Chapter 1

INTRODUCTION

Definition of Image and Video Compression

Image and video data compression[1] refers to a process in which the amount of data used to represent image and video is reduced to meet a bit rate requirement (below or at most equal to the maximum available bit rate), while the quality of the reconstructed image or video satisfies a requirement for a certain application and the complexity of computation involved is affordable for the application.

---

[1] In this book, the terms image and video data compression, image and video compression, and image and video coding are synonymous.

Functionality in Visual Transmission and Storage

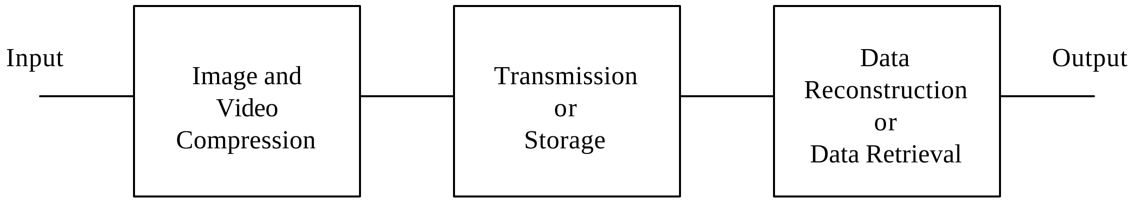


Figure 1. 1 Image and video compression for visual transmission and storage

Information and Data:

Information: Knowledge, facts and news

Data is used to represent information.

Bit rate (also known as coding rate)

bits per second (bps), for visual communication.

bits per pixel (bpp), for image storage

bits per symbol, for information source coding

1.1 Practical Needs for Image and Video Compression

An Illustrative Example:

- Assume the present switch telephone network (PSTN) modem:

  at a maximum bit rate of 56,600 bps

- Assume each video frame:

  a resolution of 288 by 352

  comparable with that of a TV picture

  referred to as common intermediate format (CIF).

- Each of the three primary colors RGB:

  8 bits per pixel

- Frame rate in transmission:

  30 frames per second

- The required bit rate:

  $288 \times 352 \times 8 \times 3 \times 30 = 72{,}990{,}720$ bps

Comments

- We have to compress the video data by at least **1289** times in order to accomplish the transmission described in this example.
- Audio not accounted yet.
- Video services such as 3-D movie, 3-D game, and HDTV.

❖ Advanced image and video data compression becomes **an enabling technology** to bridge the gap between the required huge amount of video data and the limited hardware capability.

1.2 Feasibility of Image and Video Compression

1.2.1 Interpixel Redundancy

1.2.1.1 Spatial Redundancy

Statistical correlation among pixels within an image frame: **intraframe redundancy**.

Figure 1.2 A picture: "Boy and Girl"

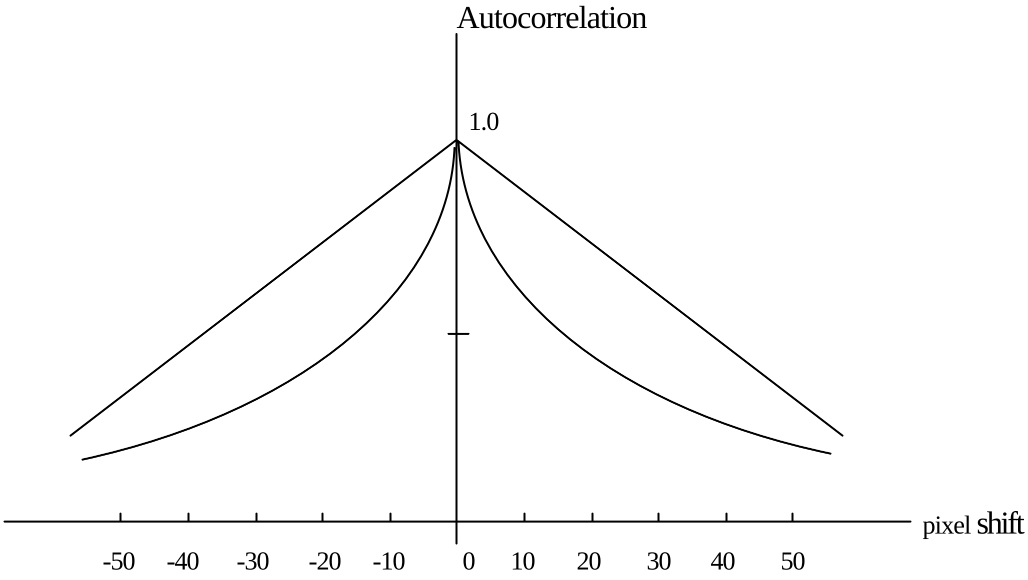


Figure 1. 3 Autocorrelation in horizontal directions for some testing pictures after [kretzmer 1952]

- 0.97 to 0.99 for one- or two-pixel shifting. (For very detailed pictures, it can be from 0.43 to 0.75.)
- Not necessary to represent each pixel in an image frame independently. Instead, one can predict a pixel from its neighbors.
- Removing a large amount of the redundancy within an image frame, we may save a lot of data in representing the frame, thus achieving data compression.

1.2.1.2 Temporal Redundancy

Statistical correlation between pixels from successive frames: **interframe redundancy**.

Figure 1.5 (a) 21st and (b) 22nd frames of Ms. America sequence

Observation made in [mounts 1969]:

- For a videophone-like signal with moderate motion in the scene, on average, less than 10% of pixels change their gray values between two consecutive frames by an amount of 1% of the peak signal.

- Removing a large amount of temporal redundancy leads to a great deal of data compression: motion compensated (MC) predictive coding.

1.2.2 Psychovisual Redundancy

- Originates from the characteristics of the human visual system (HVS).
- Visual information not perceived equally by HVS.
- If apply less data to represent less important visual information, perception will not be affected.
- In this sense, **some visual information is psychovisually redundant**.
- Eliminating psychovisual redundancy leads to data compression.
- Five maskings discussed below

1.2.2.1 Luminance Masking

- Also referred to as **luminance dependence** and **contrast masking**.
- Most fundamental one among the five.
- Concerns **brightness perception** of the HVS: the detectability of one stimulus when another stimulus is present simultaneously.
- Consider monochrome Figure 1.7.

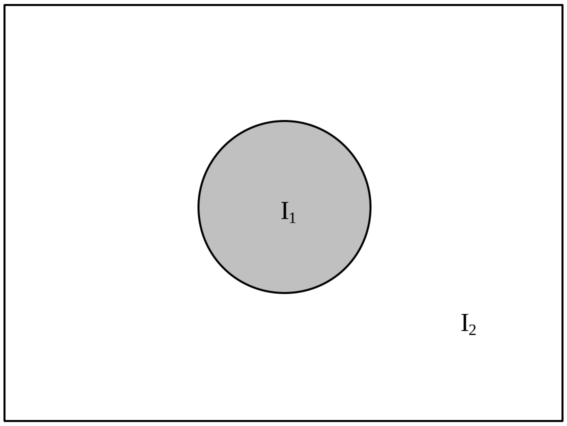


Figure 1.7 A uniform object with gray level I_1 imposed on a uniform background with gray level I_2

- ✓Question: Under what circumstances can the disk-shaped object be discriminated from the background by the HVS?
- ✓The effect of one stimulus (background, masker) on the detectability of another stimulus (disk).
- ✓Two extreme cases are obvious.
  - ❑ If the difference between the two gray levels is quite large, no problem.
  - ❑ If the two gray levels are the same, the HVS cannot identify the object.
- ➢Concern: critical threshold in the gray level difference for discrimination.
- ➢**Threshold:** $\Delta I = I_1 - I_2$. ⇔ the object can be noticed by the HVS with a 50% chance.
- ➢**Weber's law**:

$$\frac{\Delta I}{I} \approx constant \approx 0.02 \quad . \qquad (1.1)$$

- Implication: when the background is bright, a larger difference in gray levels is needed for the HVS to discriminate the object from the background. Vice verser.
- Weber’s law holds for all other human senses as well.

1.2.2.2 Texture Masking

- Also known as ***detail dependence, spatial masking,*** or ***activity masking.***
- It states that the discrimination threshold increases with increasing picture detail.
- **Coarse quantization** ⇒ **false contouring**
    - ✓No. of quantization levels decreases from 256 (8 bits) to 16 (4 bits).
    - ✓Can be explained by using texture masking.

1.2.2.3 Frequency Masking

- **Picture-independent**.
- It states that the discrimination threshold increases with frequency increase.
- Can be well illustrated by alleviating false contouring using a method called improved gray-scale (IGS) quantization [gonzalez 1992, p. 318].
  - ✓The low frequency quantization error is converted to the high frequency noise.
  - ✓HVS is less sensitive to the high frequency content.
  - ✓HVS functions like a low pass filter.
- Use in JPEG.

1.2.2.4 Temporal Masking

- Another **picture-independent** feature.

- It states that it takes a while for the HVS to adapt itself to the scene when the scene changes abruptly. During this transition the HVS is not sensitive to details.
- The masking takes place both before and after the abrupt change.
    - **Forward temporal masking** if it happens after the scene change.
    - Otherwise, **backward temporal masking**

1.2.2.5 Color Masking

- Digital color image processing
- Any visible light corresponds to an electromagnetic spectral distribution.
- **Luminance:** perception of **brightness**
- **Chrominance:** perception of: **hue** and **saturation**.
    - Hue ↔ the dominant wavelength
    - Saturation ↔ the purity of a color

RGB Model *

- **The most well-known** color system
- The color sensitive area in the HVS consists of three different sets of cones and each set is sensitive to the light of one of the three primary colors: red, green and blue.
- Many research results are available, e.g., C.I.E. (Commission Internationale de l'Eclairage) chromaticity diagram
- Used mainly in color image acquisition and display.
- In color signal processing, however, the luminance-chrominance color system is more efficient and, hence, widely used.
    - An example: histogram equalization
    - Luminance-chrominance representation agrees more with the color perception of the HVS.

- HVS is more sensitive to green than to red, and is least sensitive to blue. An equal representation of RGB leads to inefficient data representation.
- Several different luminance-chrominance color models: HSI, YUV, YCbCr and YIQ.

HSI Model

- I: stands for the intensity component,
- H: for the hue component,
- S: for saturation.
- This model is closely related to the way the HVS perceives color pictures.
- Main drawback: complicated conversion between RGB and HSI models.
- Because of this complexity, the HSI model is not used in any TV systems.

YUV Model*

- Y: luminance component,
- U and V: two chrominance components.
- $Y = 0.299R + 0.587G + 0.114B.$ (1. 2)

Note that the three weights associated with the three primary colors, R, G and B, are not the same: reflecting different responses of the HVS to the different primary colors.

- U and V, are defined as color differences:

$$U = 0.492(B - Y), \quad \text{and}$$
$$V = 0.877(R - Y).$$

- Thus, YUV model lowers computational complexity. Used in **PAL** (Phase Alternating Line) TV systems.
  - ➢Pal: an analog composite color TV standard, used in most of European

  countries, some Asian countries, and Australia.

  - Composite systems: both luminance and chrominance components of TV signals are multiplexed within the same channel.

YIQ Model

- Utilized in **NTSC** (National Television Systems Committee) TV systems.
  - NTSC: an analog composite color TV standard used in North America, Japan.
- Y component: the luminance.
- Two chrominance components:

$$I = -0.545U + 0.839V, \qquad \text{and}$$

$$Q = 0.839U + 0.545V.$$

YDbDr Model

- Used in **SECAM** (Sequential Couleur a Memoire) TV system.
  - SECAM: used in France, Russia, and some eastern European countries.

$$Db = 3.059U, \qquad \text{and}$$
$$Dr = -2.169V.$$

YCbCr Model*

- The chrominance components U and V: differences between B and Y, and R and Y, respectively.

- The chrominance component pairs I and Q, and Db and Dr: both linear transforms of U and V.

- Noted that <u>U and V may be negative</u>.

❖ To make U and V nonnegative, the Y, U and Y are scaled and shifted to produce the YCbCr model.

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix}$$

(1. 3)

❖ Used in **JPEG** and **MPEG**.

1.2.2.6 Color Masking and Its Application in Video Compression

- The HVS is **much more sensitive to the luminance component** than to the chrominance components**.**
- To quantitatively illustrate the above statement: A modified version is shown in Figure 1.10.
- Two observations:
  - ▪ First, for each of the three curves, contrast sensitivity increases when spatial frequency increases, agreeing with frequency masking.
  - ▪ Second, for the same contrast sensitivity, the luminance component corresponds to a much higher spatial frequency. This is an indication that the HVS is much more sensitive to luminance than to chrominance.

    Alternatively, examining those spatial frequencies at which all three curves have data available. Then we can see that the contrast sensitivity of luminance is much lower than that of the chrominance components.

- The direct impact of color masking on image and video coding: allocate more bits to the luminance component than to the chrominance components.
    - A common practice: using full resolution for luminance component, while using a 2 by 1 subsampling both horizontally and vertically for chrominance components.

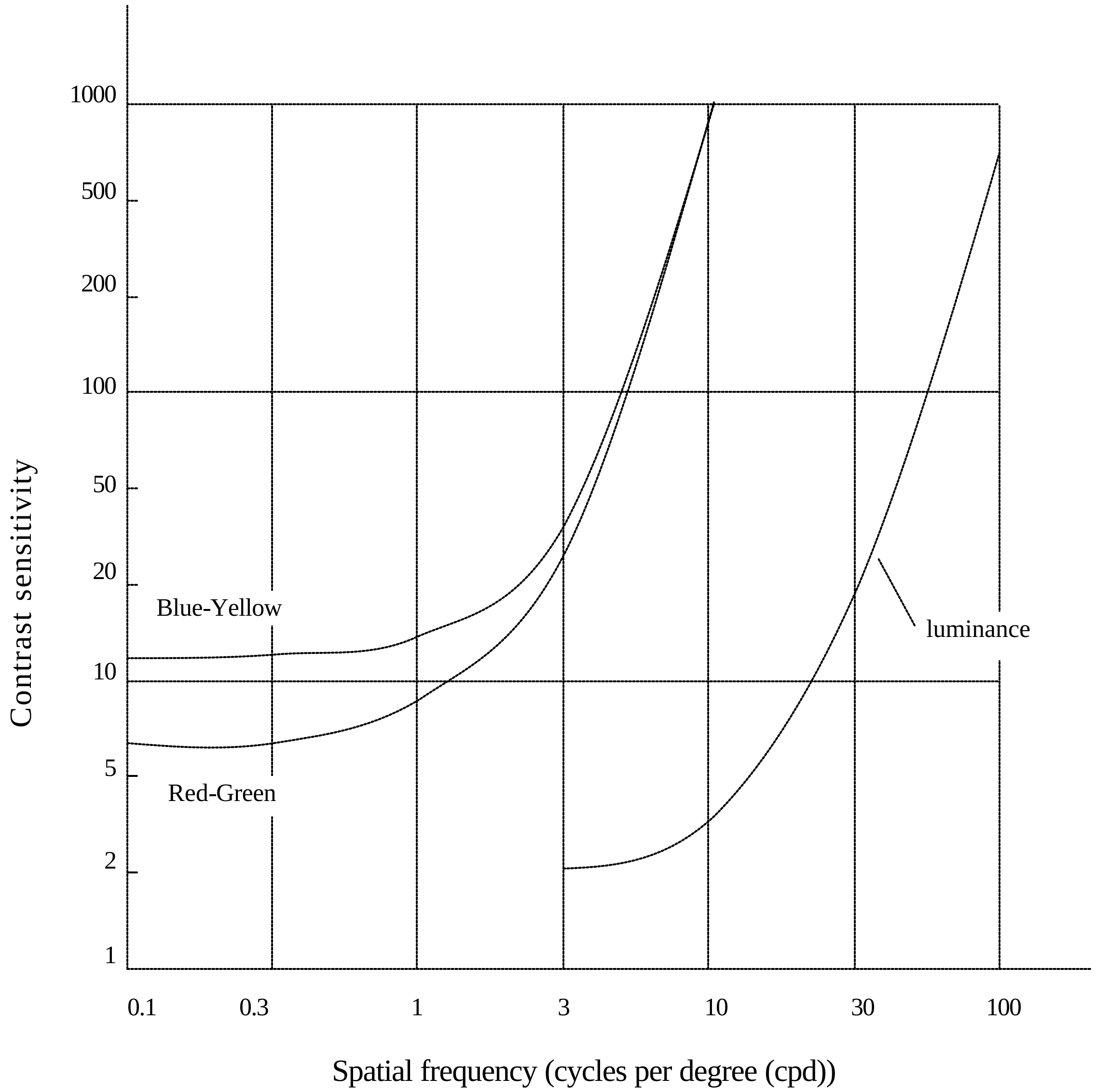


Figure 1. 10 Contrast sensitivity vs. spatial frequency (revised from [van ness 1967], [mullen 1985])

1.2.2.7 Summary: Differential Sensitivity

In this subsection, we discussed

- ❑ luminance masking,
- ❑ texture masking,
- ❑ frequency masking,
- ❑ temporal masking,
- ❑ color masking.

- Luminance masking: of fundamental importance among several types of masking.
- It states that the sensitivity of the eyes to a stimulus depends on the intensity of another stimulus.

  ⇒ a differential sensitivity.
- Both <u>texture</u> and <u>frequency</u> of another stimulus significantly influence this differential sensitivity.

- The same mechanism exists in color perception.
- Conclusion: **differential sensitivity** is the key in studying human visual perception.
- It is common to human perception. For instance, there is also forward and backward temporal masking in human audio perception.

1.2.3 Coding Redundancy

- Interpixel redundancy: the correlation between pixels. That is, some information associated with pixels is redundant.
- Psychovisual redundancy: the information that is psychovisually redundant.
- ⇒ Both statistical and psychovisual redundancies are somehow associated with some information contained in image and video.

- In this sense, the third type of redundancy: coding redundancy, is different.
- Not with information redundancy, but with the representation of information, i.e., coding itself.
- To see this, let us take a look at the following example.

Table 1. 1 An illustrative example

symbol	occurrence probability	code 1	code 2
a_1	**0.1**	**000**	**0000**
a_2	**0.2**	**001**	**01**
a_3	**0.5**	**010**	**1**
a_4	**0.05**	**011**	**0001**
a_5	**0.15**	**100**	**001**

- ➢ The 3rd column: **natural binary code**
- ➢ The 4th column: a **variable-length code**.
  - ✓ Noted that the symbol with a higher occurrence probability is encoded with a shorter length.

- Examine efficiency: which coder provides a shorter average length of codewords.

- Obvious: $L_{avg,1}=3$.

  $L_{avg,2}$ can be calculated as follows.

$$L_{avg,2} = 4\times 0.1 + 2\times 0.2 + 1\times 0.5 + 4\times 0.05 + 3\times 0.15 = 1.95 \text{ bits/symbol.} \tag{1. 4}$$

- From this example, for the same set of symbols, different codes may perform differently. Code 1 contains some redundancy.
- Huffman coding and arithmetic coding, two variable-length coding techniques, will be discussed in Chapter 5.

1.3 Visual Quality Measurement

- An important factor in visual compression.
- We have to base the evaluation of different coding methods on some definite image and video quality.
  - With the same quality of reconstructed image and video, the one that requires less data: superior.
  - With the same amount of data, the method providing a higher quality reconstructed image or video: better.
  - (Other performance criteria, such as computational complexity not considered.)
- Surprisingly, however, it turns out that the measurement of image and video quality is not straightforward.
- Two types of visual quality assessment:
  - Objective assessment
  - Subjective assessment

- A combination of these two methods is now widely utilized in practice.

1.3.1 Subjective Quality Measurement

- Visual quality **should be judged by human viewers** if they are to be the ultimate receivers of the data.
- In subjective visual quality measurement, a set of video frames is generated with varying coding parameters.
- Observers are invited to subjectively evaluate the visual quality of these frames.
  - ➢Observers rate the picture quality.
  - ➢Or, observers measure picture impairment
- **A five-scale rating system** of the degree of impairment, used by Bell Laboratories, is listed below [sakrison 1979].
  - ➢adopted as one of the standard scales in CCIR (now, ITU) Recommendation 500-3 [CCIR 1986].

- ✓Impairment is not noticeable.
- ✓Impairment is just noticeable.
- ✓Impairment is definitely noticeable, but not objectionable.
- ✓Impairment is objectionable
- ✓Impairment is extremely objectionable.

- ♦ In most applications there is a whole array of pictures simultaneously available for evaluation. These pictures are generated with different encoding parameters. By keeping some parameters fixed while making one parameter (or a subset of parameters) free to change, the resulting quality rating can be used to study **the effect of the one parameter (or the subset of parameters) on encoding.**
  - ➢An example: Effect of varying numbers of quantization levels on image quality [gonzalez 1992].
- ♦ Another possible way to study: identify pictures with the same subjective quality measure from the whole array of pictures.

From this subset of test pictures, produce, in the encoding parameter space, **isopreference curves** that can be used to study the effect of the parameter(s) under investigation.

- ➢An example: Effect of varying both image resolution and numbers of quantization levels on image quality [huang 1965].

- **Pairwise comparison,** because it is relatively easy for the eyes.

- Subjective evaluation is **costly**.
  - ➢ Needs a large number of pictures and observers.
  - ➢ Evaluation takes a long time because human eyes are easily fatigued and bored.
  - ➢Some special measures have to be taken in order to arrive at an accurate subjective quality measure. E.g., averaging subjective ratings and taking their deviation into consideration.

1.3.2 Objective Quality Measurement

1.3.2.1 Signal to Noise Ratio

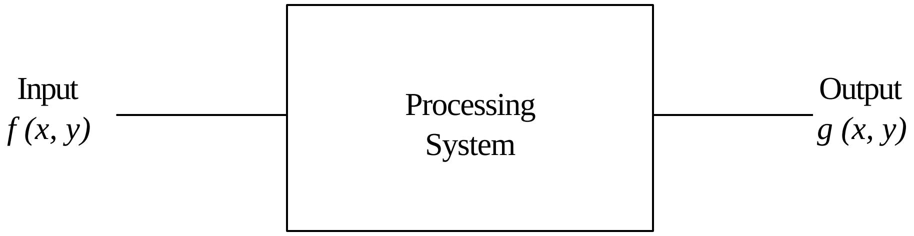


Figure 1.11 An image processing system

- Define an **error function** $e(x, y)$:

$$e(x, y) = f(x, y) - g(x, y).$$

- Mean square error, E_{ms} (MSE):

$$E_{ms} = MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} e(x, y)^2$$

- Root mean square error E_{rms} $(RMSE)$:

$$E_{rms} = RMSE = \sqrt{E_{ms}}.$$

- ❖ **SNR** widely used in objective quality measurement.
- ❖ Mean square signal to noise ratio, SNR_{ms}:

$$SNR_{ms} = 10\log_{10}\left(\frac{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} g(x,y)^2}{MN \cdot E_{ms}} \right),$$

- ❖ Root mean square SNR_{rms}:

$$SNR_{rms} = \sqrt{SNR_{ms}}.$$

- ♦ **PSNR** (peak signal to noise ratio), essentially a modified version of SNR_{ms}, is widely used. $PSNR = 10\log_{10}\left(\frac{255^2}{E_{ms}}\right)$

- The interpretation of the *SNR:*
  - ➢The larger the SNR (SNR_{ms}, SNR_{rms}, or $PSNR$), the better the quality of the processed image, $g(x,y)$.
  - ➢That is, the closer the processed image *g(x,y)* is to the original image *f(x, y)*.
- This seems correct; however, we know that the HVS does not respond to visual stimuli in a straightforward way.
  - ➢Its low-level processing unit is known to be nonlinear.
  - ➢Several masking phenomena exist.
  - ➢It is worth noting that our understanding of the high-level processing unit of the HVS is far from complete.
- Therefore, *SNR* does not always provide with reliable quality assessments.
- One example in Section 1.2.2.3, which uses the IGS quantization technique to achieve high compression (using only four bits for quantization instead of the usual

eight bits) without introducing noticeable false contouring.

In this case, the subjective quality is high, and the *SNR* decreases due to low frequency quantization noise and additive high frequency random noise.

- Other examples.
    - ➢Some additive noise in bright areas or in highly textured regions may be masked.
    - ➢Some minor artifacts in dark and uniform regions may turn out to be quite annoying.
    - ➢In these cases, the *SNR* cannot truthfully reflect visual quality.
- On the one hand, the objective quality measure does not always provide reliable picture quality assessment.
- On the other hand, however, its implementation is much **faster and easier** than that of the subjective quality measure.

Furthermore, objective assessment is repeatable.

- Owing to these merits, objective quality assessment is still widely used.

1.3.2.2 An Objective Quality Measure Based on Human Visual Perception

- A new development in visual quality assessment: [webster 1993]
- Being objective assessment, hence, merits of repeatability, fast and easy implementation.
- Because based on human visual perception, its assessment of visual quality agrees closely to that of subjective assessment.
- ⇒ new method attempts to combine merits of the two different types of assessment.

Methodology

- Figure 1.12.
- The degradation block includes various video compression codecs.
- 48 evaluators in subjective assessment
- Examples of statistical operation:
  - ✓Sobel filtering
  - ✓Laplacian operator
  - ✓first-order differencing
  - ✓moment calculation
  - ✓fast Fourier transform, etc.

- An objective assessment is formed as follows.

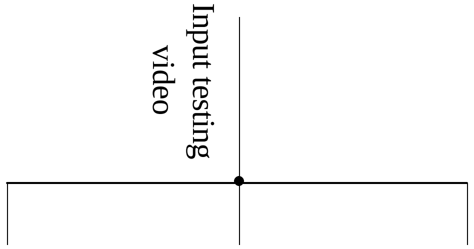
$$\hat{s} = a_0 + \sum_{i=1}^{l} a_i n_i ,$$

$\hat{s}$: the output objective measure

n_i: selected **objective measurements**

a_i: **coefficients** in the linear model

- A statistical measurement is regarded to be good if it can reduce by a significant amount the difference between the objective assessment and the subjective assessment.
- The best measurement is determined via an exhaustive search among the various measurements.
- The coefficients are treated in a similar manner.



Formulation of the Estimator

- Two selected features:
  - ✓Perceived spatial information (SI)

$$SI(f_n) = STD_s\{Sobel(f_n)\},$$

STD_s: standard deviation operator

Sobel: Sobel operation

f_n: the *n*th video frame.

  - ✓Perceived temporal information (TI)

$$TI(f_n) = STD_s\{\Delta f_n\},$$

$\Delta f_n = f_n - f_{n-1}$ (frame difference)

- Determined measurements (n_i)

Parameter l =3 in Equation 1.21, i.e.,

$$\hat{s} = a_0 + a_1 n_1 + a_2 n_2 + a_3 n_3.$$

Measurements are formulated based on SI and TI.

- Determine coefficients (a_i)

least square error procedure applied

- Final objective assessment $\hat{s}$

$$\hat{s} = 4.77 - 0.992 n_1 - 0.272 n_2 - 0.356 n_3 .$$

- Reported Experimental Results
  - Correlation coefficient between subjective assessment score and objective assessment score (an estimate of subjective score) is in the range of 0.92 to 0.94.
  - (A set of 36 testing scenes containing various amounts of spatial and temporal information was experimented, implying quite good performance)
  - This work does open a new and promising way to assess visual quality by combining subjective and objective approaches.
  - Theoretically, SI and TI are important. They reflect the most important aspect of human visual perception.

1.4 Information Theory Results

1.4.1 Entropy

Entropy is a very important concept in information theory and communications. So is it in image and video compression.

1.4.1.1 Information Content of a Source Symbol

- Carriers of information: symbols.
- Consider: a symbol with an occurrence probability p.
- Its information content, I, is:

$$I = \log_2 \frac{1}{p} \quad bits \quad or \quad I = -\log_2 p \quad bits,$$

bit : a contraction of *binary unit*.

- Results can be converted for **the r-ary case**

$$I = -\log_r 2 \cdot \log_2 p \qquad bits.$$

- The smaller the probability, the more information the symbol contains.
- The occurrence probability is somewhat related to the **uncertainty** of the symbol.

1.4.1.2 Average Information per Symbol

- A discrete memoriless information source:

$$\{s_i, i = 1, 2, \cdots, m\}.$$

  - ✓Corresponding occurrence probabilities are denoted by $\{p_i, i = 1, 2, \cdots, m\}$.
  - ✓Discreteness: a countable set of symbols.
  - ✓Memoriless: the occurrence of a symbol in the set is independent of that of its preceding symbol.

- The information content of a symbol s_i, I_i, is equal to $I_i = -\log_2 p_i$ bits.

- Entropy, H: the average information content per symbol of the source

$$H = -\sum_{i=1}^{m} p_i \log_2 p_i \qquad bits.$$

- Straightforward to show: entropy reaches the maximum when all symbols in the set are equally probable.

1.4.2 Shannon's Noiseless Source Coding Theorem

- Consider a discrete, memoriless, stationary information source.
- In source encoding, a *codeword* is assigned to each symbol in the source.
- The number of bits in the codeword is referred to as the length of the codeword.
- The average length of codewords is referred to as bit rate, expressed in the unit of bits per symbol.

- Shannon’s noiseless source coding theorem states that for a discrete, memoriless, stationary information source, **the minimum bit rate required** to encode a symbol on average is equal to the entropy of the source.
- This theorem provides us with a lower bound in source coding.
- Shannon showed that the **lower bound** can be achieved when the *encoding delay* extends to infinity.

  By encoding delay, we mean the encoder waits and then encodes a certain number of symbols at once.
- Fortunately, with finite encoding delay, we can already achieve an average codeword length fairly close to the entropy.
- That is, we do not have to actually sacrifice bit rate much to avoid long encoding delay, which involves high computational complexity and a large amount of memory space.

- Stationarity assumption is necessary in deriving the noiseless source coding theorem.

- This assumption may not be satisfied in practice. Hence, Shannon’s theorem is a theoretical guideline only.

- One way to evaluate the efficiency of a coding scheme: *efficiency* h

  $$h = \frac{H}{L_{avg}},$$

  where H is entropy, and L_{avg} denotes the average length of the codewords in the code. $h \le 1$.

- The above can be generalized to calculate the relative efficiency between two codes.

  $$h = \frac{L_{avg,1}}{L_{avg,2}},$$

  $L_{avg,1}$ and $L_{avg,2}$: the average codeword length for code 1 and code 2.

- A complementary parameter of coding efficiency: *redundancy*, z

$$z = 1 - h.$$

1.4.3 Shannon's Noisy Channel Coding Theorem

- Consider a noisy transmission channel.
  - ✓Received symbols may be erroneous due to the lack of redundancy.
  - ✓Adding redundancy may combat noise.
- Shannon's noisy channel coding theorem states that it is possible to transmit symbols over a noisy channel without error if the bit rate, R, is below a *channel capacity*, C.

$$R < C$$

The channel capacity is determined by the noise and signal power.

- The noisy channel coding theorem, Shannon's second theorem [shannon 1948], is concerned with **a noisy, memoriless channel**.
- By memoriless, we mean the channel output corresponding to the current input is independent of the output corresponding to previous input symbols.
- Channel capacity sets **an upper bound** on the bit rate.

1.4.4 Shannon's Source Coding Theorem

- Now continue to deal with discrete memoriless information sources, but we discuss the situation in which lossy coding is encountered.

- The source coding theorem [shannon 1948] states that for a given distortion D, there exists a rate distortion function $R(D)$ [berger 1971], which is the minimum bit rate required to transmit the source with distortion less than or equal to D.
- That is, in order to have distortion not larger than D, the bit rate, R, must satisfy:

$$R \geq R(D).$$

1.4.5 Information Transmission Theorem

- Combining the noisy channel coding theorem and the source coding theorem, we have the information transmission theorem [slepian 1973].

$$C \geq R(D)$$

- It states that if the channel capacity of a noisy channel, *C*, is larger than the rate distortion function *R(D)*, then it is possible to transmit an information source with distortion *D* over a noisy channel.

References

[berger 1971] T. Berger, *Rate Distortion Theory*, Englewood Cliffs, NJ, Prentice-Hall, 1971.

[CCIR 1986] CCIR Recommendation 500-3, "Method for the subjective assessment of the quality of television pictures," Recommendations and Reports of the CCIR, 1986, XVIth Plenary Assembly, Volume XI, Part 1.

[connor 1972] D. J. Connor, R. C. Brainard and J. O. Limb, ''Interframe coding for picture transmission,'' *Proceedings of The IEEE*, vol. 60, no. 7, pp. 779-790, July 1972.

[fink 1957] D. G. Fink, *Television Engineering Handbook*, New York, McGraw-Hill 1957, Sect. 10.7.

[goodall 1951] W. M. Goodall, ''Television by pulse code modulation,'' Bell System Technical Journal, pp. 33-49, January 1951.

[gonzalez 1992] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison Wesley, 1992.

[hidaka 1990] T. Hidaka and K. Ozawa, ‘‘Subjective assessment of redundancy-reduced moving images for interactive application: test methodology and report,’’ *Signal Processing: Image Communication*, 2, pp. 201-219, 1990.

[huang 1965] T. S. Huang, ‘‘PCM picture transmission,’’ *IEEE Spectrum*, vol. 2, no. 12, pp. 57-63, 1965.

[huang 1998] J. Huang and Y. Q. Shi, ‘‘Adaptive image watermarking scheme based on visual masking,’’ *IEE Electronics Letters*, vol. 34, no. 8, pp. 748-750, April 1998.

[lathi 1998] B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3rd edition, Oxford University Press, New York, Oxford, 1998.

[legge 1980] G. E. Legge and J. M. Foley, ‘‘Contrast masking in human vision,’’ *J. Opt. Soc. Am*., vol. 70, no. 12, pp. 1458-1471, December 1980.

[lim 1990] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice Hall, 1990.

[kretzmer 1952] E. R. Kretzmer, "Statistics of television signal," *Bell System Technical Journal*, vol. 31, no. 4, pp. 751-763, July 1952.

[mounts 1969] F. W. Mounts, "A video encoding system with conditional picture-element replenishment," *Bell System Technical Journal*, vol. 48, no. 7, pp. 2545-2554, September 1969.

[mullen 1985] K.T. Mullen, ‘‘The contrast sensitivity of human color vision to red-green and blue-yellow chromatic gratings,’’ *J. Physiol*., 359: 381-400, 1985.

[netravali 1977] A. N. Netravali and B. Prasada, “Adaptive quantization of picture signals using spatial masking,” *Proceedings of The IEEE*, vol. 65, pp. 536-548, Apr. 1977.

[sakrison 1979] D. J. Sakrison, ‘‘Image coding applications of vision model,’’ in *Image Transmission Techniques*, W. K. Pratt (Ed.), pp. 21-71, Academic Press, 1079.

[seyler 1962] A. J. Seyler, "The coding of visual signals to reduce channel-capacity requirements," *The Institution of Electrical Engineers Monograph*, no. 533E, July 1962.

[seyler 1965] A. J. Seyler, "Probability distributions of television frame difference," *Proceedings of IREE*, (Australia), 26, p. 335, November 1965.

[shannon 1948] C. E. Shannon, ‘‘A mathematical theory of communication,’’ *Bell System Technical Journal*, vol. 27, pp. 379-423 (Part I), July 1948, pp. 623-656 (Part II), October 1948.

[slepian 1973] *Key Papers in the Development of Information Theory*, D. Slepian (Ed.), IEEE Press, New York, 1973.

[van ness 1967] F. I. Van Ness and M. A. Bouman, ‘‘Spatial modulation transfer in the human eye,’’ *J. Opt. Soc. Am.* 57(3): 401-406, March 1967.

[watson 1987] A. B. Watson, ‘‘Efficiency of a model human image code,’’ *J. Opt. Soc. Am. A*, vol. 4, no. 12, pp. 2401-2417, December 1987.

[webster 1993] A. A. Webster, C. T. Jones and M. H. Pinson, ‘‘An objective video quality assessment system based on human perception,’’ *Proceedings of Human Vision, Visual Processing and Digital Display IV*, J. P. Allebach and B. E. Rogowitz (Ed.s), SPIE, vol. 1913, pp. 15-26, September 1993.