Dix et al. Chapter 21 is on hypertext, multimedia and the WWW. Sheiderman has a bit on multimedia document searches on pages 519-521. There is a chapter (incl. 3 papers on the Wang Freestyle system) in the Baecker et al. volume, pp. 867-896.
A Definition

“Multimedia is the use of several different media to convey information (text, audio, graphics, animation, video, and interactivity) … As the information is presented in various formats, multimedia enhances user experience and makes it easier and faster to grasp information … Presenting information in various formats is nothing new, but multimedia generally implies presenting information in various digital formats.”

→ Accessing information using more than one input channel (human sense) improves processing (understanding, remembering, learning, ...).
Images from Wikipedia

different types of information:
  • text (ASCII, formatted, HTML)
  • Images/graphics (photographs, graphics)
  • Audio/sound (earcons, speech, music)
  • video
  • animation (avatars)

→ Using multiple media in a senseful combination can make an interface or system more appealing…

For digital multimedia, physical storage is the same for all media. But media are encoded differently, which influences the type of object and level of complexity at which it can be represented. Therefore, different media are suitable for communicating different types of information.
Interactive Multimedia

“The user is in control “

- Distinction between linear and non-linear interactivity
  - (e.g. ppt presentation vs. computer games)
  - E-learning applications (sequential or explorative)

Illustration taken from: http://wonders.eburg.wednet.edu/Topics/GMMIE/grandTour.htm
Multimedia in e-learning example:
Concurrent display of medical images to facilitate learning for med-students (Harry Brenton, Imperial)
Characteristics of media

• **Richness**
  - abstract vs. concrete (icon vs. photo)
  - focus vs. context (animation vs. video)
  - static vs. dynamic (photo vs. video)

• often asserted that rich media are always better
  - Which is not generally true, as the usage of media depends on task context

• Higher quality (size, resolution) increases processing and transmission requirements
  → delay (on the Web, editing HD video, …)
Media type and use

• More complex information leads to bigger data objects
• Communication of these requires bigger channels
• i.e. more “capacity” or “bandwidth”

From:
Internet Multimedia by Andy Sloane.
Published by Palgrave-Macmillan ISBN 1403934134
Linear nature of text can be hindrance, especially for searching; compiling of indices is very time-consuming

- Visual representation of language
- Tool to communicate
- Graphic element → Typography
- Text is traditionally linear
- Regarded as more formal and authoritative (than other media)
- Synchronous vs. asynchronous usage scenarios:
  - E.g. chat vs. book
Text

SMS "language"

Adapting text-based communication to context:

- BRB       Be right back
- CUL8R     See you later
- CW2CU     Can’t wait to see you
- FYI       For your information
- IYSS      If you say so
- LMAO      Laugh my ass off
- OIC       Oh I see
- PCM       Please call me
- YTLKIN2ME Are you talking to me
- 1STY      Thirsty

-SMS usually offer restricted amount of information (160 characters)
-The input using a standard phone is tedious (even with T9)

See: http://www.txt2nite.com/ for more examples...
- The linear nature of text has been undermined through the integration of hyperlinks
- Linear structures disappear more and more in computer-based systems in favour of more complex structures
- With hypertext, interaction in terms of manipulation of the presented information became possible

Ted Nelson discovered the concept of hypertext, influencing several developers of the Internet, most notably Tim Berners-Lee. [...] As he considered the design of this system, Nelson applied his experience as a filmmaker with the conception of complex motion picture effects, moving from one shot to another, and conceived of the idea of hypertext. He became profoundly convinced of the enormous value of such a system, and has been thinking and talking about it ever since. Nelson's first job was as a photographer and film editor at a Miami laboratory where John Lilly was carrying out research on the intelligence of dolphins, using LINC microcomputers to analyze their talking, as fascinated by acoustics as J.C.R. Licklider. Nelson then moved to a job teaching sociology at Vassar College. The word "hypertext" was first coined by Nelson in 1963, and is first found in print in a college newspaper article about a lecture he gave called "Computers, Creativity, and the Nature of the Written Word" in January, 1965.
• Very abstract – purely text-based – abstraction of interaction
• Complicate to use for everyday user
• Very restricted functionality
Windows Vista (or other visually rich interfaces) → Reduced cognitive effort

• Visually rich interface
• Uses a desktop metaphor
• More and more photo-realistic icons (see the recycle bin)
• Animation of menus
• Anti-aliased text to simulate printed text (like in PDF files)

→ Combination of text, graphics, photos and animation to optimize interaction (to use the computer and „all“ its functionality)
Images

- Range from rich (photos) to very sparse representations
- Colour fidelity important for some applications (art, fashion, medical)
- High resolution and size may be important (maps)
- Big image files (still) slow down web applications
  - With growing bandwidth, size, quality and number of images on the Web increases dramatically

Pictures from: http://umbra.nascom.nasa.gov/images/latest.html
Images on the Web

- Tool for visual communication
  - Avatars, portraits, galleries, news, ...
- 1 million MMS per day in the UK
- "Flickr serves hundreds of millions of photos each day. On the highest traffic days, just over a billion photos are served" (Flickr developer Eric Costello)
  - Approx. 12,000 per second viewed
  - Approx. 1 million uploaded each day

Statistics from:
If authors have to use a range of editors for different media, those tools should be similar/consistent. Transferring data between tools should be as easy as poss. (without converting file formats, or using cumbersome import/export procedures).

“The largest cost in producing a hypermedia application is authoring. The high cost results from the time spent by a person experienced in the chosen area, selecting and collating the information, and then manually linking the data into a cognitive and pedagogical structure that is easy to navigate. Therefore as much of the process of authoring needs to be automated, thereby reducing the time required of the ‘expert’ to link the hypermedia.”

“To enable automatic re-linking of a modified document, where possible try to make use of local and generic links. Generic linking is used for explicit linking i.e. procedure names or numbers, part descriptions, etc. In generic linking the link is formed on the object and not on the location of the object, an object being a word, collection of pixels, a vector, a segment of audio or a frame in a video file. Hence, after an updating of a document, providing that the object is still in the document it will still be linked even if its position within the document has changed. Yet, in practice most links will be button (point-to-point).”

Wills et al., 1997
Other Multimedia Searches:

- **Design/Diagram Searches**
  - Some computer-assisted design packages support search of designs
  - Allows searches of diagrams, blueprints, newspapers, etc.
    - E.g. search for a red circle in a blue square or a piston in an engine
  - Document-structure recognition for searching newspapers

- **Sound Search**
  - MIR supports audio input
  - Search for phone conversations may be possible in future on speaker independent basis

- **Video Search**
  - Provide an overview
  - Segmentation into scenes and frames
  - Support multiple search methods
  - Infomedia project

- **Animation Search**
  - Prevalence increased with the popularity of Flash
  - Possible to search for specific animations like a spinning globe
  - Search for moving text on a black background
Check out: http://labs.systemone.at/retrievr/?sketchName=2006-02-01-22-36-54-688090.17
http://www.xcavator.net/index.html
http://www.like.com/
Audio

- separate channel
  - Melodies, sounds, songs
  - Voice
  - Auditory icons (familiar sounds) & earcons (abstract sounds – meaning to be learnt)

- good-quality audio can enhance user experience
  - high quality audio improves perception of overall system
  - convey personality of speaker
  - but: very irritating if not done well

- quality of speakers, mikes etc. can have impact on audio
Audio
Speech and Auditory Interfaces

- Speech recognition still does not match the fantasy of science fiction:
  - background noise problematic
  - variations in user speech performance impacts effectiveness
  - most useful in specific applications, such as to benefit handicapped users
- more demanding for user than hand/eye coordination (in terms of cognitive load)

See chapter 9.4 in Shneiderman.
Discrete word recognition
- recognize individual words
- Very limited vocabulary
- Mainly speaker-dependent training
- Very helpful for paralyzed people

Continuous-speech recognition
- recognizing boundaries between spoken words difficult

Speech store and forward
- Voice mail services
Audio
Speech and auditory interfaces

• Voice information systems
  – Low cost
  – Hardly user friendly
  – Audio books

• Speech generation
  – For short and simple messages
  – Helps user to focus on other tasks (e.g. operation)

• Audio tones, audiolization, and music
  – Sound feedback

Voice information systems
• Stored speech commonly used to provide information about tourist sites, government services, after-hours messages for organizations
• Low cost
• Voice prompts
• Deep and complex menus frustrating
• Slow pace of voice output, ephemeral nature of speech, scanning and searching problems
• Voice mail
• Handheld voice recorders
• Audio books
• Instructional systems

Speech generation
• Michaelis and Wiggins (1982) suggest that speech generation is "frequently preferable" under these circumstances:
  • The message is simple.
  • The message is short.
  • The message will not be referred to later.
  • The message deals with events in time.
  • The message requires an immediate response.
  • The visual channels of communication are overloaded.
  • The environment is too brightly lit, too poorly lit, subject to severe vibration, or otherwise unsuitable for transmission of visual information.
  • The user must be free to move around.
  • The user is subjected to high G forces or anoxia

Audio tones, audiolization, and music
• Sound feedback can be important:
  • To confirm actions
  • Offer warning
  • For visually-impaired users
  • Music used to provide mood context, e.g. in games
  • Can provide unique opportunities for user, e.g. with simulating various musical instruments
Animation

- Can be used for:
  - attracting attention
  - showing continuity in transitions
  - indicating dimensionality in transitions
  - illustrating change over time
  - visualizing three-dimensional structure

- Avatars (animated people or objects)
  - Representation of real person or virtual character
    (e.g. second life vs. Microsoft’s ‘retired’ office assistant)
  - fidelity of appearance and behaviour needs to tie up

Avatars: Garau et al, CHI 2003
A quick Salsa lesson...
And a small guide to knitting…

http://knittingclub.yokolab.com/elesson/e-scarf1.html
Flash
for animations and complete UIs

- Seamless interaction and navigation (also AJAX)
- Flash animations for novel interfaces, integration of animations and video
- Flash Video provides much more integrated experience of video on the Web and pushed the distribution and usage enormously (Youtube, Google Video, ...)
- Flash interfaces for mobile devices (e.g. Samsung)
- Desktop applications with Adobe AIR

- Lots of mobile devices (over 300) now feature Flash capabilities
- Prada phone has Flash UI
- MP3 player with Flash UI (iRiver)
- ...
Flash Interfaces
Samsung F700 with Flash Lite UI
Video

- showing things that move
- good quality requires high resolution or bandwidth
- passive viewing has higher quality requirements than active viewing (focus of attention)
- convey speakers’ personality
  - increase trust
- production quality often too low
- size in proportion of distance from display

This is based on Jacob Nielsen’s alertbox column
http://www.useit.com/alertbox/9512.html

In a later column, Nielsen points out that “Flash is 99% bad” because it encourages designers to go against these principles
Both columns are highly recommended reading.
http://www.rocketclips.com/collections.cfm?ResetAll=1
- In a business environment
- Distance collaboration
- Online teaching
- Personal communication (webcam chats)
A different kind of video conference...

As soon as your video conference with the customer is over, I’d like a word with you Ted.
Applications for Multimedia

- Multimedia Art
  - Installations, Flash, Video art, Interactive events …
- Education
  - Computer based learning, training material …
- Entertainment
  - TV, radio, games, chats …
- Edutainment
  - form of entertainment designed to educate as well as to amuse
  - instruct or socialize audience by embedding lessons in some familiar form of entertainment
About Start Mobile (taken from Website):

December 12, 2005 – San Francisco, CA. START SOMA, the San Francisco gallery for emerging artists, today announced the launch of START MOBILE, becoming the first retail art gallery in the world to sell NEW ART for cell phones. Curated by START SOMA founder John Doffing, THOUSANDS of original works of NEW ART from HUNDREDS of the world’s most prominent emerging + underground artists can now be downloaded onto mobile phones across the United States

http://www.STARTmobile.net

“Cell phone wallpaper has long been limited to banal graphics, contrived iconography, and blurry clip-art of models in bikinis,” according to Doffing. “With the launch of START MOBILE, we are bringing some absolutely amazing original artwork to a brand new medium.”

HUNDREDS OF ARTISTS + THOUSANDS OF IMAGES

START MOBILE showcases the world’s premier emerging + underground artists, quite literally from A to Z. “START MOBILE features artists whose work is in museums + artists whose work appears exclusively on the streets - and pretty much everything in between,” according to Doffing. “We have created an unprecedented mix of creative talent from around the world, and their art is now available to anyone with a cell phone for less than the price of a cup of coffee.”
MULTIMEDIA

PLEIX
digital artists community

See:
http://pleix.net/films.html
ABSTRACT
Prompt feedback is essential for beginning violin learners; however, most amateur learners can only meet with teachers and receive feedback once or twice a week. To help such learners, we have attempted an initial design of Digital Violin Tutor (DVT), an integrated system that provides the much-needed feedback when human teachers are not available. DVT combines violin audio transcription with visualization. Our transcription method is fast, accurate, and robust against noise for violin audio recorded in home environments. The visualization is designed to be intuitive and easily understandable by people with little music knowledge. The different visualization modalities—video, 2D fingerboard animation, 3D avatar animation—help learners to practice and learn more effectively. The entire system has been implemented with off-the-shelf hardware and shown to be practical in home environments. In our user study, the system has received very positive evaluation.
„Turning the pages“ is a project by the British Library that allows the user to leaf through 15 great books and magnify details.

Turning the Pages allows visitors to virtually 'turn' the pages of manuscripts in a realistic way, using touch-screen technology and interactive animation. They can zoom in on the high-quality digitised images and read or listen to notes explaining the beauty and significance of each page. There are other features specific to the individual manuscripts. In a Leonardo da Vinci notebook, for example, a button turns the text round so visitors can read his famous 'mirror' handwriting.

(see: http://www.bl.uk/onlinegallery/ttp/ttpbooks.html)
First Atlas of Europe
Compiled by Mercator in the 1570s

The British Isles map was composed using extracts from at least two copies of Mercator's 1571 map of Europe. One of Mercator's principle sources was almost certainly his 1571 map of the British Isles, created in 1570 for the Earl of Leicester. At this time, English and Scottish mapmakers had made considerable advances in mapping the British Isles in the 16th century; there were no printers in England to make their maps available in print. However, maps of the British Isles were often kept in manuscript form, and were often kept under lock and key by the government. Such sources of information could be very
Multimedia document searches

- Searches for databases and textual documents are good, but multimedia searches are in a primitive stage
- Locating of images, video, audio or animation much more difficult than search text or structured database
- Current multimedia searches require descriptive documents or metadata searches
  → Image tagging in Facebook or Flickr good example of potential solution (but requires manual effort)
- Search by date, text captions, or media is possible

Searching in mm documents/libraries still difficult - see Shneiderman pp. 519-521.

Structured databases or document search for media: e.g. by name of photographer, actor, etc.

photos: query by content (QBIC):colours, patterns, shapes
  - works best on highly specific/homogeneous images
video: similar, but more difficult; search by object/person
audio: words, sounds, tunes
  - automatic speech-to-text transcription
Multimedia document searches
Map Search

- Search by latitude/longitude is the structured-database solution
- Today’s maps are utilizing structured aspects and multiple layers
  - City, state, and site searches
  - Flight information searches
  - Weather information searches
  - Example: www.mapquest.com
- Mobile devices can allow “here” as a point of reference
Map Search
Google Maps
Mash-ups
user generated multimedia

"A **mashup** is a website or application that combines content from more than one source into an integrated experience."
(Wikipedia)

- information and presentation are being separated in ways that allow for novel forms of reuse
- Linking content from various resources to create new "services" (mainly RSS), e.g.:
  - text from blogs, search queries, etc.
  - Images from flickr & co
  - Audio from last.fm ...

---


See also:

…”Those new abilities allow the ‘bot to listen to anything his owner says to him about, say, their day just gone, process what he hears and then use the keywords to search the internet for related multimedia content.

After that, whatever online goodness PaPeRo digs up gets dumped onto the owner’s website along with a video recording of the conversation that prompted the cybernetic creativity. All the human blogger has to do is to feed PaPeRo the blog login details and a few handfuls of batteries every day…”
Summary points

(1) Select the right media for communicating different types of information.
(2) Redundancy can be useful to support understanding but can also lead to opposite effects if done bad or wrong.
(3) Dialogue styles and input/output devices must match media users interact with, but trade-off with complexity of interaction