Introduction to Image Compression

Computer Science Department
CS4481b/9628b: Image Compression
Winter 2017
Instructor: Mahmoud R. El-Sakka
Office: MC-419
Email: elsakka@csd.uwo.ca
Phone: 519-661-2111 x86996

Compression, What Does It Mean?

- In real life, compression means making things smaller by applying pressure
- Image compression is not about physically squashing images, but about finding ways to represent it in fewer bytes
- From this point of view, compression can be defined as a *process intended to yield a compact representation of a given object*
- The *objective* of image compression is to *achieve a low bit rate in the digital representation of an input image*, i.e., *compact digital image representation*, *with no, or at most minimal, perceived loss of picture quality*

Why Compression?

- Since the very beginning of digital image processing in the 1950's, image compression has been recognized as an important field
  - Due to the large amount of data which needs to be handled for ***transmission***, or ***storage***, of digital images
- To appreciate the need for image compression, consider a typical true-color digital image of size 1024 × 1024 (1 M. Pixel)
  - Such an image requires at least 3 MB of storage space
  - To ***upload*** this uncompressed image over a 1 Mbits/second modem, it would take ***at least*** 24 seconds
  - With compression, these requirements can be dramatically reduced
- Now think of a 12 Mega-Pixels images (3072 × 4096)!!
- I needless to mention uncompressed movies!!

Why Compression?

- Image compression has been, and continued to be, crucial to many important and diverse applications, including:
  - *Image communication applications*, such as
    - Facsimile machines
    - Digital TV
    - Video conferencing
  - *Image retrieval applications*, such as
    - Real estate
    - Security
    - Medical imaging
  - *Remote sensing applications*, such as
    - the use of satellite imagery for
      - Weather
      - Earth-resources

Why Compression?

- All the above mentioned applications require ***transmitting*** and ***storing*** a huge amount of *image information*, and hence image compression is needed
- Even though technology has made enormous progress in computer storage capacity, the interest in compression is still an important present-day research issue, and it is likely to stay as such for years to come
- To draw an analogy, no matter how much money you earn, you would probably hesitate to turn away your nose at the chance of tripling your salary!!!!

Data Versus Information

- ***Data*** and ***information*** are not synonymous
- Data is the means by which information is conveyed
- Various amounts of data may be used to represent the same amount of information
- Think of
  - ***data*** as raw material
  - ***information*** as final product

Data Versus Information

- Example:
  Two individuals may tell the same story by using different words
  - Words here are the *data* used to convey the *information*
  - If the used words in these two versions are not identical, then
    - Two different versions of the same basic story are created
    - One of them may include *unessential data*
    - In other words, one version may contain *redundant data* that either
      - Provides no relevant information, or
      - Restates things which is already known

Data Redundancies

- Data redundancy is a central issue in digital image compression
- In digital image compression, *three basic data redundancies* can be identified and exploited
  - *Psychovisual* redundancy
  - *Encoding* redundancy
  - *Inter-pixel* (a.k.a. *spatial*) redundancy

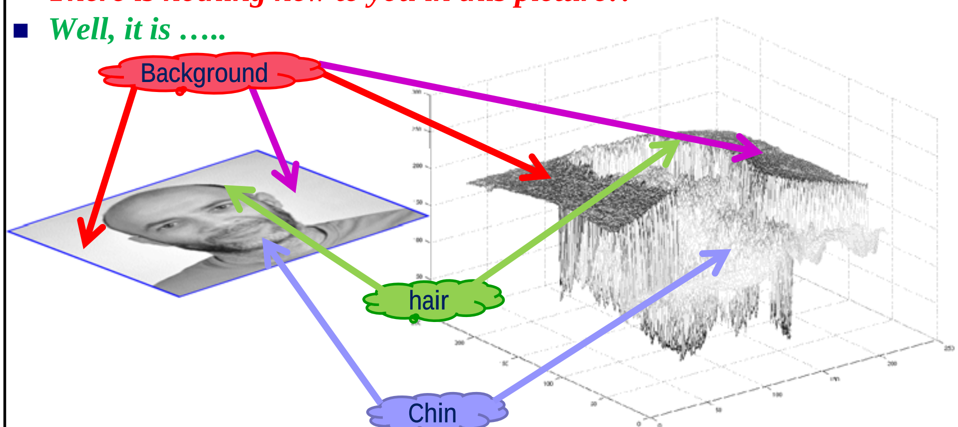
- Image compression is achieved when one, or more, of these redundancies are reduced

Psychovisual Redundancy

- The *human visual system* (HVS) does not respond with equal sensitivity to all visual information
- In a normal visual processing, certain information has less, or no, relative importance than other information
- This information can be eliminated without significantly undermining the quality of image perception
- Note that:
  - The human perception of image information normally does not involve a quantitative analysis of every pixel
  - In general, an observer searches for distinguishing features, such as edges or textural regions, and mentally combines them into recognizable groups
  - The brain then correlates these groups with a prior knowledge in order to complete the image interpretation process

Psychovisual Redundancy

- However, we need not to forget that a digital image is just an array of numbers (*human vision* vs *computer vision*)
- ***Can you tell me what is in this image?***
- ***There is nothing new to you in this picture!!***
- ***Well, it is …..***



Psychovisual Redundancy

- Reducing the *psychovisual redundancy* is associated with *eliminating* real, or quantifiable, visual information
- The reduction of the psychovisual redundancy is possible only because the information itself is not essential for the *normal* visual processing
- Since the reduction of the psychovisually redundant information results in a *loss of quantitative information*, it is commonly referred to as *quantization*
- This terminology is consistent with the normal usage of the word, which generally means *many-to-one* mapping
- Quantization results in loss of visual information, i.e., *irreversible*, or *lossy*

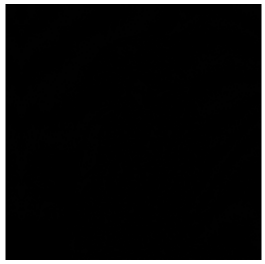
Psychovisual Redundancy

- Example
  Consider the following 256 × 256 *bird* gray-scale image

Bird image 256 × 256

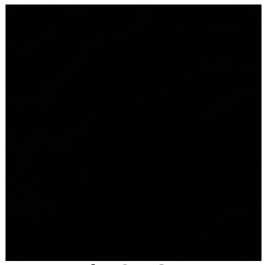
Psychovisual Redundancy

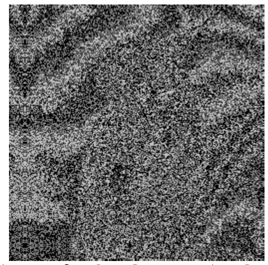

7-MSB bird image


original – 7-MSB image

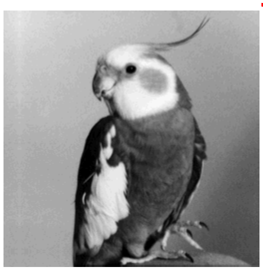

(original – 7-MSB image) * 128

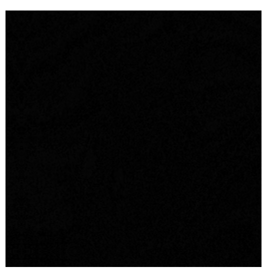

6-MSB bird image

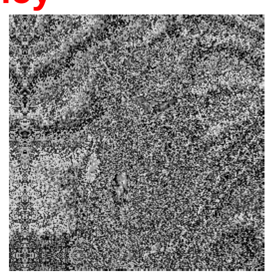

original – 6-MSB image

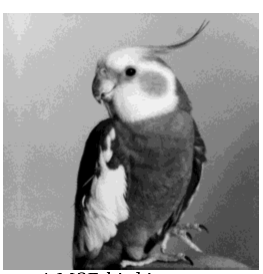

(original – 6-MSB image) * 64

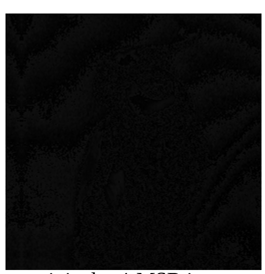
Psychovisual Redundancy


5-MSB bird image


original – 5-MSB image


(original – 5-MSB image) * 32


4-MSB bird image


original – 4-MSB image


(original – 4-MSB image) * 16

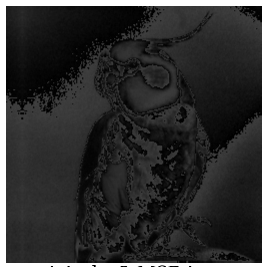
Psychovisual Redundancy

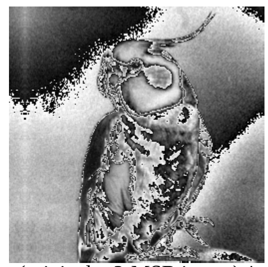

3-MSB bird image


original – 3-MSB image


(original – 3-MSB image) * 8

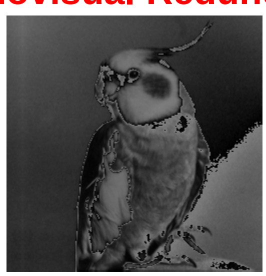

2-MSB bird image


original – 2-MSB image


(original – 2-MSB image) * 4

Psychovisual Redundancy


1-MSB bird image


original – 1-MSB image

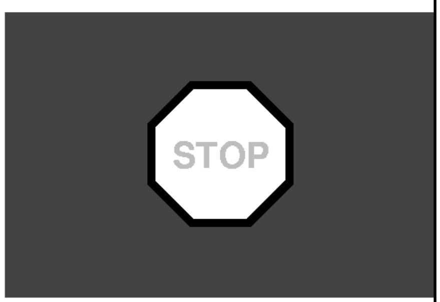

(original – 1-MSB image) * 2

Encoding Redundancy

- A *code is a system of symbols*, e.g., letters, numbers, bits, used to represent a body of information
- Each piece of information is assigned a *sequence of code-symbols*, called a *code-word*
- The length of a *code-word* is the *number of symbols in it*
- One *goal* of compression schemes is *to find a code which reduces the amount of code-symbols needed to represent code-words*
- *The shorter the average code-word length is, the higher the compression accomplished*
- The process of reducing encoding redundancy data is *reversible*, i.e., you can restore the original data back, i.e., *lossless*

Encoding Redundancy

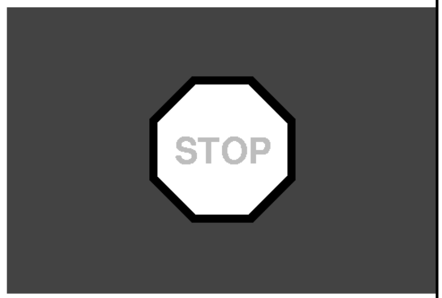
- The 512 × 400 *stop sign* image
  - Is a gray-scale image
  - Needs 512 × 400 bytes (200 KB) to represent its pixel values



Encoding Redundancy

- The image has only 4 gray-scale values, which are 0, 67, 189, and 255
- Pixel values can be represented as follows:
  - "0" to be represented by $(00)_2$
  - "67" to be represented by $(01)_2$
  - "189" to be represented by $(10)_2$
  - "255" to be represented by $(11)_2$

  in this case, only 512 × 400 × 2/8 bytes (50 KB) are needed



- Pixel values can also be represented as follows:
  - "0" to be represented by $(000)_2$ (8,156 cases)
  - "67" to be represented by $(1)_2$ (169,320 cases)
  - "189" to be represented by $(001)_2$ (2,567 cases)
  - "255" to be represented by $(01)_2$ (24,757 cases)

  in this case, only 8,156 × 3/8 + 169,320 × 1/8 + 2,567 × 3/8 + 24,757 × 2/8 bytes (30.64 KB) are needed

Inter-pixel Redundancy

- In most images, adjacent pixels are highly *correlated*
- As a result of this fact, the value of a given pixel can be reasonably *predicted* (*guessed*) from the values of its neighbor pixels
- The information carried by an individual pixel, giving the value of its neighbor pixels, is relatively small
- To reduce the inter-pixel redundancy in a given image, its pixel values array must be *transformed*, or *mapped*, into a more efficient (usually "*non-visual*") domain
- The process of reducing inter-pixel redundancy data is *reversible*, i.e., you can restore the original data back, i.e., *lossless*

Interpixel Redundancy

- The difference between adjacent pixels can be used to represent an image; by applying this scheme to the *stop sign* image using the following code:
  - “0–255” to be represented by $(0000)_2$ (178 cases)
  - “255–0” to be represented by $(0001)_2$ (178 cases)
  - “0–67” to be represented by $(0010)_2$ (202 cases)
  - “67–0” to be represented by $(0011)_2$ (202 cases)
  - “189–255” to be represented by $(010)_2$ (208 cases)
  - “255–189” to be represented by $(011)_2$ (208 cases)
  - “255–255”, “189–189”, “67–67”, or “0–0” to be represented by $(1)_2$ (203,623 cases)

in this case, we only need
1 (to encode the first pixel in the image)
$+ 178 \times 4/8 + 178 \times 4/8 + 202 \times 4/8 + 202 \times 4/8$
$+ 208 \times 3/8 + 208 \times 3/8 + 203{,}623 \times 1/8 = 25.38$ KB

This achievement is due to the reduction in interpixel redundancy and encoding redundancy.

Lossless Versus Lossy Compression

- Image compression schemes can be classified as
  - *Lossless* schemes
    - Compressing an image and expanding it again produces an image which is bit-by-bit *identical* to the original image
    - All the information is preserved
    - No quantization is applied
  - *Lossy* schemes
    - Compressing an image and expanding it again produces an image which is *close* to the original image, i.e., it is not an exact match
    - Some of the psychovisual redundant data might be lost
    - Quantization is applied

- The reason for using lossy compression schemes is that it generally gives significantly greater compression than lossless schemes do
- In many situations, small losses of data are acceptable, as long as a higher compression level will be achieved

Fidelity Criteria

- The reduction of psychovisual redundant information results in a loss of real, or quantitative, visual information
- Because information of interest may be lost, a mean of quantifying the nature of information loss is highly desirable
- To do so, two general classes of criteria are used as the basis for such an assessment
  - *Subjective* fidelity criteria
  - *Objective* fidelity criteria

Subjective Fidelity Criteria

- Most decompressed images are ultimately viewed by human beings
- Although *objective fidelity criteria* offer a simple, and convenient, mechanism for evaluating *data loss*, measuring image quality by the *subjective evaluations of human observers* is often more appropriate
- Subjective evaluations can be accomplished by
  - Asking an *appropriate cross-section of viewers* to assess the quality of a reconstructed image
  - Analyzing their evaluations and producing some sort of averaging
- The appropriate cross-section of viewers means a group of people who include
  - *Experts*, to give refined assessment of image quality
  - *Non-experts*, to represent average viewers
- *Viewer evaluation* is a *n-point rating scale* to show how pleased, with the reconstructed image, a viewer is

Subjective Fidelity Criteria

- The table below shows an example of a 5-point rating scale and the meaning of each rating value

Rating	*Rating Value*	*Rating Description*
Extremely high-quality	5	When a viewer is confident that the reconstructed image is as exactly as the original image
High-quality	4	When a viewer sees a pleasant reconstructed image with hardly to recognize degradations
Acceptable-quality	3	When a viewer sees a pleasant reconstructed image but, at the same time, recognizes some minor degradations
Poor-quality	2	When a viewer wishes he/she could improve the reconstructed image
Extremely poor-quality	1	When a viewer hardly recognizes, or even could not recognize, the reconstructed image

Subjective Fidelity Criteria

- Note that the results of subjective rating are affected by many factors, including:
  - The resolution of a viewing device
  - The size of a displayed image
  - The viewing distance
  - The viewers' vision
  - The viewers' mode
  - The viewers' level of expertise
  - The viewers' gender
  - The number of viewers
- If a standard can be established, *and followed*, for these factors, the results obtained at different locations and different times may then become comparable
- In general, subjective evaluation is a time consuming scheme
- Unless viewers are keen and honest, subjective evaluation is a waste of time

Objective Fidelity Criteria

- Objective fidelity criteria are measures attempting to assess the differences between
  - the *original* image and
  - the *compressed-then-decompressed* image
- Good examples for objective fidelity criteria include
  - Mean Absolute Error (*MAE)*
  - Root Mean Squared Error (*RMSE)*
  - Signal-to-Noise-Ratio (*SNR)*
  - Peak Signal-to-Noise-Ratio (*PSNR)*
  - Mean Structural SIMilarity *(MSSIM)* index

Objective Fidelity Criteria

$$MAE = \frac{1}{W \times H} \times \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \left| f(x,y) - \hat{f}(x,y) \right|$$

$$RMSE = \sqrt{\frac{1}{W \times H} \times \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \left(f(x,y) - \hat{f}(x,y) \right)^2}$$

$$SNR = 10 \times \log_{10} \left(\frac{\frac{1}{W \times H} \times \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \hat{f}(x,y)^2}{RMSE^2} \right) dB$$

$$PSNR = 10 \times \log_{10} \left(\frac{255^2}{RMSE^2} \right) = 20 \times \log_{10} \left(\frac{255}{RMSE} \right) dB$$

Where:
$f(x,y)$: *The original image*
$\hat{f}(x,y)$: *compressed-then-decompressed image*
W: Image width
H: Image height

Objective Fidelity Criteria

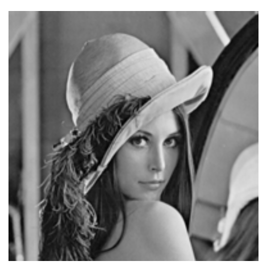
- Example 1: Consider the *bird* image and its quantized versions

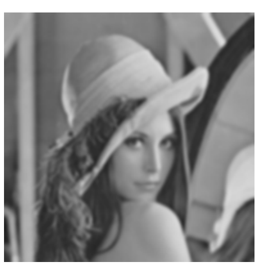
Image	MAE	RMSE	SNR	PSNR
Original bird image	0.00	0.00	∞	∞
7-MSB bird image	0.49	0.70	46.95 dB	51.24 dB
6-MSB bird image	1.48	1.85	38.44 dB	42.78 dB
5-MSB bird image	3.56	4.24	31.14 dB	35.59 dB
4-MSB bird image	7.87	9.13	24.23 dB	28.92 dB
3-MSB bird image	15.39	17.71	18.10 dB	23.17 dB
2-MSB bird image	32.97	37.82	10.48 dB	16.58 dB
1-MSB bird image	62.44	68.95	3.47 dB	11.36 dB

Objective Fidelity Criteria

- Example 2: Consider the lena images below:

(a) (b)





(c) (d)

Objective Fidelity Criteria

(a) The original Lena image

(b) A histogram flattening version of the Lena image
(PSNR = 26.29dB and RMSE = 12.36)

(c) A Gaussian blurred version of the Lena image using a mask of size 15 × 15 with a standard deviation = 2.992
(PSNR = 26.29dB and RMSE = 12.36)

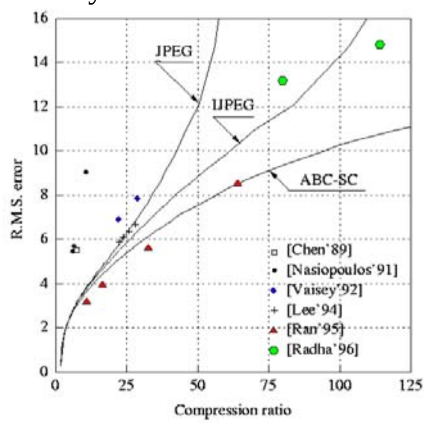
(d) A noisy version of the Lena image using random Gaussian noise
(PSNR = 26.29dB and RMSE = 12.36)

Objective Fidelity Criteria

- Although RMSE, SNR, and PSNR are widely used in lossy compression literature, they are sometime *misleading* and not indicative of the actual loss of fidelity
  - Especially at low values of SNR/PSNR (i.e., at high vales of RMSE)
- Reasons for this *misleading* situation:
  - Objective fidelity criteria ***assess the loss of data***, i.e., the changes in pixel value, ***not the loss of information***

Rate-Distortion Curve

- Usually, plotting the bit-rate (or the compression ratio) against the objective fidelity criteria is useful and meaningful



Comparison of rate distortion results between ABC-SC, IJPEG, JPEG, and some other segmentation-based encoding techniques for Lena image

Mean Structural SIMilarity *(MSSIM)* index

- The **Mean Structural SIMilarity** (MSSIM) index is a method for measuring the similarity between two images
- MSSIM is designed to improve on traditional methods like PSNR and RMSE, which have proved to be inconsistent with human perception

Structural SIMilarity *(SSIM)* index

- The measure between two windows x and y of common size $N \times N$ is:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^{\ 2} + \mu_y^{\ 2} + c_1)(\sigma_x^{\ 2} + \sigma_y^{\ 2} + c_2)}$$

- μ_x the average of x
- μ_y the average of y
- σ_x the variance of x
- σ_y the variance of y
- σ_{xy} the covariance of x and y
- $c_1=(k_1L)^2$ and $c_2=(k_2L)^2$ two variables to stabilize the division with weak denominator
- L the dynamic range of the pixel-values ($2^{\#\text{ of bits per pixel}}$ -1)
- k_1 = 0.01 and k_3 = 0.03 by default (The performance of the SSIM index algorithm is fairly insensitive to variations of these values)

Topic 02: Introduction to Image Compression

Structural SIMilarity *(SSIM)* index

- The measure between two windows x and y of common size $N \times N$ is:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^{\ 2} + \mu_y^{\ 2} + c_1)(\sigma_x^{\ 2} + \sigma_y^{\ 2} + c_2)}$$

- The resultant *SSIM* index is a decimal value between -1 and 1:
  - value 1 is only reachable in the case of two identical sets of data
- Typically it is calculated on window sizes of 8×8

- the Mean SSIM (MSSIM) index is usually utilized to evaluate the overall image quality

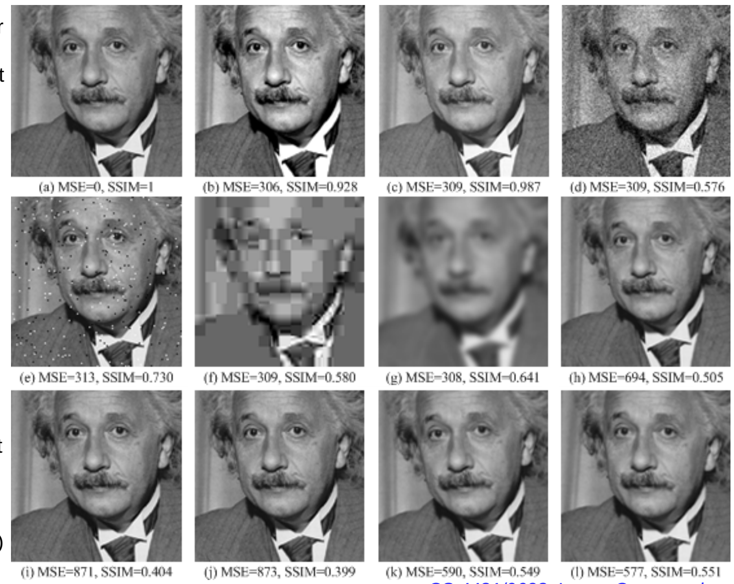
Structural SIMilarity *(SSIM)* index

Comparison of image fidelity measures for "Einstein" image altered with different types of distortions

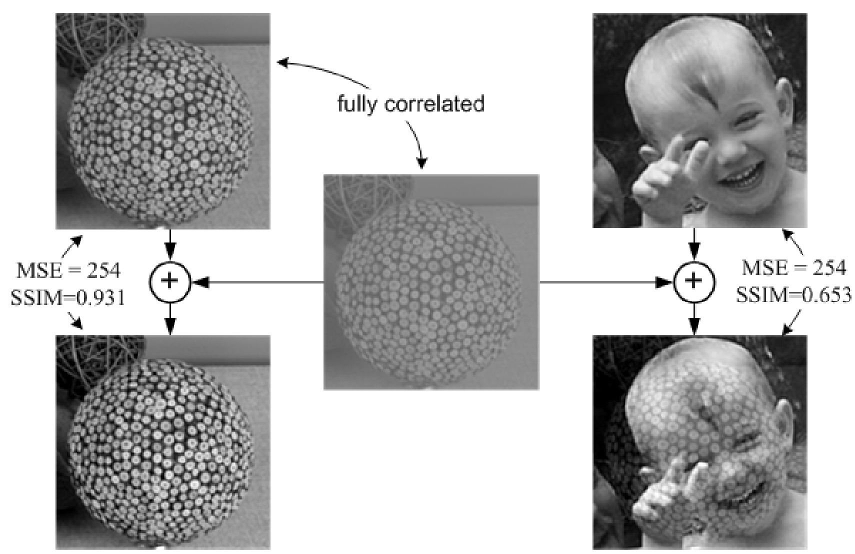
(a) Reference image (b) Mean contrast stretch (c) Luminance shift (d) Gaussian noise contamination

(e) Impulsive noise contamination (f) JPEG compression (g) Blurring (h) Spatial scaling (zooming out)

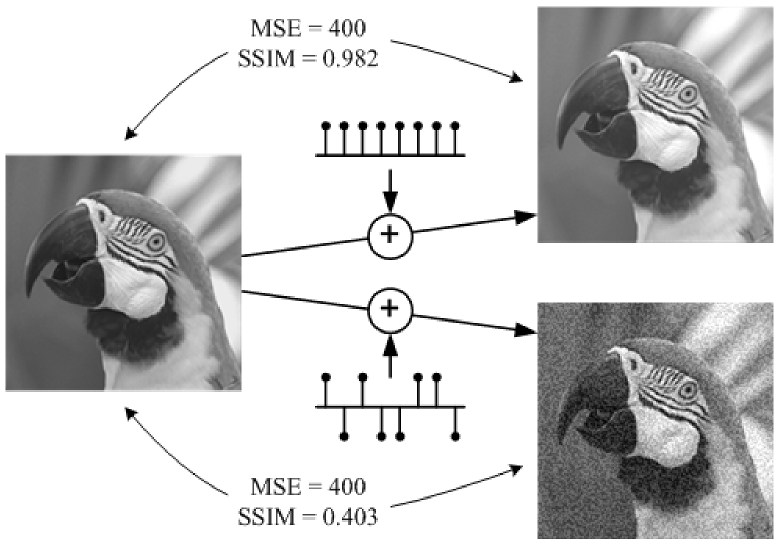
(i) Spatial shift (to the right) (j) Spatial shift (to the left) (k) Rotation (counter-clockwise) (l) Rotation (clockwise)



Structural SIMilarity *(SSIM)* index



Structural SIMilarity *(SSIM)* index



Structural SIMilarity *(SSIM)* index

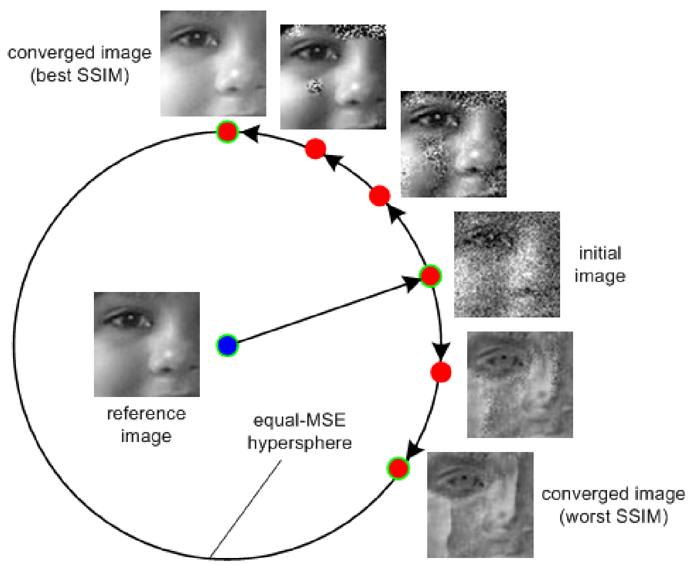
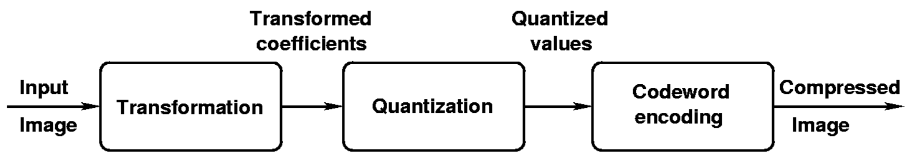


Image Compression Model

- Image compression can be characterized by three independent stages: *transformation* (a.k.a. mapping), *quantization*, and *codeword encoding*

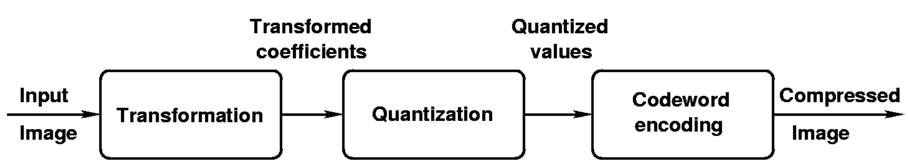


General image compression model

- The transformation stage
  - Reduces inter-pixel redundancy
  - Reversible (one-to-one mapping)
  - Usually, produces a non-visualize format
  - May, or may not, directly reduce the amount of data required to represent the image

Image Compression Model

- Image compression can be characterized by three independent stages: *transformation* (a.k.a. mapping), *quantization*, and *codeword encoding*

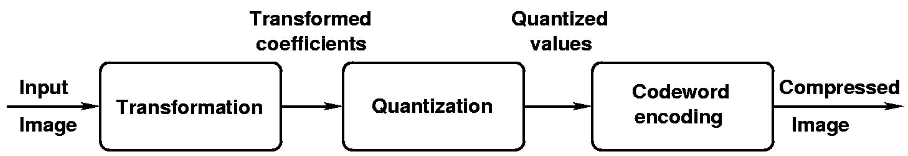


General image compression model

- The quantization stage
  - Reduces psychovisual redundancy
  - Irreversible (many-to-one mapping)
  - Reduces the accuracy of the transformed coefficients in accordance with some pre-established fidelity criterion, or compression level
  - Must be omitted when lossless compression is desired

Image Compression Model

- Image compression can be characterized by three independent stages: *transformation* (a.k.a. mapping), *quantization*, and *codeword encoding*

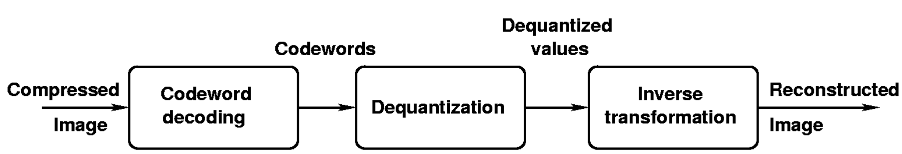


General image compression model

- The codeword encoding stage
  - Reduces encoding redundancy
  - Reversible (one-to-one mapping)
  - Creates a fixed-length, or variable-length, code to represent the quantized values
  - In case of variable-length encoding, the shortest codewords are assigned to the most frequently occurring quantized values, and thus reduces the encoding redundancy

Image Decompression Model

- Image decompression is achieved by *reversing* the effect of the above mentioned three compression stages



General image decompression model