

In The Name of God

An Introduction to Multimedia World (with emphasis on Video/Image)

Omid Fatemi, Ph.D.

fatemi@irandoc.ac.ir

ofatemi@ut.ac.ir

Iranian Research Institute for Information Science and Technology, <http://www.irandoc.ac.ir>

99/بهار

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

1

Outline

- Multimedia and Digital Signals
- Analog and Digital Image and Video
- Data Compression
- Image Compression - JPEG
- Video Coding – MPEG
- Scalability
- MPEG-2 System
- Implementations and challenges
- New Standards

99/بهار

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

2

Gartner Identifies the Top 10 Strategic Technologies for 2011

1. Cloud
 2. Mobile and Tablets
 3. Social Computing
 4. Big Data and Analytics
 5. Next-Generation Security and Identity
- Video is not a new media form, but its use as a standard media type used in non-media companies is expanding rapidly. Technology trends in digital photography, consumer electronics, the web, social software, unified communications, digital and Internet-based television and mobile computing are all reaching critical tipping points that bring video into the mainstream.

99/بهار

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

3

What Is Multimedia?

- Apps that involve more than conventional data types (e.g., text, drawing and images)
- Best examples are continuous media (e.g., animations, audio, and video)



From 00:13

99/بهار

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

4

Digital Multimedia

- Field concerned with integration of.

- Text.
- Graphics.
- Still and moving images.
- Animation.
- Sounds.



- any other medium where every piece of info can be represented, stored, transmitted, or processed digitally.
- E.g. smell (scent).

99/بهار

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

5

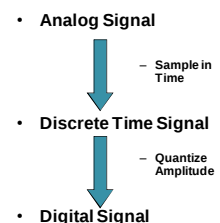
Digital Representation

- Digitization.

- Analog.
- Discrete Time.
- Digital.

- Why Digitize?

- Universality of representation.
- Robustness to error; aging, distortion, noise.



99/بهار

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

6

Advantages of Digital Representation

- Universal Storage
- A single digital network for transmission
- Computers for processing and manipulation

99/بخت

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

7

Disadvantages of Digital Representation

- Quantization distortion.
- Sampling distortion (aliasing).
- Need for large amounts of digital storage capacity → need for compression.

99/بخت

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

8

Conventional Classification of MM Types

Time Based:	Audio, Video	ANIMATION (2D, 3D)
Space Based:	STILL IMAGES	TEXT, GRAPHICS
	Captured from the real world.	Synthesized by computer.

99/بخت

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

10

Image and Video

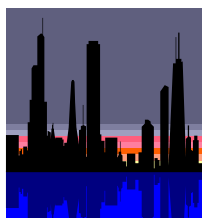
Analog and Digital Representation

99/بخت

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

10

Why Visual?



=

Buildings
Night

•
•
•

Shadow

A PICTURE IS WORTH A THOUSAND WORDS

99/بخت

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

99/بخت

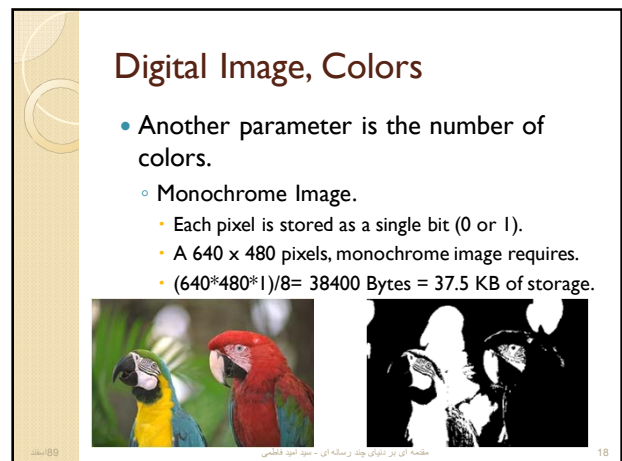
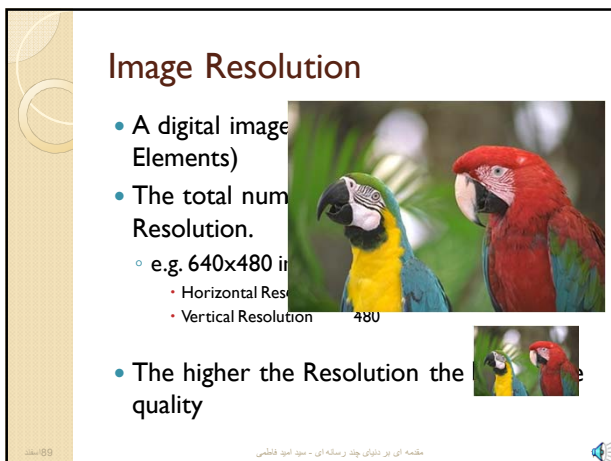
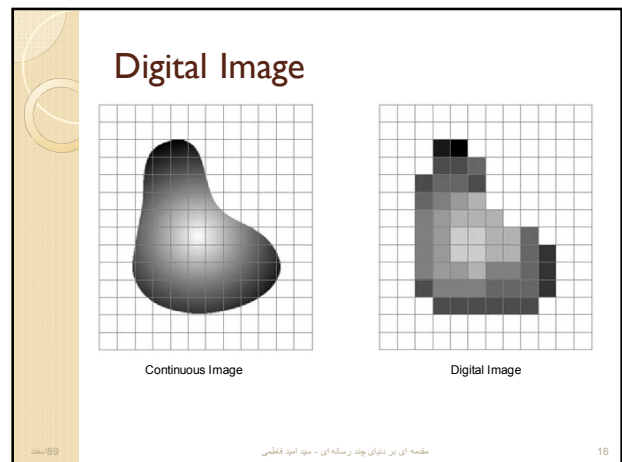
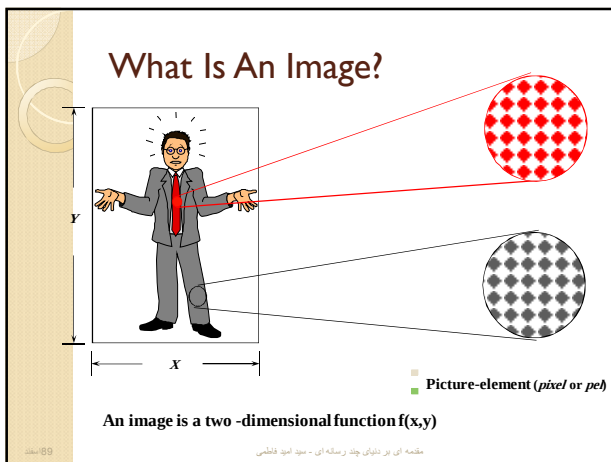
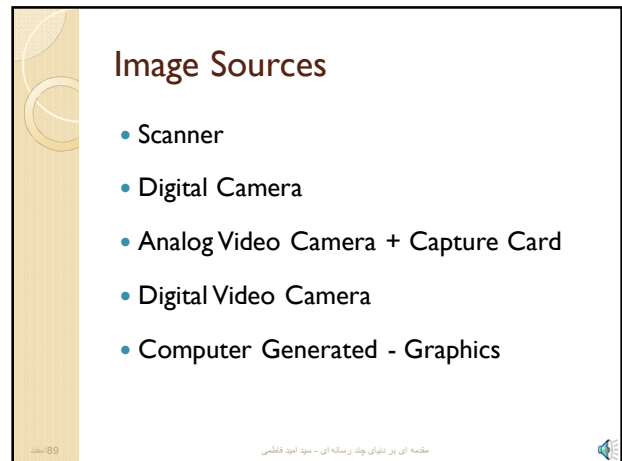
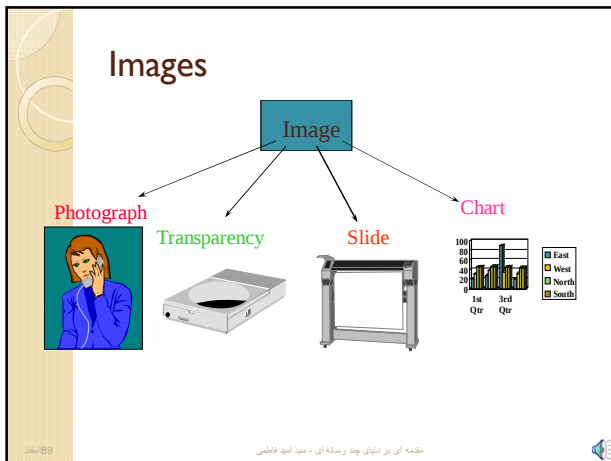
Seeing Is Believing ...



99/بخت

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

99/بخت



Grayscale Images

- Each pixel is usually stored as a byte (value between 0 to 255).
- We can perceive only few dozen gray levels.
- A 640 x 480 grayscale image requires over 300 KB of storage.
- $(640 \times 480 \times 8) / 8 = 307200 \text{ Bytes} = 300 \text{ KB}$ of storage.



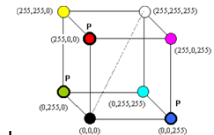
99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

21

Color Models

- RGB (Red- Green- Blue) Color Model.
 - Most commonly used in computer graphics.
 - The RGB colour model has three basic primary (basic) colours:
 - red,
 - Green,
 - blue.



- All other colours are obtained by combining them. This model can be thought as a cube, where 3 non-adjacent and perpendicular corners are R, G and B, like in the figure.

99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

22

Color Images

- Another alternative is to have one byte for each pixel
- Supports 256 out of the million colours possible, acceptable colour quality
- Requires Colour Look-Up Tables (LUTs)
 - Palette
- A 640 x 480 8-bit colour image requires 307.2 KB of storage (the same as 8-bit gray-scale)

99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

21

Color Images

24 Bits



8 Bit- Standard Palette



8 Bit- Selected Palette

99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

22

YUV Colour Model

- Initially, for PAL analog video, it is now also used in CCIR 601 standard for digital video
- Y (luminance) is the CIE Y primary. $Y = 0.299R + 0.587G + 0.114B$
- Chrominance is defined as the difference between a colour and luminance. It can be represented by U and V -- the colour differences.
 - $U = B - Y$
 - $V = R - Y$
- If b/w image, then $U = V = 0$. --> No chrominance!

99/بیت

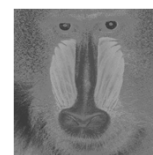
مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

23

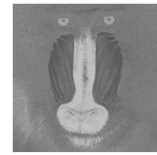
Sample YUV Decomposition



Y



U



V

99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

24

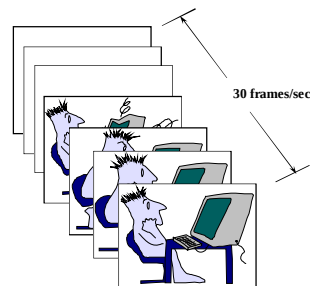
Video

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

25

A Video Sequence



99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

26

Standards Groups

- ITU - International Telecommunication Union
 - Formerly CCITT
- FCC
- SMPTE - Society of Motion Picture and Television Engineers
- ... and many more!

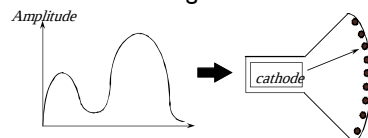
99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

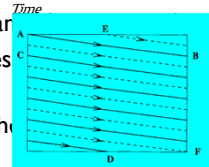
27

Video Display Scanning

- Analog video is a continuous signal that drives a CRT



- Video composed of luma and chroma
- Composite video combines luma and chroma
- Component video sends the luma and chroma signals separately



99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

28

Representations

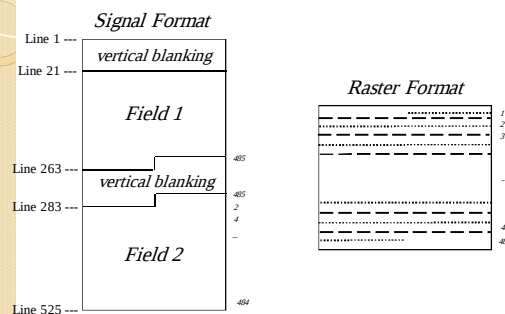
- Composite
 - NTSC - 6MHz (4.2MHz video), 29.97 frames/second
 - PAL - 6-8MHz (4.2-6MHz video), 25 frames/second
- Component
 - Separation video (luma, chroma) – svideo, HDMI
 - RGB, YUV, YIQ, ...
 - YCRCB - used for most compressed representations

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

29

Interlaced Fields



99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

30

Aspect Ratio/Refresh Rate

- Aspect ratio
 - Conventional TV is 4:3 (1.33)
 - HDTV is 16:9 (2.11)
 - Cinema uses 1.85:1 or 2.35:1
- Refresh Rate
 - NTSC is 60Hz (actually 59.94Hz)
 - PAL/SECAM is 50Hz
 - Cinema is 48Hz

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاضلی

31

Digitizing (Spatial and Temporal)

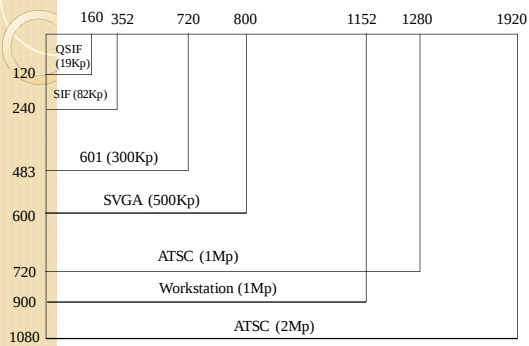
- Analog TV is a continuous signal
 - Camera grabs fields so small movement between fields
- Digital TV uses discrete numeric values
- Signal is sampled
- Samples are quantized
- Small, discrete regions are digitized
- Image represented by pixel array
- DVBT is now broadcasting in Tehran

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاضلی

32

Pixel Arrays

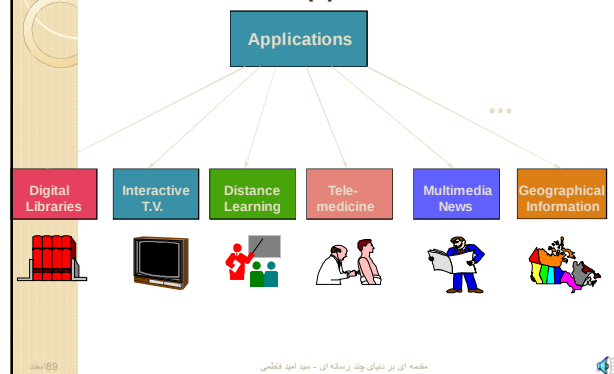


99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاضلی

33

Multimedia Applications



99/بستان

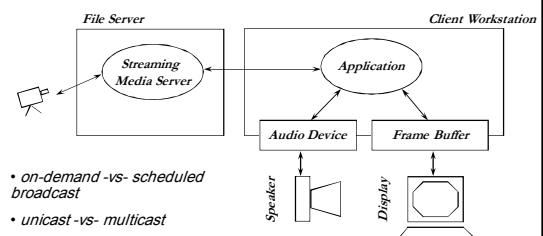
مقدمه ای بر فناوری چند رسانه ای - سید امید فاضلی

34

Multimedia in Consumer and Embedded Markets



Streaming Playback



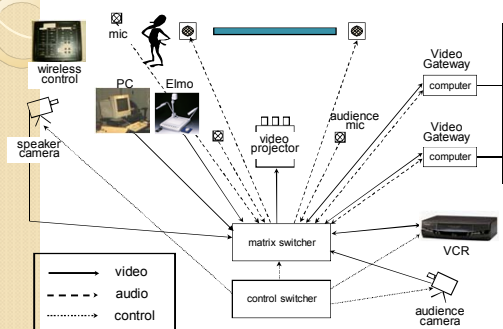
- on-demand -vs- scheduled broadcast
- unicast -vs- multicast
- bandwidth -vs- quality
- hardware -vs- software CODEC
- multiplexed -vs- single streams
- control?

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاضلی

36

Multimedia Presentation



99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

37

Google Earth 3D, Street View



99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

38

Sixth Sense



Part one from 6:23

99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

39

Second life



99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

40

Video Archive (VoD): iransima.ir

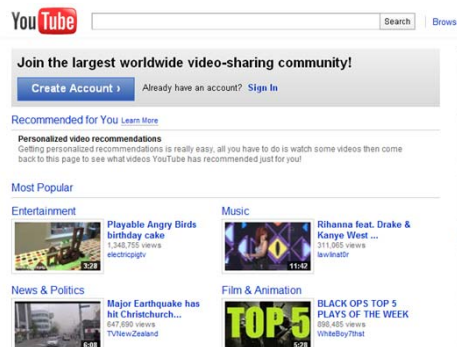


99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

41

YouTube



99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

42

Interactive TV



99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

43

Fiber to Home – Verizon FIOS TV

Get more HD channels.

FIOS TV gives you over 130 HD channels, and access to more than 10,000 On Demand movies and programs each month. Watch what you want, when you want it. You can always access channels like the Food Network, the Disney Channel, and all your premium movie content, anytime.

See why fiber-optic TV is better.

Simply put, the FIOS network is built for HD TV—so picture quality is sharper, uncompressed, and more reliable in bad weather. It's higher quality TV, brought to you by the nation's largest 100% fiber-optic network.

Watch a smarter version of TV.

FIOS TV comes with an interactive media guide, tools and lots of widgets. Choosing a big game on ESPN? Get statistics and ESPN fantasy football updates on-screen while you're watching. Or access Facebook and Twitter from your TV and update your friends on the action. Also get instant local weather, traffic, and much more. And with Flex View, you can take your favorite entertainment with you when you go.



Iron Man 2 now playing on FIOS On Demand

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

44

Multimedia Messaging Services

Messaging interface



All messaging services including SMS, MMS & e-mail collected under one application

Photo album



Pictures can be stored in the photo album

MMS message



MMS messaging can combine text, pictures and audio

MMS delivery



The address formats supported are global phone numbers and simple e-mail addresses

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

45

4G Networks?

- Bandwidth of 100Mbit/s
- Due 2010
- But: Too many 3G standards now, slows development
- Permanent internet connection (IPv6 address) while mobile
- All digital, packet switched, with tight security
- Enables full interactive video services

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

46

The current product is...

- High Definition DVD
 - 1080 pixels vertical resolution, 60 Hz
- Blu-Ray discs
 - 25/50GB of storage
 - 4 hours of HD video
- H.264 and Windows Media Video 9

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

47



99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

48



Why is Multimedia Hot?

- Multimedia is commonly used
- Convergence of computers, telecom, and TV
 - Caused by technology and competition (cable and phone companies)
 - Dramatic changes in products, infrastructure, etc.
- Application potential
 - Huge markets
 - Potential for improving our lives (e.g., learning, entertainment, and work)

Why Is Multimedia Hot? (Cont.)

- Digital audio and video is revolutionizing music, film, and video industries
- Collaboration, virtual environments, and webcasting
- Interesting technical issues
- It is fun

Are Multimedia Applications Hard?

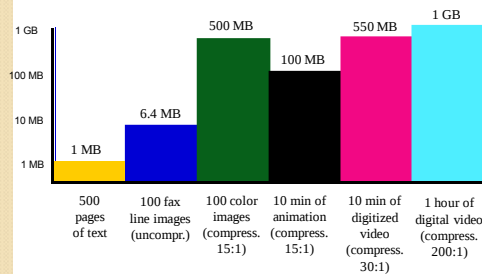
- **Higher data rates**
 - Fast I/O subsystem (fast/wide SCSI-2, fiber channel, PCI, AGP)
 - High speed backplane (PCI or faster)
 - Faster network (1-25Mbps per video stream)
 - **Real-time, distributed software**
 - **More storage required**
 - 32-128 MB main memory
 - 10-30 GB secondary storage
 - TB's of tertiary storage
 - **Hardware CODEC, modified CPU, and modified frame buffer/graphics subsystem**
- ...essentially new hardware and software*

Enabling Technologies

- Compression
- Why Compression?
- (a) large storage (b) slow storage device
- (c) network bandwidth
- ex: 620×560 pixel frame of color video
- 24 bits/pixel --> 1MB
- 30 f/s --> 30 MB/s data rate of storage device; CDROM 300 KB/s
- 30 min --> 50 GB storage

Storage requirements for a typical multimedia application

Storage Requirements



Data Compression

Information Theory

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

56

Lossless / Lossy

- Lossless
 - $\text{decompress}(\text{compress}(\text{data})) = \text{data}$
- Lossy
 - $\text{decompress}(\text{compress}(\text{data})) \neq \text{data}$
 - A small change in pixel values may be invisible

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

57

Lossless vs. Lossy

- Lossless:
 - Decompressed and original are identical
 - Pros: No artifacts, more reliable, Editable
 - Cons: Lower compression ratios (e.g. 3-4 for images)
 - Usually used for documents, and binary data
- Lossy:
 - Decompressed and original are slightly different.
 - We try to keep artefacts as not noticeable by the human eye as possible.
 - Pros: High compression rate (e.g. Up to 100)
 - Cons: The artefacts are irreversible. Can't repeat the process many times, makes editing very difficult.

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

58

Run-Length Encoding

- Repeating string of characters, called a *run*, is coded into two bytes
 - First byte contains the *run count*, one less than the number of repetitions
 - Second byte contains the *run value*, the character being repeated

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

59

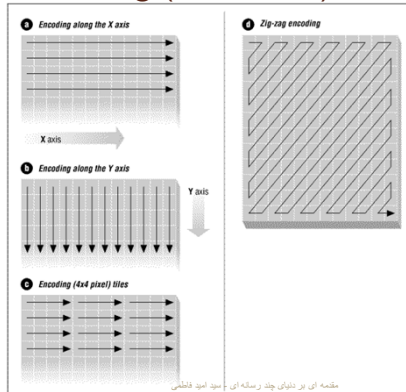
RLC Example

- '77777zzzzzzzzV' becomes '472z5y0V'
 - 15 byte string becomes 8 bytes long
 - Compression ratio of almost 2 to 1
- Some strings become twice as long
 - '7fu5JLY9jhYlujG'

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

60

Scanning (2D → 1D)



99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

61

Measure of Information Content

- Given a source (image) which generates a discrete set of dependent x_i messages with probabilities p_i , the information associated with each message is:

$$I_i = -\log_2 p_i$$

- The entropy of a source is defined as the average information generated by the source:

$$H = -\sum_{k=1}^{MN} p_k \log_2 p_k \quad \text{bits/pel}$$

- Shannon's Source Coding Theorem states that the number of bits per symbol
- required to encode the output of a source is lower bounded by the entropy of that source.

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

62

Entropy Coding Algorithms

- An entropy encoding is a coding scheme that assigns codes to symbols so as to match code lengths with the probabilities of the symbols.
- The most common entropy encoding techniques:
 - Huffman Coding
 - Arithmetic Coding

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

63

Image Compression

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

64

Image compression (Transform Based)



Compression is accomplished by:

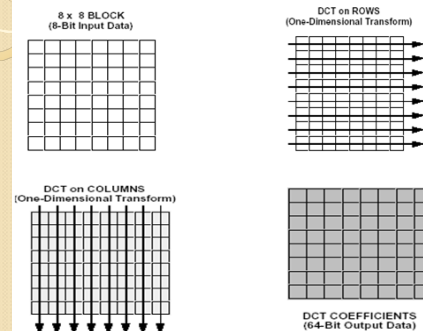
- applying a linear transform to decorrelate the image data
- Quantizing the resulting transform coefficients
- Entropy coding the quantized values.

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

65

2-D Discrete Cosine Transform



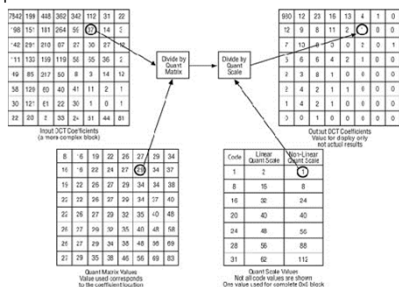
99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

66

Quantization

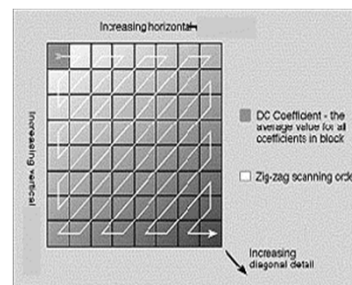
- Each coefficient from the DCT is divided by the value in corresponding location in a quantization matrix.



مقدمه ای بر دنیای جلد و سازه ای - سید امید فاطمی

67

Zig-zag Scanning



مقدمه ای بر دنیای جلد ۲ ساله ای - سید امید قاضی

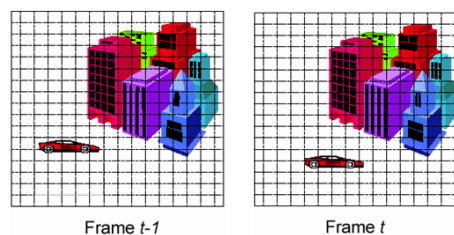
68

Video Coding

مقدمه ای بر دنیای جنگ و ساله ای - سید امید فاطمی

69

Characteristics of Typical Videos



Adjacent frames are similar and changes are due to object or camera motion

مقدمه ای بر دنیای جلد ۲ شماره ای - سید امجد قاسمی

70

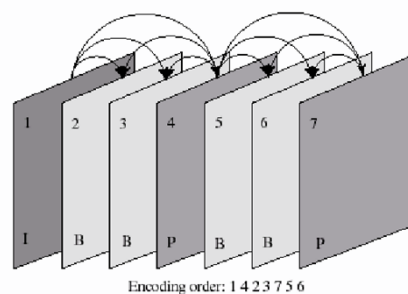
Key Idea in Video Compression

- Predict a new frame from a previous frame and only code the prediction error
- Prediction error will be coded using the DCT method
- Prediction errors have smaller energy than the original pixel values and can be coded with fewer bits

مقدمه ای بر دنیای جنگ و ساله ای - سید امید فاطمی

71

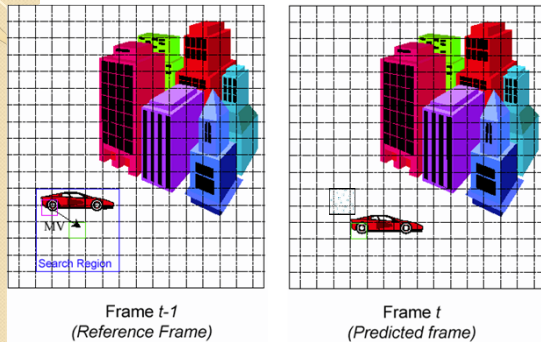
Different Coding Methods



مقدمه ای بر دنیای جلد رسانه ای - سید امجد قاضی

72

Block Matching Algorithm for Motion Estimation

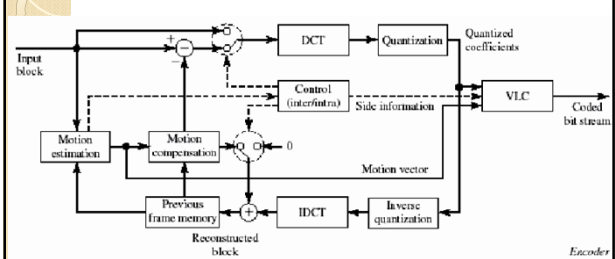


99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

73

Encoder Block Diagram

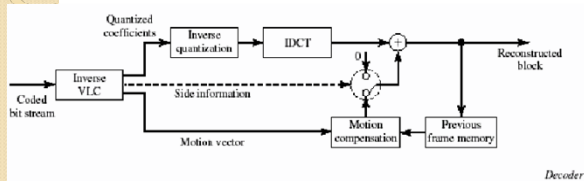


99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

74

Decoder Block Diagram



99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

75

Rate Control

- The coding method necessarily yields variable bit rate
- constant bit rate (CBR) channel
- Rate control is necessary
- The fluctuation within the period can be smoothed by a buffer at encoder.

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

76

Scalability

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

77

Scalable Coding

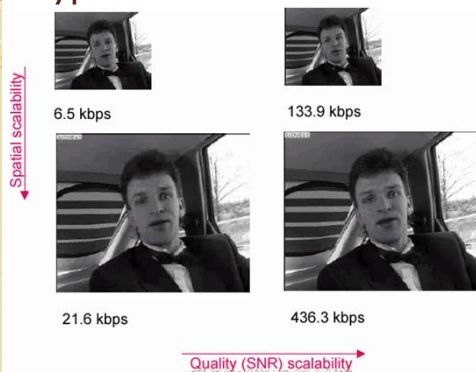
- Motivation
 - Real networks are heterogeneous in rate
- streaming video from home (56 kbps) using modem vs. corporate LAN (10- 100 mbps)
- Scalable video coding
 - Ideal goal (embedded stream): Creating a bitstream that can be accessed at any rate

99/بستان

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

78

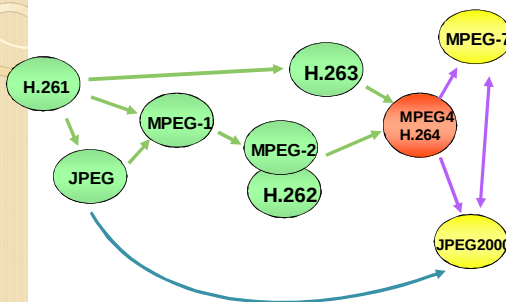
Typical Scalabilities



89

79

Compression Standards

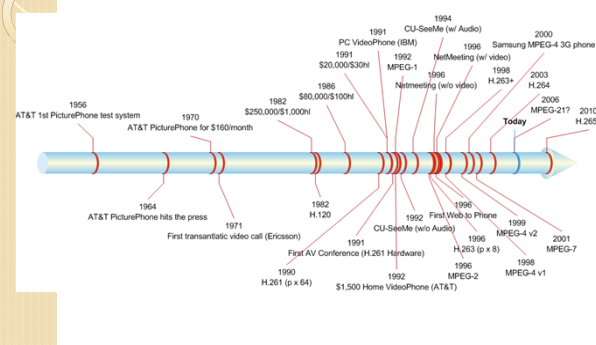


89

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

80

The history

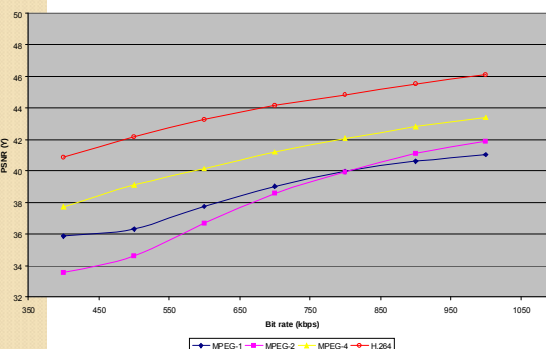


89

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

81

R-D Performance of MPEG Codecs



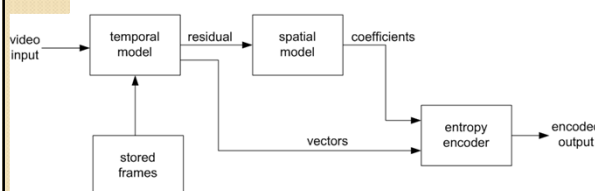
Mission Impossible 3 HDTV trailer (1080p) resized to 720x304. Encoders: TMPGEnc (MPEG-1/2), XviD (MPEG-4), x264 (H.264)

89

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

82

CODEC Design

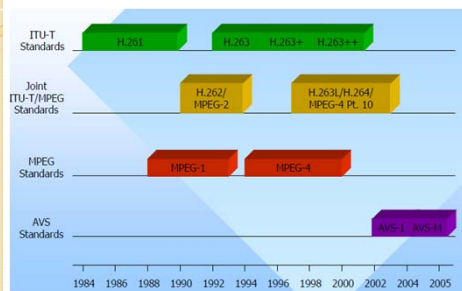


Reproduced from "H.264 and MPEG-4 Video Compression" by I. E. G. Richardson.

89

83

Audio Video coding Standard of (AVS) China



89

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

84

Development stages of AVS [3]

December 2003

- In the 7th AVS meeting, AVS-video (part-2) and AVS-system (part-1) was finalized.

December 2004

- In the 11th AVS meeting, AVS-M (part-7) was finalized

March 2005

- Authentication of 'AVS101' – high definition decoding chip

May 2005

- AVS Industry Alliance was set up.

June 2005

- Joint AVS/ISMA workshop on IPTV standard and industry forum

February 2006

- AVS part-2 was announced as a national standard.

85

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

85

Standards

- The standards in image/video compression :
 - JPEG (Joint Photographic Experts Group)
 - MPEG-1 and MPEG-2
 - H.261 and H.263
 - MPEG-4
- all use Discrete Cosine transform and Block transform-based techniques.

86

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

86

Implementations

challenges

87

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

87

Characteristics of Multimedia Apps (I)

- Requirement for real-time response
 - "Incorrect" result often preferred to slow result
 - Unpredictability can be bad (e.g. dynamic execution)
- Narrow data-types
 - Typical width of data in memory: 8 to 16 bits
 - Typical width of data during computation: 16 to 32 bits
 - 64-bit data types rarely needed
 - Fixed-point arithmetic often replaces floating-point
- Fine-grain (data) parallelism
 - Identical operation applied on streams of input data
 - Branches have high predictability
 - High instruction locality in small loops or kernels

88

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

88

Characteristics of Multimedia Apps (2)

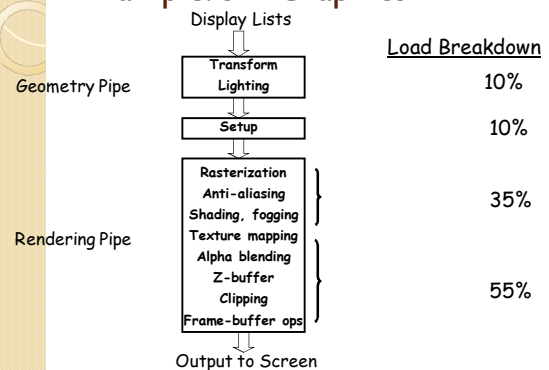
- Coarse-grain parallelism
 - Most apps organized as a pipeline of functions
 - Multiple threads of execution can be used
- Memory requirements
 - High bandwidth requirements but can tolerate high latency
 - High spatial locality (predictable pattern) but low temporal locality
 - Cache bypassing and prefetching can be crucial

89

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

89

Example: 3D Graphics

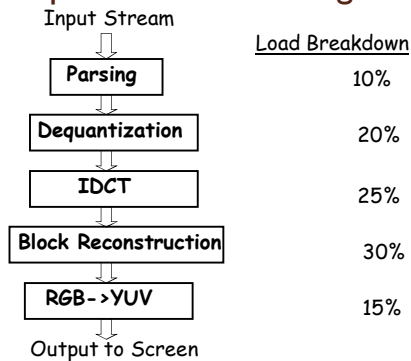


90

مقدمه ای بر فناوری چند رسانه ای - سید امجد فاطمی

90

Example: MPEG Decoding

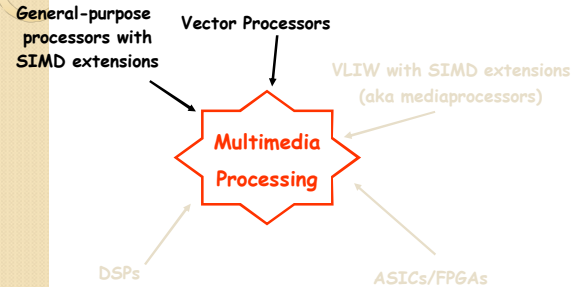


99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

91

Approaches to Media Processing



99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

92

New Standards

Multimedia Standards

99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

93

Other standards

- MPEG-7
- The increasing demand for searching multimedia content on the web has opened up new opportunities for creation and delivery of these items on the Internet. Today many elements exist to build an infrastructure for the delivery and consumption of multimedia content. The main goal of MPEG-7 is to specify a standard set of descriptors that can be used to **describe various types of multimedia information coded with the standard codecs**, as well as other databases and even analogue audio-visual information.

99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

94

Other standards

- MPEG-21
- Today multimedia technology is so advanced that access to the vast amount of information and services from almost anywhere at any time, through ubiquitous terminals and networks, is possible. However, no complete picture exists of how different communities can best interact with each other in a complex infrastructure. Developing a **common multimedia framework** will facilitate cooperation between these sectors and support a more efficient implementation and integration of different models, rules, interests and content formats. This is the task given to the multimedia framework project under the name of MPEG-21.

99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

95

Other standards

- JPEG2000
- The JPEG2000 standard is devised with the aim of providing the best quality or performance and capabilities to market evolution that the current JPEG standard fails to cater for

99/99

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

96

Research in Multimedia

IT context

99

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

97

Information within a video

- Cognitive
 - News
 - News reporter
 - Trevor McDonald
 - News report
 - Robbery
 - Money \$
 - Heathrow
 - Get away van



99

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

Information within a video

- Formative
 - Pattern, shape, form
 - Colour statistics
 - % skin colour
 - Spatial textures
 - Global motion



99

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

99

Semantic gap

lack of coincidence between formative and cognitive information

99

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

100

Multimedia Understanding

- Video mining
- Multimedia understanding
 - Video summary
 - Concept query
- Multimedia ontology

99

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

101

Content Based Video Analysis

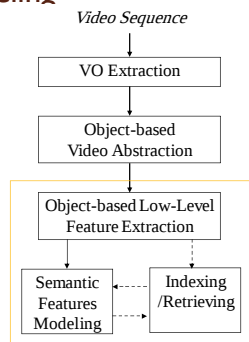
- Using automatic analysis techniques
 - shot detection, video segmentation
 - key frame selection
 - object segmentation and recognition
 - visual/audio feature extraction
 - speech recognition, video text, VOICR

99

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

102

Object-based Semantic Video Modeling



99/138

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

103

Image Understanding Functions

- Scene segmentation
- Similarity matching
- Camera motion determination and object tracking
- Optical Character Recognition (OCR) on video text and titles
- Face detection and recognition

99/138

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

104

Content-based Video Retrieval

- Simple visual feature query
 - Retrieve video with key-frame: Color-R(80%),G(10%),B(10%)
- Feature combination query
 - Retrieve video with high motion upward(70%),Blue(30%)
- Query by example (QBE)
 - Retrieve video which is similar to example
- Localized feature query
 - Retrieve video with a running car toward right
- Object relationship query
 - Retrieve video with a student reading a book
- Concept query (query by keyword)
 - Retrieve explosion, Libya uprising
- Semantic description:
 - MPEG-7

99/138

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

105

Optical Character Recognition (OCR)

- First patent in OCR in 19th century
- First applications in post-office and banks
- Documents easier to distribute, search, organize, and edit in digital form
- State of the art not perfect, but 99% accurate on cleanly printed pages
- Farsi OCR?

99/138

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

106

Text Area Detection



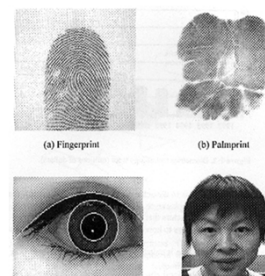
99/138

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

107

Biometrics

- Technology for identification
 - Finger/palm print
 - Iris
 - Face

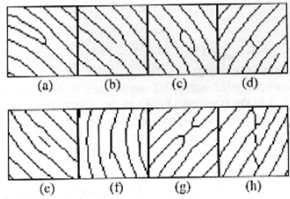


99/138

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

108

Fingerprints



99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

109

Face Identification

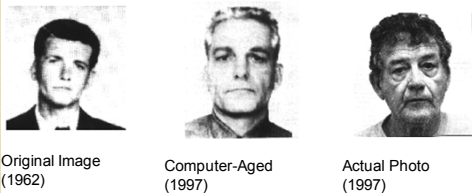
- Not quite reliable yet.
 - Performance degrades rapidly with uncontrolled lighting, facial expression, and size of database
- Commercial systems installed in London and Brazil to catch criminals

99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

110

Automatic Age Progression



99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

111

Automatic Target Recognition (ATR)

- Finding mines, tanks, etc.
- Billion dollar a year industry
 - Martin-Lockheed, TSR, Northrup-Grumman, other aerospace contractors.
- Various types of imagery:
 - Synthetic Aperture Radar (SAR), Sonar, hyper-spectral imagery (more than 3 colors)

99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

112

Better Security Cameras

- Cameras that are responsive to the environment
 - Track and zoom on moving objects
 - Automatic adjustment of contrast

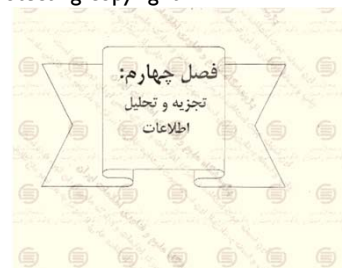
99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

113

Image Processing: Watermarking

- Information hiding
 - Protecting copyright



99/بیت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

114

Computer Vision: Guiding Motion

- Visually guided manipulation
 - Hand-eye coordination
- Visually guided locomotion
 - robotic vehicles



CMU NavLab II

99/11/23

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

115

Find Images With Similar Colors



99/11/23

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

116

Find Images with Similar Shape



99/11/23

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

117

Goal: Find Images with Similar Content



99/11/23

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

118

Status of Image Search

- Typical Search Features
 - Color
 - Texture
 - Shape
 - Spatial attributes (local color regions, less common than global color, texture, shape metrics)

99/11/23

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

119

Google Image Search

Google Advanced Image Search [About Google](#)

Find results related to all of the words, related to the exact phrase, related to any of the words, not related to the words

google image query

Content types Return images that contain

☒ any content ☐ faces ☐ photo content ☐ clip art

☐ line drawings

Size Return images that are

Any size Width: Height:

Exact size Return images exactly the size

Use my desktop size

Aspect ratio Return images with an aspect ratio that is

Any aspect ratio

Filetypes Return only image files formatted as

any filetype

Coloration Return only images in

any colors

Domain Return images from the site or domain

Usage Rights Return images that are

not filtered by license

More info

SafeSearch ☒ No filtering ☐ Use moderate filtering ☐ Use strict filtering

99/11/23

مقدمه ای بر دنیای چند رسانه ای - سید امید فاطمی

120

Conclusions

- Multimedia is becoming important
- Image and Video are the main part of Multimedia systems
- Lots of applications
- Future is video in content
- Lots of research areas for Iranian Research Institute for Information Science and Technology

99/بخت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

121

Thank you for your attention

Comments, Suggestions, Questions?

fatemi@irandoc.ac.ir

99/بخت

مقدمه ای بر فناوری چند رسانه ای - سید امید فاطمی

122