Multimedia - A few Aspects

STTP on Image Processing and Multimedia

Sunil Kumar Kopparapu

SunilKumar.Kopparapu@TCS.COM

Speech and Natural Language Processing Group

TCS Innovation Labs - Mumbai,

Yantra Park, Thane (West), Maharastra

June 2007

Multimedia - In a nutshell

It’s as large as life, and twice as natural!

Multimedia - In a nutshell

It's as large as life, and twice as natural!

Lewis Carroll,
Through the Looking-Glass

Multimedia - In a nutshell

It's as large as life, and twice as natural!

Lewis Carroll,
Through the Looking-Glass

Next set of slides ....

Multimedia - In a nutshell

It's as large as life, and twice as natural!

Lewis Carroll,
Through the Looking-Glass

Next set of slides ....

- cover a small aspect of it
- neither comprehensive
- nor complete!
- material borrowed from *all* sources

What is Media?

What is Media?

- What is Media?
  - A plural of medium [Dictionary]
  - Formats for **presenting information** [Wiki]
  - a means or instrumentality for *storing* or *communicating information* [WordNet]

What is Media?

- What is Media?
  - A plural of medium [Dictionary]
  - Formats for **presenting information** [Wiki]
  - a means or instrumentality for *storing* or *communicating information* [WordNet]
- Information is the key element
- Media can be loosely defined as any information in electronic form

What is Media?

- What is Media?
  - A plural of medium [Dictionary]
  - Formats for **presenting information** [Wiki]
  - a means or instrumentality for *storing* or *communicating information* [WordNet]
- Information is the key element
- Media can be loosely defined as any information in electronic form

So, what holds information?

What is Multimedia?

- What are the different media that hold information?
  - text
  - graphics (the digital representation of an *imaginary scene*)
  - audio / sounds
  - images (the digital representation of a *real scene*)
  - videos (moving images or graphics)

What is Multimedia?

- What are the different media that hold information?
  - text
  - graphics (the digital representation of an *imaginary scene*)
  - audio / sounds
  - images (the digital representation of a *real scene*)
  - videos (moving images or graphics)
- What is Multimedia?
  - the ability to combine text, graphics, audio, and (moving) images in meaningful ways.
  - Probably this is one of the powerful aspect of computing technology

Multimedia - Components, Aspects

- Components
  - text
  - audio
  - images (sequence of images - video)

Multimedia - Components, Aspects

- Components
  - text
  - audio
  - images (sequence of images - video)
- Some Aspects
  - Storage / Compression (mp3, jpg, mpeg)
  - Transmission (codec, mpeg4)
  - Multimedia signal processing (including annotation)
  - Search / Retrieval
  - Algorithms, Architecture, Software Development

More Aspects

- Real-time Multimedia, 3D and Motion, Multiple camera System

More Aspects

- Real-time Multimedia, 3D and Motion, Multiple camera System
- Visualization, Interactive Media and Games, 3D-TV, Stereo Systems,

More Aspects

- Real-time Multimedia, 3D and Motion, Multiple camera System
- Visualization, Interactive Media and Games, 3D-TV, Stereo Systems,
- Content Management, Watermarking, DRM

More Aspects

- Real-time Multimedia, 3D and Motion, Multiple camera System
- Visualization, Interactive Media and Games, 3D-TV, Stereo Systems,
- Content Management, Watermarking, DRM
- Databases and File Systems, System Integration

More Aspects

- Real-time Multimedia, 3D and Motion, Multiple camera System
- Visualization, Interactive Media and Games, 3D-TV, Stereo Systems,
- Content Management, Watermarking, DRM
- Databases and File Systems, System Integration
- Internet Telephony, Audio/video Streaming

More Aspects

- Real-time Multimedia, 3D and Motion, Multiple camera System
- Visualization, Interactive Media and Games, 3D-TV, Stereo Systems,
- Content Management, Watermarking, DRM
- Databases and File Systems, System Integration
- Internet Telephony, Audio/video Streaming
- Multimedia Content Distribution, Wireless Multimedia

More Aspects

- Real-time Multimedia, 3D and Motion, Multiple camera System
- Visualization, Interactive Media and Games, 3D-TV, Stereo Systems,
- Content Management, Watermarking, DRM
- Databases and File Systems, System Integration
- Internet Telephony, Audio/video Streaming
- Multimedia Content Distribution, Wireless Multimedia
- Artificial Intelligence in Multimedia Technologies

Trend in Multimedia (1)

The next wave of multimedia could be in terms of creating new approaches for the

- acquisition,

Trend in Multimedia (1)

The next wave of multimedia could be in terms of creating new approaches for the

- acquisition,
- processing, and

Trend in Multimedia (1)

The next wave of multimedia could be in terms of creating new approaches for the

- acquisition,
- processing, and
- display of new types of image

Trend in Multimedia (1)

The next wave of multimedia could be in terms of creating new approaches for the

- acquisition,
- processing, and
- display of new types of image

The ultimate goal is to

Trend in Multimedia (1)

The next wave of multimedia could be in terms of creating new approaches for the

- acquisition,
- processing, and
- display of new types of image

The ultimate goal is to

- *change the way visual information is captured from real scenes* and

Trend in Multimedia (1)

The next wave of multimedia could be in terms of creating new approaches for the

- acquisition,
- processing, and
- display of new types of image

The ultimate goal is to

- *change the way visual information is captured from real scenes* and
- *presented to the human observer*

Example: Trend in Multimedia (2)

- depth-perception,

Example: Trend in Multimedia (2)

- depth-perception,
- visualization without aids (eg. virtual reality glasses),

Example: Trend in Multimedia (2)

- depth-perception,
- visualization without aids (eg. virtual reality glasses),
- offering the viewer the capacity to visually *fly around* an object or a scene

Example: Trend in Multimedia (2)

- depth-perception,
- visualization without aids (eg. virtual reality glasses),
- offering the viewer the capacity to visually *fly around* an object or a scene
- view a complete sphere of view at multiple points in a natural scene, while preserving photo-realistic quality

Example: Trend in Multimedia (2)

- depth-perception,
- visualization without aids (eg. virtual reality glasses),
- offering the viewer the capacity to visually *fly around* an object or a scene
- view a complete sphere of view at multiple points in a natural scene, while preserving photo-realistic quality

Achieving this goal will rely on rendering visual information generated by using input images captured from *real* scenes

Trend in Multimedia (3)

A challenging research objective is to look at efficient delivery of multimedia applications and services over emerging diverse and heterogeneous wireless networks

- 2.5G/3G/4G
- Wi-Max,
- WLANs,
- adhoc

Trend in Multimedia (4)

There are several research issues

- encoding techniques (New audio/video coding techniques over wireless networks)

Trend in Multimedia (4)

There are several research issues

- encoding techniques (New audio/video coding techniques over wireless networks)
- cross-layer optimization,

Trend in Multimedia (4)

There are several research issues

- encoding techniques (New audio/video coding techniques over wireless networks)
- cross-layer optimization,
- error control - Protocols for efficient delivery of multimedia applications and services over mobile networks

Trend in Multimedia (4)

There are several research issues

- encoding techniques (New audio/video coding techniques over wireless networks)
- cross-layer optimization,
- error control - Protocols for efficient delivery of multimedia applications and services over mobile networks
- security - Watermarking

Trend in Multimedia (4)

There are several research issues

- encoding techniques (New audio/video coding techniques over wireless networks)
- cross-layer optimization,
- error control - Protocols for efficient delivery of multimedia applications and services over mobile networks
- security - Watermarking
- Performance evaluation of multimedia services via analysis, simulation, and experiments

Trend in Multimedia (5)

Internet and wireless access technologies have opened up new opportunities to serve high quality, on-demand multimedia applications

Trend in Multimedia (5)

Internet and wireless access technologies have opened up new opportunities to serve high quality, on-demand multimedia applications

Content distribution and consumption landscape has changed because of

Trend in Multimedia (5)

Internet and wireless access technologies have opened up new opportunities to serve high quality, on-demand multimedia applications

Content distribution and consumption landscape has changed because of

- mobile TV,
- IPTV,
- on-demand streaming,
- peer-to-peer video sharing

Trend in Multimedia (5)

Internet and wireless access technologies have opened up new opportunities to serve high quality, on-demand multimedia applications

Content distribution and consumption landscape has changed because of

- mobile TV,
- IPTV,
- on-demand streaming,
- peer-to-peer video sharing

Seamless mobile access, and any time anywhere media consumption is the goal for complete freedom.

Trend in Multimedia (6)

- New models, metrics, and methodologies in both source, and channel coding,

Trend in Multimedia (6)

- New models, metrics, and methodologies in both source, and channel coding,
- Advances in distributed and collaborative communications (to intelligently adapt the multimedia content to suit user preferences, meet device and network constraints, and achieve better communication resource utilization)

Trend in Multimedia (6)

- New models, metrics, and methodologies in both source, and channel coding,
- Advances in distributed and collaborative communications (to intelligently adapt the multimedia content to suit user preferences, meet device and network constraints, and achieve better communication resource utilization)
- Distributed resource allocation and management for multimedia communication

Trend in Multimedia (6)

- New models, metrics, and methodologies in both source, and channel coding,
- Advances in distributed and collaborative communications (to intelligently adapt the multimedia content to suit user preferences, meet device and network constraints, and achieve better communication resource utilization)
- Distributed resource allocation and management for multimedia communication
- Game theoretical models for multimedia contents interactions

Example

Sharing multimedia content among networked devices and even *handing over* from one device to another.

Example

Sharing multimedia content among networked devices and even *handing over* from one device to another.

Example

Sharing multimedia content among networked devices and even *handing over* from one device to another.

Media playback, a task that can be handed over from a mobile MP3 player to a hi-fi system in a living room

Interactive Multimedia in Education

People only retain 20% of what they see and 30% of what they hear. But they remember 50% of what they see and hear, and as much as 80% of what they see, hear, and do simultaneously.
-Computer Technology Research, 1993

Interactive Multimedia in Education

People only retain 20% of what they see and 30% of what they hear. But they remember 50% of what they see and hear, and as much as 80% of what they see, hear, and do simultaneously.
-Computer Technology Research, 1993

Multimedia when integrated with links and tools that let the user navigate, interact, and communicate with the computer makes it very powerful and interactive.

Interactive Multimedia in Education

People only retain 20% of what they see and 30% of what they hear. But they remember 50% of what they see and hear, and as much as 80% of what they see, hear, and do simultaneously.
-Computer Technology Research, 1993

Multimedia when integrated with links and tools that let the user navigate, interact, and communicate with the computer makes it very powerful and interactive.

Interactive Multimedia is a great teaching tool. Retention rates increase by 25% to 50%.

Interactive Multimedia in Education

- Interactive multimedia changes the nature of reading

Interactive Multimedia in Education

- Interactive multimedia changes the nature of reading
- Instead of limiting to the linear presentation of text as printed in books, multimedia makes reading dynamic

Interactive Multimedia in Education

- Interactive multimedia changes the nature of reading
- Instead of limiting to the linear presentation of text as printed in books, multimedia makes reading dynamic
- Text convey meaning but additionally serve as triggers for further learning

Interactive Multimedia in Education

- Interactive multimedia changes the nature of reading
- Instead of limiting to the linear presentation of text as printed in books, multimedia makes reading dynamic
- Text convey meaning but additionally serve as triggers for further learning
- accomplished not only by providing more text but by bringing it to life with sound, pictures, music, and video.

Examples

- CyberMath Animation, Plug-in, VRML (3D)

Examples

- CyberMath Animation, Plug-in, VRML (3D)
- Yahooligans - Several topics explained with complete class notes

Examples

- CyberMath Animation, Plug-in, VRML (3D)
- Yahooligans - Several topics explained with complete class notes
- Scientific American Easy to understand science

Examples

- CyberMath Animation, Plug-in, VRML (3D)
- Yahooligans - Several topics explained with complete class notes
- Scientific American Easy to understand science
- Discovery Channel On-Line Latest and greatest about the world we live in

Examples

- CyberMath Animation, Plug-in, VRML (3D)
- Yahooligans - Several topics explained with complete class notes
- Scientific American Easy to understand science
- Discovery Channel On-Line Latest and greatest about the world we live in
- National Geographic Good multimedia - RealAudio

Examples

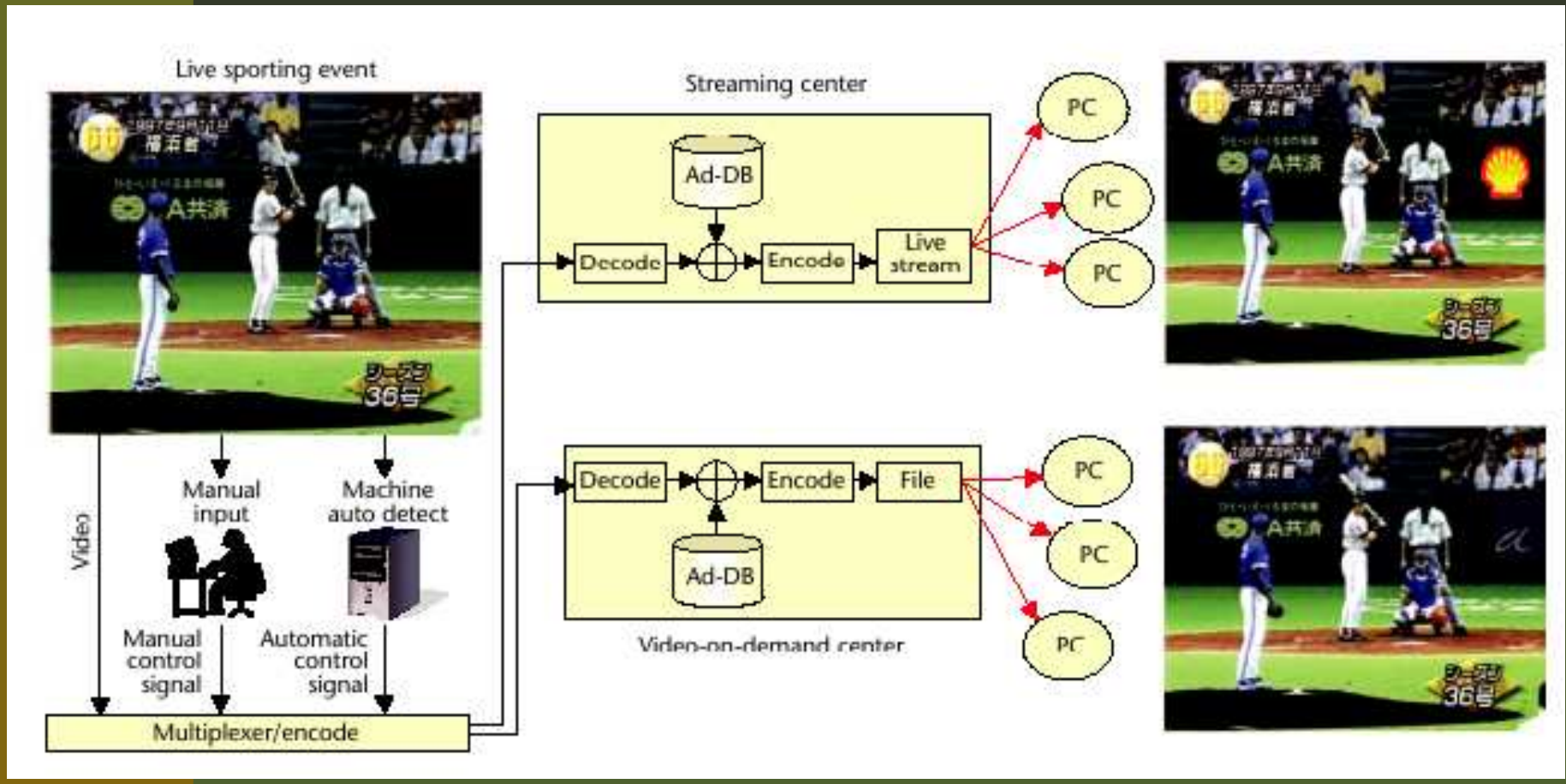
- CyberMath Animation, Plug-in, VRML (3D)
- Yahooligans - Several topics explained with complete class notes
- Scientific American Easy to understand science
- Discovery Channel On-Line Latest and greatest about the world we live in
- National Geographic Good multimedia - RealAudio
- Wiki (user participation)

Multimedia at Work

Advertising Insertion in Sports Webcasts (IEEE Multimedia)

Multimedia at Work

Advertising Insertion in Sports Webcasts (IEEE Multimedia)



Advertising Insertion in Webcasts

- Webcasts can reach a demographic segment.

Advertising Insertion in Webcasts

- Webcasts can reach a demographic segment.
- Webcasts combine traditional media's familiarity with the Internet's one-to-one interactivity.

Advertising Insertion in Webcasts

- Webcasts can reach a demographic segment.
- Webcasts combine traditional media's familiarity with the Internet's one-to-one interactivity.
- Webcast viewers are likely to be more tolerant of extraneous video effects such as advertising insertions than TV viewers because webcasts' video quality is generally lower than TV's.

Advertising Insertion in Webcasts

- Webcasts can reach a demographic segment.
- Webcasts combine traditional media's familiarity with the Internet's one-to-one interactivity.
- Webcast viewers are likely to be more tolerant of extraneous video effects such as advertising insertions than TV viewers because webcasts' video quality is generally lower than TV's.
- Webcast audiences are generally more technologically savvy, affluent, and likely to spend money on advertised items than TV audiences.

Multimedia at Work

- Multimedia Video Blog (Hosted on IEEE Multimedia website)

Multimedia at Work

- Multimedia Video Blog (Hosted on IEEE Multimedia website)

- New mode of knowledge distribution

Multimedia at Work

- Multimedia Video Blog (Hosted on IEEE Multimedia website)

- New mode of knowledge distribution

How do we access this kind of knowledge?

Enabling Easy Access to Multimedia

- Universal Multimedia Access (UMA)

Enabling Easy Access to Multimedia

- Universal Multimedia Access (UMA)
- adaptively deliver multimedia content to users according to their needs (device, context, or preferences)

Enabling Easy Access to Multimedia

- Universal Multimedia Access (UMA)
- adaptively deliver multimedia content to users according to their needs (device, context, or preferences)
- Central to UMA is the development of meta-data standards for describing multimedia resources

Enabling Easy Access to Multimedia

- Universal Multimedia Access (UMA)
- adaptively deliver multimedia content to users according to their needs (device, context, or preferences)
- Central to UMA is the development of meta-data standards for describing multimedia resources

Need to have a universal standard to enable collaborative tagging of Multimedia - this will enable easy access to multimedia knowledge base. *Manual Tagging.*

Enabling Easy Access to Multimedia

- Universal Multimedia Access (UMA)
- adaptively deliver multimedia content to users according to their needs (device, context, or preferences)
- Central to UMA is the development of meta-data standards for describing multimedia resources

Need to have a universal standard to enable collaborative tagging of Multimedia - this will enable easy access to multimedia knowledge base. *Manual Tagging.*

Automatic Tagging needs media understanding!

Basics: Information in Speech

- **non-linguistic**, (*who said it*)
  gender, emotional states, speaker name

Basics: Information in Speech

- **non-linguistic**, (*who said it*)
  gender, emotional states, speaker name
- **linguistic** (*what he said*)
  Language name and what was said (written language)

Basics: Information in Speech

- **non-linguistic**, (*who said it*)
  gender, emotional states, speaker name
- **linguistic** (*what he said*)
  Language name and what was said (written language)
- **paralinguistic** (*how well said* – manner, clarity or accent, aspects related to quality)
  deliberately added by the speaker, and not inferable from the written text.

Basics: Information in Speech

- **non-linguistic**, (*who said it*)
  gender, emotional states, speaker name
- **linguistic** (*what he said*)
  Language name and what was said (written language)
- **paralinguistic** (*how well said* – manner, clarity or accent, aspects related to quality)
  deliberately added by the speaker, and not inferable from the written text.

Goal: Automatically extract information in speech signal

Audio Understanding

First step is to convert audio signal into searchable text and then tag it.

- Speech Recognition (recognize text from audio)
- Speaker Recognition
- Accent Recognition
- Webcast transcriptions
- Audio songs (+ background music) to lyrics
- emotion recognition

Image Understanding for Tagging

Understanding usually attempts to mimic the human visual system in extracting meaning from an image

Image Understanding for Tagging

Understanding usually attempts to mimic the human visual system in extracting meaning from an image
Given a reason for looking at a particular scene, understanding systems produce descriptions of (a) the image and (b) the world scene that the image represent.

Image Understanding for Tagging

Understanding usually attempts to mimic the human visual system in extracting meaning from an image
Given a reason for looking at a particular scene, understanding systems produce descriptions of (a) the image and (b) the world scene that the image represent.




Image Understanding for Tagging

Understanding usually attempts to mimic the human visual system in extracting meaning from an image
Given a reason for looking at a particular scene, understanding systems produce descriptions of (a) the image and (b) the world scene that the image represent.





System should be able to describe all these images as say **red car**.

Media Retrieval

- Data is being generated at an unbelievable pace (in an organization and on the web; database and unstructured info-base).
- Obtaining *useful* information that one is looking for from this ever growing data has become very crucial world over.
- Requirement
  - The information has to be retrieved Quickly and Precisely.
  - Advantage if there is no constraint on the way the information is asked for (natural language).

Why Natural Language Processing?

- Human interaction with computers is very constrained!

  anything we want (ideas) the computers to do we first need to translate (our ideas) into a form understandable by the computer

- Natural Language Processing technology enables humans to use computers in much the same way that they would use a human assistant to get their work done

  without having to talk to computers in a predefined manner

Evolution of Website

A website is an effective way of show casing offerings, products, technology ...

- Set of hyperlink documents (not really organized)
- Hyperlinked documents (major design effort)
- Conventional search engines (how to search tips!) have found their way onto web pages

Why? ... enable their customers access information they are looking for as quickly and as easily as possible

Right Step .... Is it good enough?

- Search engines require the user to choose one or several key words.
- The search engines presents to the user (one or several documents) which contains the keywords the user has asked for.
- The user now has to go through several hyperlink clicks and search manually to converge to the queried information.

Uhmm... need to search the search engine results

Drawbacks?

- User needs to remember keywords the system knows!
- Keyword ambiguity (Bank - Money, River?)
- Need to manually browse *(search!)* and hyperlink click to reach the information (he is looking for)
- Good probability that the document containing the search word doesn’t contain the information one is looking for!

Very restrictive and not very friendly ;-)

Beyond keyword search - What?

- Enable user to ask for information from a particular website in natural *(free style)* English.
- Understand the *intent* of the query rather than constraining the user to use a set of predefined keywords
- Provide nearest useful information in the event of exact information not being present.
- Should not require further browsing for the exact information

.. enabled using Natural Language Processing (NLP) techniques

At Tata Consultancy Services

- Aim to develop a *(generic)* Natural Language Processing based system which can understand the intent of the query posed in natural language.
- Use this understanding to query data from either a structured database or unstructured info-base.
- Give most appropriate answers when exact answers are not present or available

Ongoing Research and Development activity

Information Retrieval Systems for ....

- Textual Info-base
  - Corporate Websites
- Conventional Database
  - Airline; Yellow Pages; Railway
- Ebooks
  - Health Fitness
  - Technical / Educational(Management Information Systems)
- Multimedia
  - Ring-tone or song download
  - Wall-paper or image, video download

Advantage of NLP: Example

Task: *Can you give me the location address of your Chennai branch office?*

Current 2 clicks, 1 search. Go to the branchlocator.htm page (click), search (manual search) for a location (say Chennai) and then click on the Chennai link (click), to see the information

NLP One click - Address of Chennai branch displayed

Advantage of NLP: Example

"What are the facilities for passengers with restricted mobility?"
Currently, it would require a user to first click the navigation bar related to Products and services ... then search for a link, say, On ground Services .... browse through all the information on that page and pick out relevant information manually.

NLP understands what the user wants ... it is capable of picking up and displaying only the relevant paragraph of information ... *saving the time of the user also saving the user from seeing information that he is not looking for!*

Advantage of NLP: Example

- Ring-tone download
  - Lata Mangeshkar’s mere watan ke log
  - Amitabh and Rani song in Black
  - Title song of Shivaji
- Wall paper download
  - Red car on the road
  - Sun at dawn
  - Taj Mahal in moon light
  - Raj Kapoor as joker

Advantage of using NLP

- Saves time for the information seeker by giving focused and exact answers
- If no suitable answer found ... displays *nearest* answer by understanding the intent of the question
- Enables user to ask question in the preferred Natural Language
- Faster access to information in terms of reduced number of clicks and manual search ....

Multimedia in Indian Scenario

- Multilinguism in Multimedia
- Applications for the masses (to impart multi-sensory knowledge to the users)
  - applications enriched with high quality imageries
  - immersive 3D environments,
  - animated scenes
  - synchronized digital sound tracks,
  - interactivity

Places to catch up ...

- Journals
  - Multimedia Tools and Applications
  - Journal of Multimedia (JMM) http://www.academypublisher.com/jmm/
  - Journal of Educational Multimedia and Hypermedia
  - IEEE Multimedia
  - Advanced in Multimedia
- Conferences
  - International Conference on Multimedia & Expo
  - etc ...
- World Wide Web

Thank You

SunilKumar.Kopparapu@TCS.Com
TCS Innovation Labs - Mumbai
Advanced Technology Application Group
Yantra Park, Thane (West), Maharastra 400 601
http://www.tcs.com

