Reading and Collaboration in a Digital Age

Cathy Marshall
MSR-SVC

Portland CHIFOO
8 July 2009
first, a brief history...
the short and wondrous life of eBooks, generation 1: the excitement of multimedia

At the electrifying conclusion of his hour-long work

PLAY CONCLUSION OF THE NINTH

the audience broke out in tumultuous applause. Oblivious to everything around him, Beethoven continued to beat time after the music had stopped.

His friends on stage had to motion him to stop. They then turned him gently around so that he could at least see the ovation he had earned.

Circa 1989-1991
and the excitement of reading on a computer: e.g. the ability to do in place dictionary use
the short and wondrous life of eBooks, generation 2: mobility, retiring the pBook [sic]
Backlash!
the first two generations of eBooks fostered an atmosphere of anxiety and skepticism...

• *anxiety* in the wake of predictions of the death of text, the end of books (Coover, Bolter, ...), or even the end of reading as we know it

• *skepticism* – why would I read on a computer screen? (Gass, Birkerts, Harper & Sellen, Crawford, and at times Nielsen...)
anxiety brought by technology-driven visions of the future...
anxiety brought by technology-driven visions of the future... or not!
The second wave of eBooks provoked a profound sense of anxiety—anxiety that will require *at least* 50 years of therapy.
“Given the historic price-performance trajectories for storage, in a few years at least some high-end appliances will house hundreds, if not thousands, of books simultaneously, and certainly laptops with software book readers will house thousands or tens of thousands of books at once. *Think of portable personal digital libraries, not portable electronic books*, as the future role of these appliances.”

the short and wondrous life of eBooks, generation 3: recapturing lost innocence

Circa 2006-2009
Will the third generation of eBooks see the re-emergence of the *social* side of reading?

Why is it so hard to see what’s coming?

What did we learn from generations one and two?
Why is it so hard to see what’s coming?

Reason 1:

Changes aren’t always in the technology
the usual story about the future: the Memex from As We May Think, Vannevar Bush’s 1945 article
“The advanced arithmetic machines of the future... will have enormous appetites. One of them will take instructions from a whole roomful of girls armed with simple keyboard punches and will deliver sheets of computed results every few minutes. There will always be plenty of things to compute in the detailed affairs of millions of people doing complicated things.”

Vannevar Bush in “As We May Think”, Atlantic Monthly, 1945
a typing class from the mid-1930s...
Why is it hard to answer this question?

Reason 2:

We often don’t understand the full cost/benefit picture & who the user is
Dear NABISCO Shredded Wheat Users:

NABISCO Shredded Wheat is joining the POST family of cereals.

POST Cereals has a long tradition of making high quality, nutritious and good tasting products. In keeping with this tradition, POST is proud to welcome THE ORIGINAL SHREDDED WHEAT™ to our family of cereals.

For over 100 years, THE ORIGINAL SHREDDED WHEAT™ has been made from 100% pure whole grain wheat with no added salt or sugar. See side for information about calories and other nutrients. Each crunchy biscuit is lightly toasted for a delicious taste that’s 98% fat free.

POST is committed to following the same recipe that has provided the great taste and nutrition you have come to enjoy. In fact, the only change you’ll notice is the POST logo.

Please know your satisfaction is our most important goal.

Sincerely,
POST Cereals
Who is the user?
consider textbooks!
Why is it hard to answer this question?

Reason 3:

Reading is invisible
Nothing is more commonplace than the reading experience, and yet nothing is more unknown. Reading is such a matter of course that at first glance, it seems there is nothing to say about it.

Tzvetan Todorov, quoted by Nicholas Howe in *the Ethnography of Reading*
Observing reading:
I didn’t steal his magazine (honestly!)
In this talk, I’ll try to summarize 15 years of field studies, cooperative prototyping, and technology interventions.
Reading.

is it:

Private?

Individual?

Stationary?

Passive?

Immersive?

Soggy?
Reading
is a
stationary
information-centric
passive
immersive
individual activity
Reading is a mobile information-centric passive immersive individual activity
mobility overwhelms many other factors!

“if I’m going home to Colorado, I have to really be sure I’m going to read something if I’m going to bring it. Otherwise, why should I bring it? This thing [a handheld with e-book software], I was like, ‘I’ll bring it, and if I read it, I read it; if I don’t, I don’t.’ It doesn’t matter. It’s small, it’s handy.”

quote from a college student talking about a Pocket PC with his course texts
Reading is a mobile information-centric passive immersive individual activity
Reading is a mobile material passive immersive individual activity
“Reading what people have had to say about the future of knowledge in an electronic world, you sometimes have the picture of somebody holding all the books in the library by their spines and shaking them until the sentences fall out loose in space...”

Geoff Nunberg in “The Places of Books in the Age of Electronic Reproduction” from *Representations*

April 25, 1999. Wife finds on Alibris and puts in family minivan.
“You get this little screen, so you get no sense of even how long the work is... You have 600 pages, which means what? No-one knows. And so ... I definitely don’t see it as a literary experience.”

an English lit grad student talking about reading on the Jornada Pocket PC
navigation is fundamental to the materiality of paper

“Something else that I think that I sometimes do is an article, I’ll feel like, ‘boy, this has been going on a long time.’ ... Sometimes I’ll even flip ahead and think, ‘boy, I’m sort of tired of reading this.’ And, ‘How many more pages do I have?’ And if it’s going to end on this page, then I may just read it. But if I see it’s three more pages, then ... I may just either give up. Or just go into a scan mode, where I just flip, you know, see what grabs my attention.”
Reading has a basic physicality
“I usually read in one of the chairs in [the living room]. That’s partly because I don’t have a desk lamp in here... [The chairs are] very comfortable. They’re occasionally much too comfortable. That’s why I have blankets around every chair in the house... So I’m always comfortable. Prepared to sleep.”

-an English Lit grad student talks about why she doesn’t read in the room where she writes
Reading is a mobile material passive immersive individual activity
Reading is a mobile material interactive immersive individual activity
“The compact disc. It doesn’t wear out, even if you use it. Terrifying. It’s as though you’d never used it. So it’s as though you didn’t exist. If things don’t get old any more, then that’s because it’s you who are dead.”

Jean Baudrillard, Cool Memories II
a reader’s unselfconscious epiphany:
the text is a mirror.
oppressive tendencies. The ethnographer must question the value of the project before contributing to it.

Second, the ethnographer’s bias presents a more subtle and pragmatic ethical problem. Presenting a wide variety of options to a group or a client—even options with which the ethnographer may disagree or may dislike—is easy. However, the ethnographer must somehow control subliminal differences in presentation that reflect personal preference. The ethnographer may unintentionally focus on one alternative more than another, spending more time on the preferred choice or presenting it more cogently and convincingly. Information overload is one way—however unintentional—to dissuade people from considering an option. Hundreds of intentional and unintentional tactics are available for influencing a decision while creating the illusion of free choice.

The ethnographer can control this subtle bias once individual preferences become explicit. For example, piloting or testing significant presentations on others can help the ethnographer determine that time or verbal emphasis is the same for each topic. The use of partners to share in the presentation and place a check on how each communicates information can reduce subtle but patterned and observable manipulation. Self-discipline and self-criticism are the only tools to control the unconscious influences that will occur in casual conversations with participants.

Advocate Ethnographers

Advocate ethnography is as much a stage in a researcher’s life as it is another type of research. Advocate ethnographers allow participants to define their reality, consider their view about the ideal solution to their problems, and then take an active role in making social change happen. These ethnographers serve as advocates for the group. They write in public forums to change public opinion, embarrass power brokers, and provide relevant information about a situation at opportune moments in the policy decision-making forum.

After conducting the dropout study and determining that the programs merited continued funding, I actively disseminated the largely positive findings about the programs to individuals in government and quasi-government institutions. The team of researchers on this project prepared a Joint Dissemination Review Panel Submission substantially based on the ethnographic findings to improve the programs’ credibility and potential to secure future funding.
of Love which persists throughout the dialogue, though often little more than a figure of speech, is a further fact for which allowance must be made by an English reader, unaccustomed to such a manner of treating psychology and metaphysics. To Phaedrus the nature of Love presents no difficulties: he is the oldest of the gods, and the supreme benefactor of mankind, inspiring both a high sense of honour, because a man is particularly afraid of being detected by his lover or beloved in any mean or cowardly action, and also the spirit of self-sacrifice. These conclusions are illustrated by examples from history and mythology, and woman, in the person of Alcestis, is allowed a place in the category of those who may be led to sacrifice their lives by love.

Pausanias, though hardly more profound, is a good deal more subtle, and introduces a distinction between a nobler and a baser kind of love which in a sense prepares the way for Socrates. The baser love aims at nothing beyond sensual gratification; it finds the means to this in women and young boys, and in the latter case it is to be severely discouraged. The nobler love is directed exclusively towards young men, and its object is a lifelong association productive of such good results as have been described by Phaedrus. In the light of this distinction the attitude of various states and forms of government towards homosexuality is analysed, and the apparent inconsistency of public opinion on the subject at Athens explained. But the importance of the distinction drawn by Pausanias should not blind us to the fact that the nobler sort of love no more precludes sexual relations than the baser, and it is possible to see in Pausanias a clever pleader for homosexual licence, who employs high-sounding but sophistical reasoning to justify the satisfaction of physical desire. His principle that all actions are morally indifferent in themselves, and becomes good or bad only through their circumstances or motives, particularly lays him open to this charge, and is fundamentally opposed to the teaching of Plato.
Annotations may lose their value or be forgotten...

“Some of them [the annotations] are absolutely ridiculous and I can’t believe that I actually wrote this in pen in this book. Some of them are – I have no idea what I’m talking about. Some of them are really interesting, and it’s something I’d forgotten. It just depends on the notes. ... when I did Milton, we were doing the epithets about Satan or something, so I underlined all of them. And when I was going back through it, I’m like ‘what on earth!'”

- A graduate student talks about annotations she made as an undergrad
Reading is a mobile, interactive, immersive individual activity.
Reading is a mobile material interactive interrupted & variable individual activity
...we do not read everything with the same intensity ... a rhythm is established, casual, unconcerned with the integrity of the text; our very avidity for knowledge impels us to skim or to skip certain passages (anticipated as ‘boring’) in order to get more quickly to the warmer parts of the anecdote ...

And yet, it is the very rhythm of what is read and what is not read that creates the pleasure of great narratives: has anyone ever read Proust, Balzac, War and Peace, word for word? (Proust’s good fortune: from one reading to the next, we never skip the same passages.)

from Barthes, The Pleasure of the Text
a look at turning a page:
it doesn’t get much simpler than that...
does it?

(from Marshall and Bly, 2005)
turning a page as a complex of lightweight navigational acts (1)

Constance is reading the first page of a review of Bob Woodward’s *Plan of Attack*. She finishes the first page and reaches for the binding in preparation to turn the page. She can still see what she’s reading…
She turns over the magazine and encounters a full page graphic: Bush with a flyswatter. She glances at the graphic…
turning a page as a complex of lightweight navigational acts (3)

...and reaches to turn to the continuation of the review’s text.
turning a page as a complex of lightweight navigational acts (4)

...lingering over the entertaining caricature before she completes the page turn.
turning a page as a complex of lightweight navigational acts (5)

Constance opens to the next two pages of the review’s textual content giving her a glimpse of what’s to come and quickly informing her that there’s much more to the review.
turning a page as a complex of lightweight navigational acts (6)

She folds the left page (the next page in the article) under...
turning a page as a complex of lightweight navigational acts (7)

and continues to re-orient the magazine to continue reading (she glances at the ad column, as is her habit, to see if there are ads for new fiction; there aren’t)...
turning a page as a complex of lightweight navigational acts (8)

she successfully flips over the magazine so the left page—next to read—is on top…
turning a page as a complex of lightweight navigational acts (9)

and changes the position of her hands so she’s holding the magazine comfortably, ready to continue reading the second page of the review. Readers frequently move their hands to their faces or hair while they read…
Reading is a mobile material interactive interrupted & variable individual activity
Reading
is a
mobile
material
interactive
interrupted & variable
social
Wait!

Reading is *social*.

Are eBooks ready?
“It is also worth noting that solitary reading always was, and still is, inherently social: how we read is ultimately determined by social conventions and community membership”

-David Levy in *Scrolling Forward*
Our old friend: the CSCW matrix

<table>
<thead>
<tr>
<th>when?</th>
<th>where?</th>
</tr>
</thead>
<tbody>
<tr>
<td>same time</td>
<td>same place</td>
</tr>
<tr>
<td></td>
<td>different place</td>
</tr>
<tr>
<td>different time</td>
<td></td>
</tr>
</tbody>
</table>
In the upper left: reading together

<table>
<thead>
<tr>
<th>when?</th>
<th>where?</th>
</tr>
</thead>
<tbody>
<tr>
<td>same time</td>
<td>same place</td>
</tr>
<tr>
<td>different time</td>
<td>different place</td>
</tr>
</tbody>
</table>

When?:
- same time
- different time

Where?:
- same place
- different place

In the upper left, reading together.
reading together
reading together: the importance of shared reference

navigation is vital to getting everyone to the same place

“turn to p.47” is difficult when each class member is using a different edition or reading device

there are many modes of referring within a text (e.g. page numbers, line numbers, by content, by structure...)

...everyone may not be on the same page...
reading together: on-the-spot research enhancing discussion or digression?

“Did they really hang dogs as witches?”

---

More Wonders of the Invisible World, by Robert...

Mr. Dudley Bradstreet, a Justice of Peace in Andover, having granted warrants against and committed thirty or forty to jail, saw cause after which they (by them) were committed. Soon he himself was accused of Witchcraft, and found it his safest course to make his escape.

A Dog being afflicted at Salem-Village, those that had the Spectral sight being sent for, they accused Mr. John Bradstreet (Brother to the Justice) that he afflicted the said Dog, and now rid upon him: He made his escape into Pescataqua-Government, and the Dog was put to death, and was all of the Afflicted that suffered death.

At Andover, the Afflicted complained of a Dog, as afflicting them, and would fall into their Fists at the Dogs looking upon them; the Dog was put to death.

A worthy Gentleman of Boston, being about this time accused by those at Andover, he sent by some particular Friends a Writ to Arrest those Accusers in a Thousand Pound Action for Defamation, with instructions to them, to inform themselves of the certainty of the proof, in doing which their business was perceived, and from thence forward the Accusations at Andover generally ceased.

In October some of these Accusers were sent for to Gloucester, and occasioned four Women to be sent to Prison, but Salem Prison being so full it could receive no more, two were sent to Ipswich Prison. In November they were sent for again by Lieutenant Stephens, who was told that a Sister of his was
sharing reading materials

• a way of reclaiming physicality
• hand-to-hand quality of passing reading materials
• captured the students’ imagination in the Pocket PC investigation
• A reason to question DRM’s role...
Back to the CSCW matrix...

<table>
<thead>
<tr>
<th>when?</th>
<th>where?</th>
</tr>
</thead>
<tbody>
<tr>
<td>same time</td>
<td>same place</td>
</tr>
<tr>
<td>different time</td>
<td>different place</td>
</tr>
</tbody>
</table>

When? and Where? matrix:
- **Same time, same place**
- **Same time, different place**
- **Different time, same place**
- **Different time, different place**

Red oval in the **different time, different place** quadrant.
Sharing the artifacts of reading
Sharing the artifacts of reading:

sharing encountered information
(from Marshall and Bly, 2004)
SYLVIA By Nicole Hollander

**From the Journal of the High-Strung Cat**

A.M. 5:02 - 5:03.
I wait patiently for the sound of food falling into my bowl.

5:04 to 5:05:
Just because it's the weekend she thinks she can sleep in... make the high-strung cat wait for her food.

5:05 to 5:06
Push lip gloss off glass shelf in the bathroom... push bottle of foundation makeup off... break glass shelf.

5:07 to 5:09
Leap onto her bed... rip up duvet... feathers all over. Too subtle. Pull down curtains and rod. She stirs. I run to sit Buddha-like at my bowl.

Who's the sweetest cat... sitting so patiently at her bowl?

LA TIMES 5/20/02
Encountering is an important counterpart to browsing and searching (e.g. Erdelez ’96)
3 questions about sharing encountered information

• how important/ubiquitous is sharing encountered material relative to saving it?

• what’s the function and value of the stuff that gets shared?

• what’s the social role of this kind of sharing?
importance/ubiquity

- **ubiquity**: all of our participants saved and shared encountered information; about 40% of the total examples we collected were about sharing.

- **content importance**: personal saving was more likely to be information-centric than sharing was.

- **personal v. shared**: for some, saving the material for themselves was more compelling; for some, sharing was.
function/value

- sharing for mutual awareness
  - at work, in customer-focused jobs
  - at home, keeping up with friends and family

- sharing to educate or raise consciousness
  - valued by sender, perhaps not by receiver
  - mostly occurred for personal topics/home

- sharing to strengthen social ties
  - “I’m thinking of you.”
  - “We have common concerns.”
  - “We have the same sense of humor.”
sharing to keep in touch or strengthen social ties

P6, a high school student, receives links to online articles from her dad sometimes as often as 2 or 3 times a day. She usually reads the articles on the screen and doesn’t keep them. For example, her dad recently sent her an article from the *NY Times* comparing the war in Afghanistan to Vietnam.

“Sending me the article is like a little note... I don’t know why we starting doing that, but it’s a habit now. And it’s nice to feel like someone’s thinking about you. It’s his way of saying ‘hello’ during the day.”
sharing for mutual awareness

P8 reads the *Wall Street Journal* and the *New York Times* several times a week to see if his company is mentioned. The day before our interview, he heard that the *New York Times* had an article mentioning his company’s chairman. He looked in the paper but didn’t find it. An assistant found the story online and printed it out for everyone in the office.

“Those are two papers our company shows up in a lot, so usually we get an email that says, ‘hey. We’re listed today. Check out the article.’”
P15’s pre-adolescent son had been diagnosed with Asperger’s Syndrome, a form of high-functioning autism. She’d photocopied a “really good” article from *Time* and said “I mailed this to so many people. Because it was very, very good.” She put a note on the copy to her son’s teacher that said “You need to read this.”
social role of sharing: myth busting

- All of our participants shared encountered information
- No real brokers – all of our participants acted as both givers and recipients
- Younger participants are no less likely to share encountered information
- Findings apply to sharing personal (home) and professional (work) information
Sharing encountered information: why is this difficult in the digital realm?
At first blush, it seems to be a solved problem...
but it’s more complicated than that!

Do I have the recipient’s email address at hand?

How about that ad? Will that be included too?

What will it look like?

Will the newspaper keep our names and track our interests?

Will the email look like spam?

Will this seem impersonal?

[Rioux, 2000]
A technological solution for sharing should:

Preserve a sense of layout and article boundaries.

Allow the sender to limit or expand scope or context (compare sending a photo to sending a whole newspaper section).
"My plan is to actually give it [a hardcopy of an article from Nature Online] to him [a project manager] and talk to him about it, rather than just put it in his in-basket because he’d kind of wonder where it came from or why he was getting it. And I’d rather say, ‘hey, I saw this online and it’s pretty interesting. Check it out.’"
Sharing the artifacts of reading:

sharing annotations
“The reason I encouraged such annotations [in FRESS] was that I remembered that when I was in college with Ted [Nelson], I would always grab the dirtiest copy of a book from the library, rather than the cleanest one, because the dirtiest ones had the most marginalia, which I found helpful.”

Andy van Dam
Hypertext ‘87 Keynote
“I have come to view margins as a literary commons with grazing room for everyone – the more, the merrier.”

“Not everyone likes used books. The smears, smudges, underlinings, and ossified toast scintillae left by their previous owners may strike daintier readers as a little icky, like secondhand underwear.”

Anne Fadiman in
Ex Libris: Confessions of a Common Reader
of course, sharing annotations is more complicated than it looks...

see, for example, Shipman et al., ECDL 2003
“publish my annotations now.”

3.1 Computers Becoming a Commodity

As Moore’s law continues to hold, computers available have become fast enough to perform anything that can do. The trick that has been used by computer scientists is to buy expensive, high-performance computers available to the public in five or 10 years now. Computers available to the public are often a research scientist’s. This may have a profound impact on science research is performed. Furthermore, the increased performance makes a qualitative change in interfaces possible. For example, it has now inexpensive palm-size computers, and single-chip power of a 68000 that cost only about 30 cents. Various devices. Another impact of the high performance is becoming more interesting with

Uh-uh.
Sharing annotations (from Marshall and Brush, 2004)

<table>
<thead>
<tr>
<th>Annotation type (on paper)</th>
<th># annotations (on paper)</th>
<th>Corresponding public annotations used in an online discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anchor only (e.g. underline)</td>
<td>1262</td>
<td>42 (3.3%)</td>
</tr>
<tr>
<td>Content-only (e.g. note)</td>
<td>120</td>
<td>23 (19.2%)</td>
</tr>
<tr>
<td>Compound</td>
<td>153</td>
<td>55 (35.9%)</td>
</tr>
<tr>
<td>Total</td>
<td>1535</td>
<td>120 (7.8%)</td>
</tr>
</tbody>
</table>
What about harnessing the wisdom of crowds?
Many ways to implement a 'wisdom of crowds' approach

<table>
<thead>
<tr>
<th>Explicit judgments</th>
<th>Whole text</th>
<th>Within-text</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recommending a book; rating a movie</td>
<td>Quoting a passage</td>
</tr>
<tr>
<td>Implicit actions</td>
<td>Sharing a newspaper article; linking to a web page</td>
<td>Annotating a selection</td>
</tr>
</tbody>
</table>
Reproduced studies of used books, technical articles, and other readings...
Annotations in the aggregate

- consensus is significantly more common than predicted by a strict probabilistic calculation of overlap

- annotators converge on important text that is different than the text that the authors and publisher designate as important
We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,

1) $\max_{w \in C_i} d(x, w) \leq \delta$

2) $\max_{y \in C_i} d_i(x, y) \leq T$

3) $d(x, w) > \delta$ or $d_i(x, w) > T$,
   for all $w \in C_j, j \neq i$. 

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,

1) $\max_{w \in C_i} d(x, w) \leq \delta$

2) $\max_{y \in C_i} d_i(x, y) \leq T$

3) $d(x, w) > \delta$ or $d_i(x, w) > T$,
   for all $w \in C_j, j \neq i$. 

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,

1) $\max_{w \in C_i} d(x, w) \leq \delta$

2) $\max_{y \in C_i} d_i(x, y) \leq T$

3) $d(x, w) > \delta$ or $d_i(x, w) > T$,
   for all $w \in C_j, j \neq i$. 

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,

1) $\max_{w \in C_i} d(x, w) \leq \delta$

2) $\max_{y \in C_i} d_i(x, y) \leq T$

3) $d(x, w) > \delta$ or $d_i(x, w) > T$,
   for all $w \in C_j, j \neq i$. 

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,

1) $\max_{w \in C_i} d(x, w) \leq \delta$

2) $\max_{y \in C_i} d_i(x, y) \leq T$

3) $d(x, w) > \delta$ or $d_i(x, w) > T$,
   for all $w \in C_j, j \neq i$. 

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,

1) $\max_{w \in C_i} d(x, w) \leq \delta$

2) $\max_{y \in C_i} d_i(x, y) \leq T$

3) $d(x, w) > \delta$ or $d_i(x, w) > T$,
   for all $w \in C_j, j \neq i$. 

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,

1) $\max_{w \in C_i} d(x, w) \leq \delta$

2) $\max_{y \in C_i} d_i(x, y) \leq T$

3) $d(x, w) > \delta$ or $d_i(x, w) > T$,
   for all $w \in C_j, j \neq i$. 

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,

1) $\max_{w \in C_i} d(x, w) \leq \delta$

2) $\max_{y \in C_i} d_i(x, y) \leq T$

3) $d(x, w) > \delta$ or $d_i(x, w) > T$,
   for all $w \in C_j, j \neq i$. 

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let $C_i$ be the $i$th cluster, and $w, x,$ and $y$ be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all $x \in C_i$,
III. PROCESSING AND ANALYSIS OF VIDEO

In this section, we give a summary of the processing performed on digital video for the extraction of key information used for the creation of pictorial summarizations. The goals of the processing are to identify the more important segments of video and to assign dominance to each segment, and to segment video into higher and meaningful units beyond the shots, such as scenes or events, so that video poster can be constructed for each individual unit.

A. Segmentation of Video into Shots

To analyze the content of a given video sequence, we need to first segment the sequence into individual shots. Such analysis is computationally expensive when full resolution video frames are used because of the huge data size. In addition, most video sequences are captured and stored in various standard compression formats. For compressed video in motion JPEG or MPEG-1 standard compression formats, we use the algorithms in [16] based on dc image extracted directly from compressed video to segment a video into shots. The detected shots form the basis of further analysis steps described in the following sections.

B. Visual and Temporal Classification of Shots

A first step toward video processing is the semantic labeling of video shots. We want to associate with each shot a label Lm, which provides description of the content of the shot. For example, a shot of a news anchor person could be labeled as “news anchor,” “news room,” or “man behind table.” Such semantic descriptions, however, is difficult to derive at by automatic means. Instead, we make use of a direct consequence of the scene presentation of film—that parallel or simultaneous events are depicted one after another, with substantial repetitions of shots with similar contents.

To capture such repetitions, clustering of video shots can be performed.

We have proposed in [15] a clustering process called the time-constrained clustering which takes into account both visual characteristics and temporal locality of shots. Let \( C_0 \) be the ith cluster, and \( m \), \( n \), and \( y \) be elements in a cluster, each of which is a video shot, we impose the clustering criteria as follows. For all \( x \in C_i \):

1. \( \max d(x, y) \leq t_y \)
2. \( \max d(x, y) \leq T \)
3. \( d(x, y) > \theta \) or \( d(x, z) > T \) for all \( z \in C_j, j \neq i \).

Here \( d(\cdot, \cdot) \) is the visual dissimilarity between two shots. \( d_r(\cdot, \cdot) \) is the temporal distance between two shots, \( d_r \) is the maximum visual separation between two shots, \( T \) is the maximum temporal distance in number frames, between the end of the earlier to the start of the later shot. This formulation reflects the scenario where two shots may share similar visual characteristics in the content, but are located far apart in time, which may potentially represent different events or occur at different scenes. The criteria imply that any other shot outside of the cluster must have a distance greater than \( \theta \) relative to any shots in the cluster; in addition, each cluster, no two shots can be separated by more than \( T \) frames apart. By combining \( d(\cdot, \cdot) \) and \( d_r(\cdot, \cdot) \) into a new distance, and using the hierarchical clustering method based on complete-link, we can generate clusters that satisfy the condition listed above [15]. In our clustering processing, matching between two video shots also takes into account the temporal variations as proposed in [17].

Given video sequence, suppose that the time-constrained clustering process partitions the sequence of shots into \( N \) clusters \( C_1, C_2, \ldots, C_N \). We then assign a label \( j \) to a shot \( m \), \( n \in C_j \). In this case, the label \( L_m \) is a scalar quantity. In general, it can be a multidimensional vector. Each dimension will then correspond to a different feature. Other features may include shot duration, color distribution of the shot, dominant motion characteristics, dominant texture patterns, spatial moments, etc. In the following sections, we shall describe how the aggregate statistics of label sequence \( L_n \) permit segmentation of a video into distinct story units and characterization of different types of meaningful segments in a sequence.

C. Finding Story Units in Video

A shot is the fundamental unit of a video production, but a shot by itself does not have much meaning in a story. Instead, a basic unit of the story can be a scene, which is comprised of a number of fragmented shots unified by location or dramatic incident. In a given scene, we often find multiple paths coexist, and the shots of these paths are juxtaposed and linked together with multiple shots of the same party. Shots from different scenes are not juxtaposed, except at the transition from one scene to the next. Because of the intense interactions between shots in a scene, the use of a label sequence can be used to segment a video into large logical units, called the story units, each of which closely approximates a scene.

Let \( (A, b) = \text{max}(a, c, d) \) represent the shot index of the last occurrence of label \( L \) starting at shot \( a \). The algorithm to detect a story unit proceeds as follows:

[Algorithm Detect Story Unit]

1. Set \( l = m \)
2. Set \( c = \text{last}(L_m, m) \)
3. While \( l \leq c \)
   1. If \( (A, b) = \text{max}(a, c, d) \) or \( (A, b) = \text{last}(L_m, m) \)
      1. \( l = c + 1 \)
   2. Shots \( a, a_1, a_2, \ldots, a_n \) constitute a story unit.

The algorithm examines the sequence of labels and identifies the subsequences of labels that are of minimal length and which contains all the identical (or recurring) labels. For a video sequence with \( n \) shots with the following labels \( A, B, C, D, F \), the first story unit consists of shots \#1-\#4. Each label in a video sequence is associated with a first and last shot. The first story unit that has to contain the first shot; in addition, it has to contain the last shot which has the same label as the first shot, and all intermediate shots. Such a method of
finally...

what does this mean for eBooks?

are we getting the social part right?
collaboration and reading technologies

What of displays – are we thinking enough about “looking on” or shared focus?

How do social expectations interact with restrictions introduced by Digital Rights Management?

Which collaboration architectures will work for people using the same collections (e.g. annotation, reading room, bookmark servers)?

Are there new modes of collaboration enabled by digital devices?
My collaborators

XLibris studies: Morgan Price, Bill Schilit, and Gene Golovchinsky at FXPAL

MS Reader study: Chris Ruotolo, University of Virginia

Clipping and navigation studies: Sara Bly

Personal/Collaborative annotation study: AJ Brush, University of Washington

New York Times News Reader study: Rob Hunt, Kathy Mahoney
Contact info (if you’re so inclined):

http://www.csdl.tamu.edu/~marshall
http://research.microsoft.com/~cathymar
    cathymar@microsoft.com