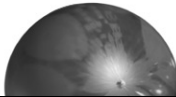




Multimedia IR in Context

Sharon Flank, CTO
eMotion, Inc.
sharon.flank@emotion.com



Copyright © 2001 eMotion, Inc

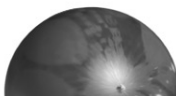


www.emotion.com • 866-2eMotion



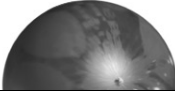
Overview

- Who uses multimedia IR?
- What is important to those users?
- Search issues
- Usability issues
- Application: digital media management



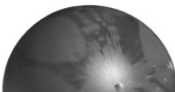


User Base

- Creative professionals
 - Kids doing homework
 - Web surfers
 - Celebrities
 - Pornography
 - Travel
 - E-commerce
- 



Requirements for Success

- Appropriate technology
 - Solvable problems
 - Follows customer's business model
 - User-centered interactions
- 



Priorities

- Speed
 - 1-2 seconds
- Usability
 - Comprehensible at a glance, usable for hours
- Adequacy
 - Some retrieval technology is adequate for sports, but not for broadcast news or feature film



Priorities, cont.

- Accuracy
 - Focus on precision over recall
 - Users' confidence is eroded if they see junk
 - Few users know what is in the collection, so recall is of less importance
 - First screen is critical
 - Users are lazy
 - "Precision at 20" may be the best measure



Multiple Search Modalities

:: text search ::

:: voice ::



:: drawing ::



:: fielded metadata ::



Advanced Search Modalities

“like this but with dark hair”

singing



natural language

“left, seated”

touch screen



Technologies Behind Search

- Voice input
 - Speech recognition
 - Advances will require moving beyond the signal and a vocabulary list, to phonemes and syntax, even discourse analysis
- Drawing
 - Template matching
 - Multidimensional feature space



Search Technologies

- Fielded metadata
 - Standards efforts for field names
 - Open-ended if possible
 - Combine with other modalities
 - Users do not like to fill out big forms
- Text search
 - Frequency-based Boolean keyword search



Search Technologies

- Singing
 - “Match that tune” audio retrieval
 - Similar to voice recognition
- Touch screen
 - Identify a region of an existing file and ask for matches (color, shape, texture)



Search Technologies

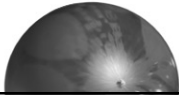
- Natural language
 - Part of speech tagging
 - Stemming, morphology
 - Pattern-matching
 - Partial or even full parsing
 - Semantic net



Natural Language Search



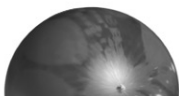
- Part of speech tagging
 - Distinguish nouns, verbs, adjectives, adverbs
 - *crane* (your neck)
 - *crane* (bird or equipment)
 - Limits spurious choices and bad synonyms
 - 99% accurate
 - Primarily statistical



Natural Language Search



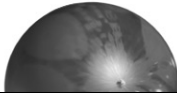
- Stemming, morphology
 - *run-s*
 - *run-ning*, not *runn-ing*
 - *ran* yields *run*



Natural Language Search



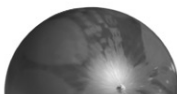
- Pattern-matching
 - Names
 - First name variants, for English and other languages:
Bill, William, Billy, Will, Willy, Guillaume
 - English name pattern: middle name or initial, *Jr.*
 - Spanish: mother's surname at end
 - Adj-noun groupings
 - *tiger shark, fire truck*
 - Different from *tiger, truck fire*



Natural Language Search



- Partial or even full parsing
 - *a mountain covered with snow*
 - *snow-covered mountain*
 - *cart pulled by a horse*
 - *horse pulling a cart*
- Complexity tradeoff



Natural Language Search

- Semantic net
 - Beyond a thesaurus
 - Start with a resource for English and then tailor it, rather than starting with nothing
 - WordNet: 100,000+ terms
 - Gazetteer, lexicons
 - A large manually-created resource may only reach 30,000 terms, with enormous maintenance overhead



Semantic net

- Multiple relationships
 - Synonymy
 - *cougar - puma - mountain lion*
 - Hyponymy
 - *bird - owl*
 - Meronymy
 - *snow - snowflake, beach - sand, car - brakes*
- Avoid full-scale AI knowledge base



Search Technologies

- Discourse-based strategies
 - “Like this but with dark hair”
 - Assumes a model of an existing query or file
 - Knowledge of where substitution should be made
 - “Left, seated”
 - Use info provided in captions
 - Srihari (various, 1991-1995)



Metadata Mantra

- When creating the metadata design, ask:
 - How are people going to search for this?
 - What are the reuse options?
- Not:
 - What categories of information can I attach?
 - What will people need to know once they have found what they want?



Usability Strategies

- Model the types of users
- Determine what functions they use
 - How often
 - For how long at a stretch
 - Often enough to remember shortcuts?



Usability

- Layer the information
 - Group functions sensibly
 - Hide functions from the top-level intro user
 - Images should not be tiny
 - Watermark as needed to protect intellectual property
 - Simple backgrounds, focus on assets



Usability

- Ask what the technology buys you
 - Will someone want your AI engine to crank through the relationship between *Barbra Streisand* and *director*, or can you just list those facts separately? (cf. MPEG-7 Semantics Description)
 - Keyframes: do you really need sophisticated strategies, or will every 30 seconds do?



Accuracy

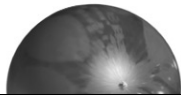
- Is face recognition mature enough to be helpful? Object recognition?
- In what constrained environments could they be useful?
 - E.g. voice recognition “Say or press 1”



What is Digital Media Management?



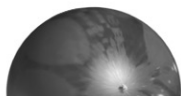
- Software that allows media companies to better manage their digital and analog media.
- Solutions that can enable effective e-commerce, link to media production systems and rights management, and provide the infrastructure for an effective digital media environment.



What does DMM do?



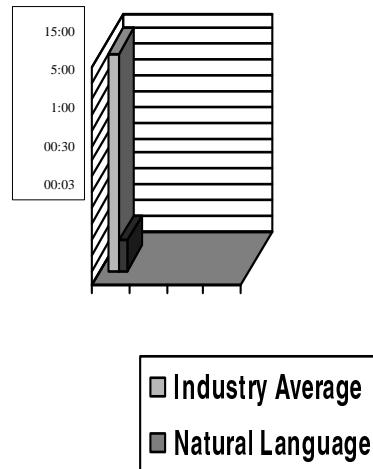
- Helps organizations that create, use or reuse digital media
- Allows clients to save money and become more productive, by helping them to more effectively
 - *Find*
 - *Collaborate*
 - *Retrieve*
 - *Protect*
 - *Re-purpose*
 - *Deliver/Distribute*their digital assets



DMM Saves Time

- Dramatically reduces the time **wasted searching** for files from **minutes to seconds**.

According to Gistics, it takes an average of 15 minutes to find a file in a file hierarchy if you don't know where it is. With Natural Language searching, it takes less than 3 seconds.



eMotion®
managing digital media