Virtual Paper

Andrew Birrell & Paul McJones

Goal

- Making on-line reading more attractive than paper

On-line Reading Offers ...

Potentially:

- speed

- cheap color and pictures

- search, browsing, retrieval

- universal availability

Often:

- slowness

- illegibility

- awkward user interface

- restricted scope

- poor ergonomics

# The Competition: Books and Hardcopy

- Single resolution

- Legible

- Easy user interface

- Fast and convenient access

- Mostly treated as read-only

Our Approach

- *Performance:* similar to a book (multiple pages per second)

- *Legibility:* use gray-scale (anti-aliasing & sub-pixel positioning)

- *User Interface:* uncluttered; single keystroke shortcuts

- *Generality:* reduce all content to raster images

- *Power:* full-text indexing; inter- & intra- document links

Organization

PostScript

Scanned
Images

Image
Files

Lectern
Document

Repository

Index
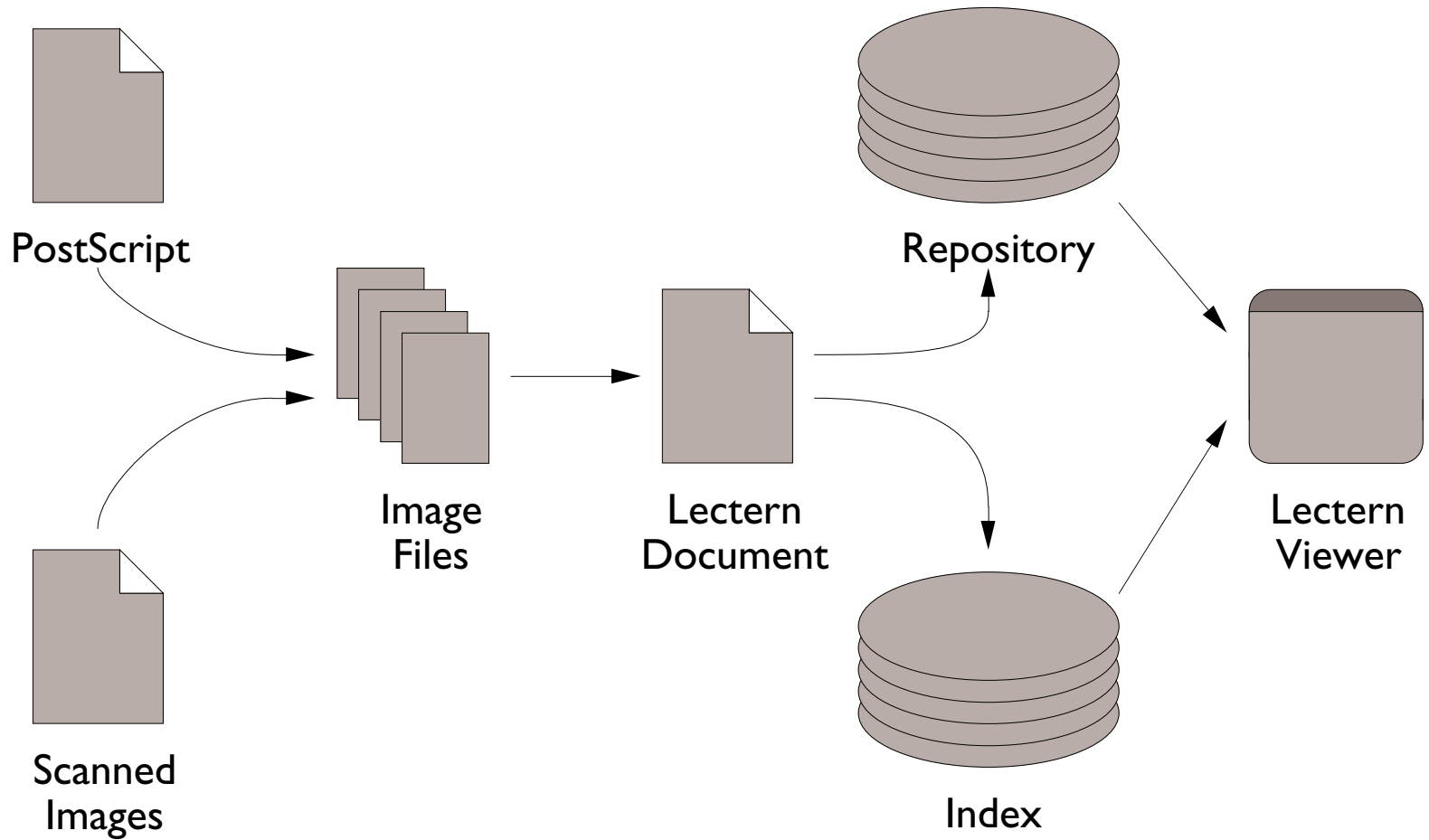
Lectern
Viewer

Technical Details

- Lectern document = scaled images + text + rectangles

- Anti-aliasing and color reduction:
  - — B&W becomes 2-bit gray
  - — Color becomes 6-bit color (diffusion dither)

- Compression to reduce file size and bandwidth:
  - — 100 DPI 2-bit grayscale is typically 30 KBytes
  - — 4 pages/sec = 1 MBit/sec
  - — Total size = 100 KBytes/page

What We Have

Good:

- speed (2 pages/sec at 20 MIPS)

- color and pictures

- good search and retrieval

- comfortable user interface

- good legibility

- can handle any document

Marginal:

- build time (9 secs/page)

- size & bandwidth

- limited OCR accuracy

- physical ergonomics

Some of the Alternatives

- Adobe Acrobat™

- Common Ground™

- HTML

- xdvi

- Books and Hardcopy

SRC

Potential Applications

- Publishing books, papers and journals on the net

- CD-ROM publishing of books, papers and journals

- Replacing microfilm (libraries, newspaper morgues)

- Replacing hardcopy collections (reserves, delicate books)

- Paperless (or less-paper) office

SRC

Possibilities for Further Work

- Page-at-a-time access over Internet via HTTP

- Make our links inter-operate with HTML

- Provide annotations

- More powerful links

- Different document types: text, PostScript, ...

digital

SRC

Conclusions (1) — Imaging

- Can read, decompress, and render in 2-3 instructions/pixel

- Gray-scale and sub-pixel positioning help legibility a lot

- ... but hand-tuned bitmap fonts can be better if available

- Monitor resolution is still marginal for continuous reading

SRC

Conclusions (2) — Online Reading

- Online reading can really work

- There's a sharp threshold for performance acceptability

- A smooth, fast, user interface is critical

- OCR is good enough for searching and indexing

- Full-text indexing is a pretty good search & retrieval tool

digital

SRC