

# Midterm Study Topics & Questions

**You are allowed 2 pages of notes for the test.**

1. What is multimedia? [L01]
  1. Give examples of MM data in an application (combinations of text, sound, images, animation, graphics)
    - a. Data is the recorded facts
      1. So in Instagram the data is the images and using filters we can alter the data
  2. Discuss a multimedia system you use on one of the following platforms: mobile phone, home PC, TV, Social media, streaming video, etc.
    - a. The tv with hundreds of digital channels available with a remote that takes user input
3. Explain three state-of-the-art MM research topics
  - a. Processing and coding
    1. Multimedia content analysis, retrieval, security, audio/image/video processing, compression, etc.
  - b. Multimedia system support and networking
    1. Network protocols, Internet, operating systems, servers and clients, quality of service, and databases
  - c. Multi-modal interaction and integration
    1. "ubiquity" – web-everywhere devices, multimedia education including Computer Supported Collaborative Learning, and design and applications of virtual environments
2. What are Multimedia Information Systems?

Consider their advanced features: supports problem solving (in the information domain), contains large amounts of MM data, interactive or immersive, smart features, intelligent assistance, and user interface (and interaction) modalities. [L01,L13]

  1. Couple data management capabilities with media management (compression, standardization, transmission).
3. Name a multimedia application that focusses on (and explain why):
  1. Data
    - a. Systems that give suggestions or recommendations
  2. Information
    - a. Closed Captions
  3. Knowledge
    - a. Wolfram Alpha
  4. Wisdom
    - a. Genie or prophet
4. What characterizes an interactive multimedia information system?
  1. Computer-delivered electronic system that allows the user to control, combine, and manipulate different types of media.

5. Explain adaptive behavior in machine learning systems.
  1. Machine learning systems are able to read in large amounts of data and adapt by learning from their experience
6. How does multimedia enable augmented, virtual, and extended reality?
  1. Knowing the camera pose enables augmentation, introducing virtual objects in the scene at correct locations, with correct orientation.
7. Explain how we model the web as a directed graph [L02]
  1. Each page is a node with connections made through hyperlinks
8. What is information retrieval?

Explain searching, role of metadata, information content, etc.

1. Is the science of searching for documents, for information within documents, as well as that of searching relational databases and the World Wide Web
9. Which of the following represents “the fraction of relevant instances that are retrieved”?

a. Precision

**b. Recall**

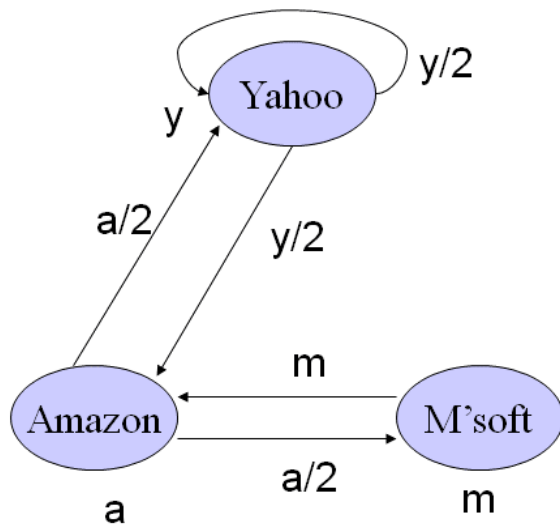
c. Relevance d. Importance

10. In web search which one is more important, precision or recall?
  1. precision
11. What is the difference between web search and information retrieval?
  1. Scale of web search is way beyond information retrieval
  2. Web is very dynamic
  3. Quality of web pages is not uniform
  4. Web is hypertextual
12. What is a web crawler?
  1. Program for downloading web pages
13. Define relevance and importance of web documents in a web query, how does that relate to the

PageRank algorithm?

1. Very important for gaining information
  2. PageRank can then figure out how difficult it is to find this page by scanning the web documents
14. Given the web/link graph below, write the normalized connectivity matrix, the initial values of

$W_0$ , and the formula for the PageRank calculation with damping.



Citing Page (to)		Citing page (from)		
	Yahoo	Amazon	Microsoft	Number
Yahoo	1	1	-	2
Amazon	1	-	1	2
Microsoft	-	1	-	1
Number	2	2	1	-

Citing Page (to)		Citing page (from)	
	Yahoo	Amazon	Microsoft
Yahoo	0.5	0.5	-
Amazon	0.5	-	1
Microsoft	-	0.5	-

$W_0 = 1/3 = 0.33333$  (always equal to  $1/(\text{\# of nodes})$ )

Formula =  $w_k = (1-d)w_0 + dBw_{k-1}$

15. Define NLP and give two examples of NLP applications [L03]

1. Field of computer science to get computers to perform useful tasks involving human language
  1. Speech Recognition
  2. Spelling/grammar/edit correction, auto-completion

16. What is folksonomy, taxonomy, and what is their relevance to NLP?

1. Folksonomy
  1. Users apply public tags to online items which give rise to classification system useful in search and retrieval
    1. Relevance to NLP is it can then be used in research and social bookmarking
2. Taxonomy
  1. Hierarchical way to distinguish levels of something
    1. Used for research with NLP

17. What is a key task (the hardest problem) of any NLP system or algorithm which still remains

unsolved? Give an example

1. Ambiguity
18. Name three types of ambiguity
  1. Lexical
  2. Syntactic
  3. Referential
19. What is the difference between syntactic and semantic analysis?
  1. Syntactic
    1. Task of recognizing a sentence and assigning a syntactic structure to it.
  2. Semantic
    1. Representing meaning of words
20. What is the difference between question answering and summarization?
  1. Question answering
    1. System that automatically return answers for a user's question by retrieving information from collected documents
  2. Question summarization
    1. Distill the most important information from a text to produce an abridged version for a particular task or user
21. What is the difference between abstractive and extractive summarization?
  1. Extraction takes actual data from the text while abstraction takes the meaning from the data and creates summarizations
22. What is the difference between text categorization and text clustering? [L04]
  1. Categorization
    1. Process of associating a document with one or more pre-existing subject categories
  2. Clustering
    1. Process of grouping documents based on similarity of features as interpreted typically by a machine learning algorithm
23. Explain the concept of bag of words and its application
  1. Bag of words is the complete list of words from given documents
24. What is a lexicon and how is it used in text categorization?
  1. Lexicon is the words ranked by their usage from given documents
25. Is it possible to do clustering and classification without a similarity/distance metric?
  1. NO
26. Explain a metric to compute document similarity. [L04-05]
27. Explain text classification by k nearest-neighbors
  1. You classify documents by the nearest neighbors, so you get the similarity between the documents and the documents most similar and classified together
28. Explain the basis of sentiment analysis and one application
  1. The basis of sentiment analysis is to read the context of a text document and be able to tell the overall sentiment of the document. An application is something that can read customer reviews and predict the number of stars given for the product based on the text review.
29. Explain the following concepts in sound signals [L06, and L07]
  - a. Frequency
    1. Number of times per second that a sound pressure wave repeats itself

- b. Amplitude
  - 1. Strength or level of sound pressure

30. Explain the following concepts in sound perception

a. Pitch

- 1. Related to frequency

b. Loudness

- 1. Intensity

c. Timbre

- 1. Classes of sounds

31. What does HRTF stand for?

- 1. Head Related Transfer Function

32. Explain the three stages of sound reproduction

- 1. ADC
  - 1. Analog to Digital processing
- 2. Processor
  - 1. Storage
- 3. DAC
  - 1. Digital to Analog

33. Define

- 1. Nyquist rate
  - 1. Minimum rate at which a signal can be sampled without introducing errors
- 2. dB
  - 1. Decibels which measures sound intensity
- 3. SNR
  - 1. Signal to noise ratio, measurement used to describe how much desired sound is present in an audio recording
- 4. Quantization
  - 1. Process of converting a continuous analog audio signal to a digital signal with discrete numerical values

34. Explain why we must sample above Nyquist rate

- 1. To avoid losing information in the signal

35. What is quantization noise, what is the difference between quantization noise and quantization error?

- 1. Noise is the rounding error when representing a signal as a discrete number
- 2. Error is the difference between the analog signal and the closest available digital value at each sampling instant from the A/D converter

36. For a signal uniformly quantized using 7 bits/sample what is the signal to quantization noise

ratio? [L06-P51]

6.02 x (7)

37. Why does non-linear quantization improve SNR? [L06-P33]

38. Define and explain PCM

1. Pulse Code Modulation

1. Analog signal is quantized into a number of discrete levels

39. Write the PCM binary code received for the signal below:

1. 1101 1111 1011 0110 0001 0001 0101 1010

40. Define

a. AM

1. Amplitude Modulations

b. FM

1. Frequency Modulation

c. MIDI

1. Musical Instrument Digital Interface

41. Define voiced vs. non-voiced sounds [L07, P33-34]

1. Voiced

1. Make our vocal cords vibrate

2. Non-voiced

1. Produced from air passing through the mouth at different points

42. True or false:

1. voiced sounds are characterized by vibration of the vocal cords

1. T

2. unvoiced sounds involve no vibration, just random air pressure through the larynx

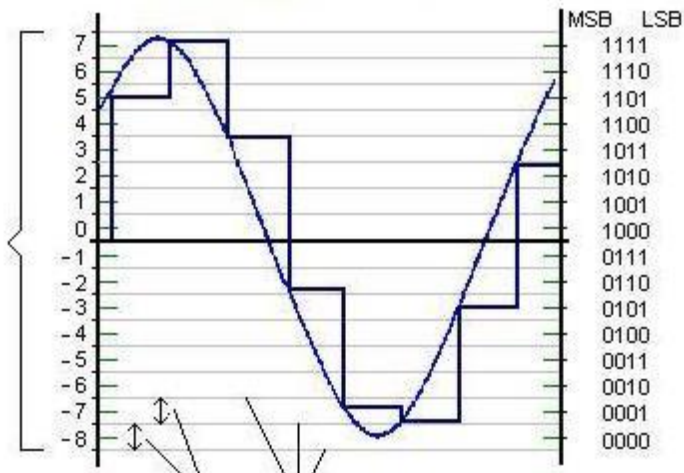
1. T

3. Most vowels are voiced

1. T

4. Consonants can be voiced or voiceless

1. T



43. What is CEPSTRUM analysis, and why is it related to deconvolution?

1. Tool for detection of periodicity in a frequency spectrum
2. Concerned with the deconvolution

44. What are the 4 steps of Cepstral Analysis

1. Take the Fourier transform of a time history
2. Take the magnitude of the Fourier transform
3. Take the natural logarithm of the magnitude
4. Take the inverse Fourier transform of the natural logarithm

45. What is homomorphic filtering

1. Nonlinear transformation usually applied to image and speech processing used to convert a signal obtained from a convolution of two original signals into the sum of two signals

46. What is the role of metadata in music information retrieval systems? (L09-P09)

1. Estimate artist similarity and track similarity. Also important for factual metadata (dates, artist, title)

47. True/False – the following are examples of low level audio features

1. Spectrogram
2. Harmonicity
3. Zero crossing rate
4. Melody
5. Lyrics

48. True/False

a. High level music features are more robust than low level features

T

b. Silence ratio is a frequency domain feature

F

c. Music features are extracted on time segments

T

d. Similarity metrics enable visualization of music content

T

e. The bandwidth of music is smaller than the bandwidth of speech

F

49. In information theory, what is the definition of entropy? [L09-P41]

1. Expected value of information. If no change there is no information

50. Which compression algorithm has higher SNR for the same input signal?

1. Lossless compression

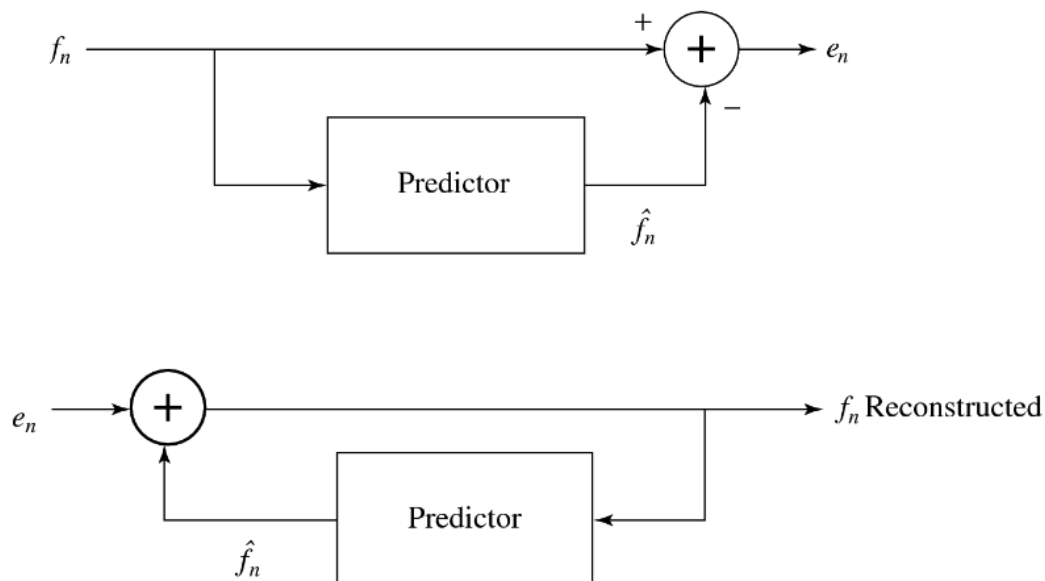
2. Lossy compression

51. Explain the principle of differential coding for audio

1. It makes data to be transmitted depend not only on the current signal state (or symbol), but also on the previous signal state.

52. Draw the diagram and explain lossless predictive coding

1. Transmitting differences – predict the next sample as being equal to the current sample; send the difference between previous and next





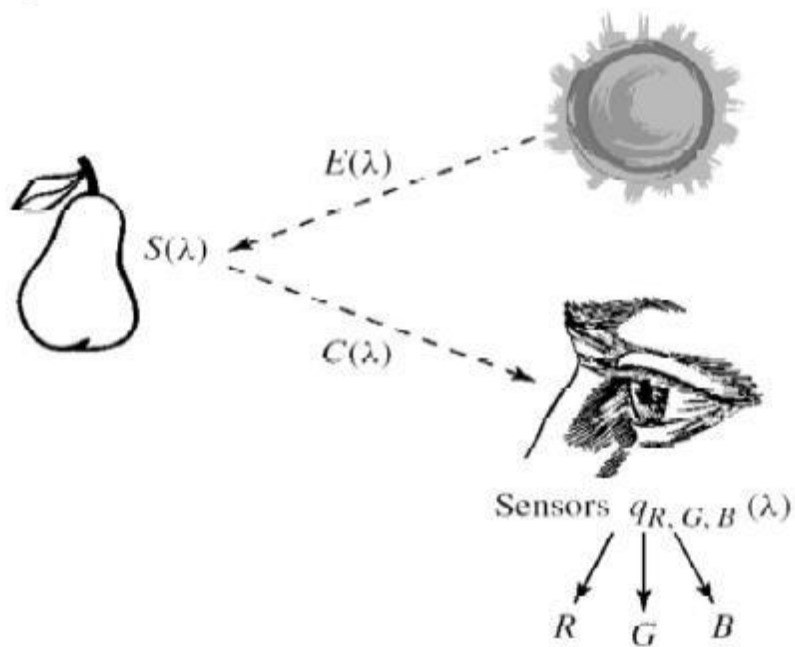
53. What is DPCM (L09-P48,P56)?

1. Same as PCM except that it incorporates a quantizer step

54. What is the range of visible light in nm (of wavelength) [L10]

1. 400 - 750

55. In the image formation model



Explain  $E(\lambda)$ ,  $S(\lambda)$ ,  $C(\lambda)$ ,  $q_{R, G, B}(\lambda)$

$S(\lambda)$ , = reflectance functions

$E(\lambda)$  = Illuminant

$C(\lambda)$  = Color signal

56. Which of the two images below has been gamma corrected?



The first one since gamma correction improves contrast

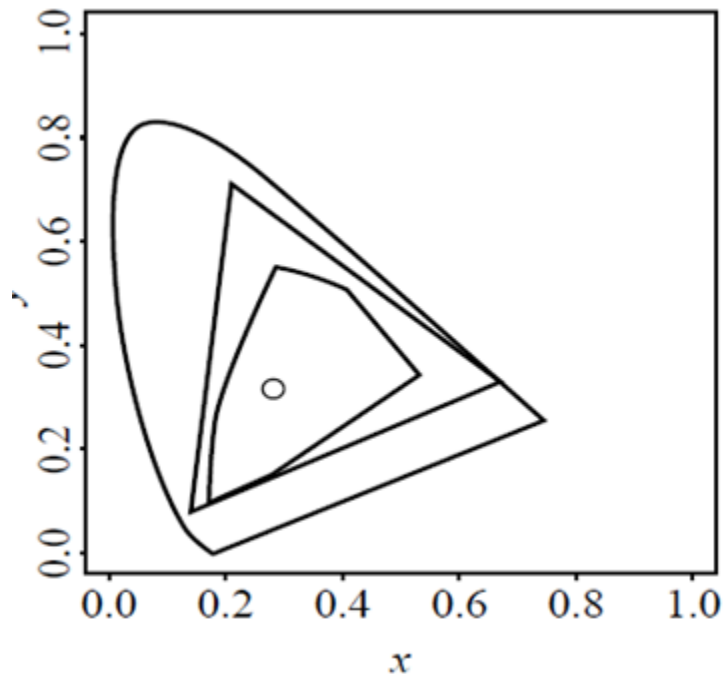
57. Explain contrast in luminance (black-white) and color (opposites: Y-B, R-G).
  1. Difference in luminance or color that makes an object distinguishable from the background
58. Given the equations to convert from RGB to XYZ, convert the color Magenta in 8-bit RGB (255,0,255) to XYZ (hint: convert RGB to 0-1 scale before converting color spaces)

$$X = 0.3935 \cdot R + 0.3653 \cdot G + 0.1916 \cdot B$$

$$Y = 0.2124 \cdot R + 0.7011 \cdot G + 0.0866 \cdot B$$

$$Z = 0.0187 \cdot R + 0.1119 \cdot G + 0.9582 \cdot B$$

59. The outside shape is the chromaticity diagram, the small circle is the reference white, of the remaining figures which is the gamut of the display and which the gamut of the printer:



The polygon

60. Define image resolution, aspect ratio, and size [L11]

1. Image Resolution
  1. Determines both the amount of detail as well as its storage requirements
2. Aspect Ratio
  1. Expresses resolution as the product of number of horizontal pixels by number of vertical pixels
3. Size
  1. Refers to both the picture resolution and physical dimensions of the image

61. What is the dynamic range of an image, and how does it affect quality?

1. Maximum contrast of an image, effects quality by if it is a low dynamic range then the contrast can't be high quality

62. What is pixelation?

1. Used to describe blurry patches or fuzziness in an image due to visibility of single colored pixels

63. What is dithering and what is it used for? Does it work for color images?

1. Dithering is used to reduce the color range of images
2. It does work for color and grey-scale images

64. What is the difference between visual quality (measured in subjective score) and quality of experience? [L12]

1. Visual quality can be given a score while quality of experience is more preference and hard to score

65. Name 3 desirable image features and 3 undesirable image features

Desirable:

1. Sharpness
2. Resolution
3. Colorfulness
4. Brightness
5. Contrast

Undesirable:

1. Noise
2. Artifacts
3. impairment

66. Name 3 network parameters that are adjusted during quality of service (QoS) management

1. Buffering
2. Bitrate
3. Lag duration
4. AV sync

67. What are the three steps of a basic edge detector? (L12-P55)

1. Filter noise using mean filter
2. Compute spatial gradient
3. Mark points > threshold as edges

68. Name four types of image regions and how can we identify them? (L12-P59)

1. Flat
2. Edges
3. Details (fine structures)
4. Textures (surface complexity as it appears in 2D images)

69. In MMIS video data falls in two categories, produced and observed, explain the difference. [L13, p.10-11]

1. Produced are created by an author who is actively selecting content and has control over the appearance of the video.

2. Observed is typical in surveillance, it is also found in broadcasting. EX. Sports.

70. In video-based MMIS, explain the role of ontologies. [L13, p. 13,25]

1. Way to store organized knowledge so that the meaning of a concept is uniquely defined and used consistently

71. What is the semantic perspective in a produced video document? [L13, p. 14]

1. Which ideas did the author have

72. What the content perspective in produced video? [L13 p.16]

1. indicating which content is used

73. Explain what the semantic gap in MMIS is. [L13, P19-20]—related to the difference between perceptual and semantic similarity [L13, p.23]
1. Semantic gap – lack of coincidence between the information that one can extract from the sensory data and the interpretation that the same data have for a user in a given situation.
  - 2.
74. What is Word2Vec used for? [L13-P26]
1. Used as a similarity or distance metric
75. Explain the process of video content segmentation. [L13, P29-31]
1. Segmentation made that decomposes the video in its layout and content elements
  2. Use algorithms to detect cuts in the video and transistions
  3. Content – Detect people, objects, and settings
76. What is video summarization? [L13-P36]
1. Summarize content and present users with its essence in the form of a moving storyboard (video skimming)
77. The perspective used in semantic video indexing comes from (select one answer): a. the author,
- b. the location, c. the camera. [L13, p. 4]