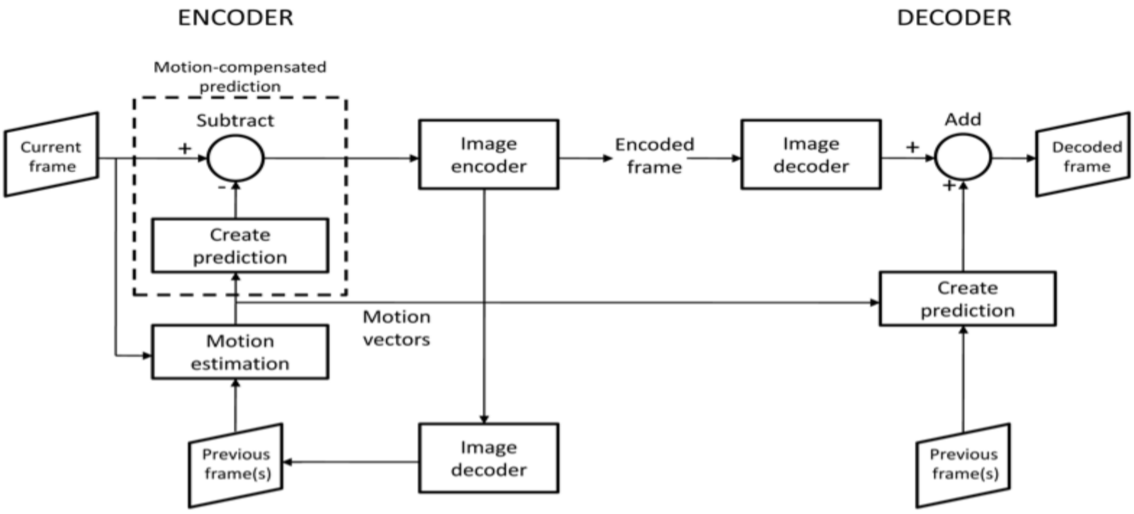
11. Why DCT is better than DFT? Difference between DFT and DCT.

- DCT
    - Fast algorithm
    - Good energy compaction
    - Only real coefficients
- DFT
    - Complex Coefficients
    - Energy compaction is worse than DCT.

12. Why do we use DCT for images?

- DCT is the most commonly used type in video and image coding because the higher bands have zeros of their frequency response at DC, hence suppressing the high DC energy of usual pictures for the higher subbands sufficiently. In this way, we do not have to encode the DC energy in higher subbands again.
- DCT is relatively simple to compute, it is separable and has pretty good "energy compaction" properties.

13. Motion estimation (Block diagram + explain the principle) and components



Motion estimation is the process of determining motion vectors that describe the transformation from one 2D image to another; usually from adjacent frames in a video sequence.

To use this model assumption, we search the piscture for displacements (movements) and estimate motion vectors for the image. These motion vectors then increase the accuracy of the prediction over the time dimension.

The prediction error is then computed based on the motion-compensated predicted image. The encoder then sends the **motion vectors** and the **prediction error** (the so-called **residual frame**) to the decoder. The decoder can compute the new frame using the motion vectors, the previous frame(s) (and perhaps also future frames, if we also compute motion vectors based on future frames), plus the residual frame, to compute the current frame.

14. Explain: motion vector, motion estimation, motion compensation, motion prediction

- **Motion Vector**: The vector which describes the motion of an image block
- **Motion prediction**: The algorithm which computes the predicted image
- **Motion Compensation**: Computing the prediction error using moton prediction
- **Motion Estimation**: The algorithm for estimating the motion vector

15. What is Interpolation.? Explain the differences between the different interpolation methods

Interpolation is a method of constructing new data points within the range of a discrete set of known data points.

- Piecewise constant interpolation(nearest neighbour Interpolation):
  - It locates the nearest data value, and assign the same value.
- Linear interpolation
  - It is a method of curve fitting using linear polynomials to construct new data points within the range of a discrete set of known data points.
- Polynomial interpolation:
  - It is the interpolation of a given data set by the polynomial of lowest possible degree that passes through the points of the dataset.
- Spline interpolation
  - It is a form of interpolation where the interpolant is a special type of piecewise polynomial called a spline. Spline interpolation is often preferred over polynomial interpolation because the interpolation error can be made small even when using low degree polynomials for the spline.

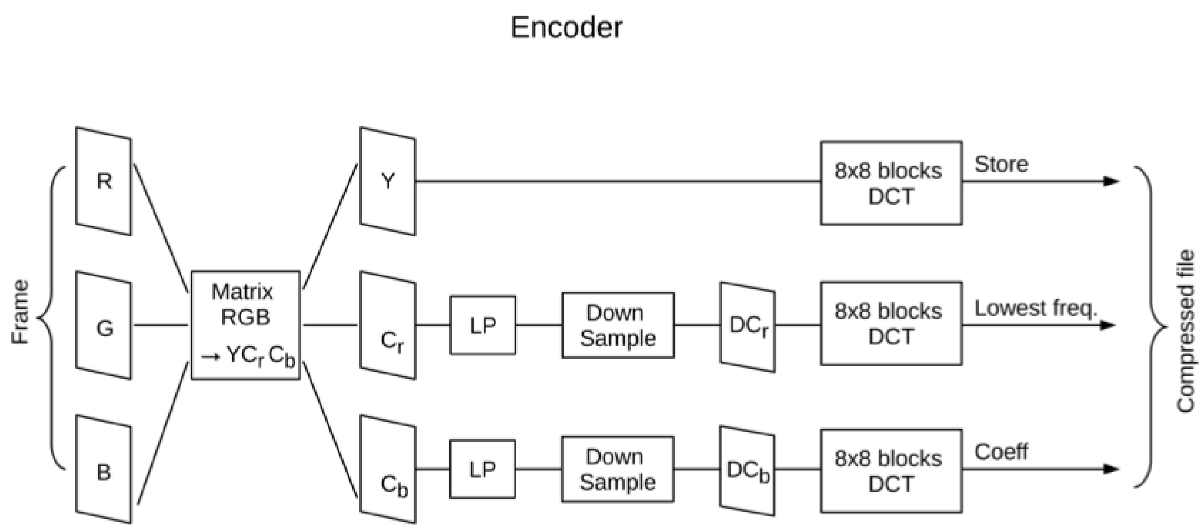
16. prediction frame

- **Inter Frames**
  - **Forward prediction (P-Frames)**: prediction is based on past frames (also more than 1 frame in the past).
    - **Advantage:** we don't need any additional delay for our prediction.
    - **Disadvantage:** For instance for scene cuts, it doesn't work for the following frame, because of missing similarity
    - -**Backwards prediction**: prediction based on future frames.
    - **Advantage**: can be used if there is no or not sufficient similarity to the past frame(s).
    - Disadvantage: We need a delay to wait for the future frame to predict from.
  - **Bidirectional prediction (B-Frames)**: We use both directions to predict the current frame (weighted average of both directions, or taking the best of

both). For each macroblock of the motion compensation, the encoder transmits a mode information
    - Mode a: forward prediction
    - Mode b: backward prediction
    - Mode c: an average of the 2 predictions.
    - **Advantage:** We obtain the best prediction because we have the most information as a basis for it. It improves compression efficiency by up to 50%.
    - **Disadvantage:** It needs side information for the signalling, and it need delay for the future frame.
- **Intra Frames (I-Frames):** No prediction using motion compensation :
    - **Advantage**: independent of other frames, can be used to start decoding the video sequence.
    - **Disadvantage:** Need the highest bit-rate.

17. Explain how the WHT is used in H.264 + compare it to generic DCT

- We have a 16x16 macroblock and 16 4x4 transforms in the macroblock. The **16 DC coefficients** of these transforms are taken into a new block, which is then again transformed using WHT.
- WHT is simpler to implement and leads to smaller subband coeffients, which need fewer bits. In DCT, DC coefficients have no advantage compared to the WHT.



18. Draw/explain how the signal matrix looks like after the DFT transform

19. What is the optimal transform matrix?

A transform matrix, which leads to a **minimized number of bits** for coding an image is Optial Transform Matrix

20. Calculate WHT of the given matrix ----
21. Calculate Axx of given x -----
22. Calculate the KLT from the Axx
23. 2D Z Transform Calculation
24. 2D Z transform Formula

$$X_z(z_1, z_2) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} x(n_1, n_2) z_1^{-n_1} z_2^{-n_2}$$

Downsampling

Yds = np.zeros(Y.shape)

Yds[0::N, 0::N] = Y[0::N,0::N]

Upsample

Yus[0::N, 0::N] = Yds[0::N,0::N]