The Lifetime Reader

George Nagy
Rensselaer Polytechnic Institute

This is an exposition of the hardware and software hurdles that must be overcome to deploy a wearable sensor that collects all the text that it sees or hears. The text is uploaded to a standard private platform (a smartphone or a laptop) or to a cloud service for character and speech recognition followed by indexing. Selected segments are retrieved on demand. Compared to queries on the World Wide Web, searching the prospective microcosm of information is analogous to finding one's way in one's backyard instead of in a national forest.

"I cannot remember the books I've read any more than the meals I have eaten; even so, they have made me," said Ralph Waldo Emerson**.** Surrounded as we always are by natural and computer-mediated visual and auditory stimuli, much of our information diet is still based on words. In addition to newspapers, magazines, books, and pamphlets, we browse and listen to smartphones, tablets, laptops, and television. When we drive or walk, we cannot avoid posters and sign. Far more text passes before our eyes and through our ears than we can remember or even assimilate. Can we preserve it all for recall at will?

The Lifetime Reader is now where the internet and drones were in 1970, the web and the smartphone in the early nineties, and self-driving cars about ten years ago. Must we be taken by surprise, again, when the Reader is deployed?

In a 1945 issue of *The Atlantic*, Vannevar Bush anticipated the wearable camera and the storage and retrieval of personal libraries. Bush's *memex* (memory extender) was an optical device. Inspired by the confluence of camera-based optical character recognition, document image analysis, speech recognition, wearable electronics, information retrieval, web science, and cognitive computing, contemporary researchers revived the notion of keeping track of what we read or see.[1–3]

The Lifetime Reader has read everything that we have, as well as most of what we barely glanced at (see Figure 1). Unlike us, it can recall moat anything at the right prompt.

The Reader will keep us from being inadvertent liars and from misquoting Shakespeare, Shannon, and our spouses. It could save us from a frustrating search on-line for that play where the old curmudgeon insists on noodles without salt. It will quickly recap the article we browsed in the dentist's waiting room that explained how bitcoin is mined and pointed out that no nation accepts bitcoin for tax payments.

We leaf through many pages but read only a few stories in thousands of issues of the *New York Times* and of our local paper, and remember even less. We revisit conference papers and articles when we forget crucial details or, surely seldom, their major point.

Some of us skim advertising pamphlets, colorful political flyers, and notices from our credit card issuers. Occasionally, we regret that we had thrown them away. We may want to dig out old email or

Published by the IEEE Computer Society
1536-1268/18/$33.00 ©2018 IEEE

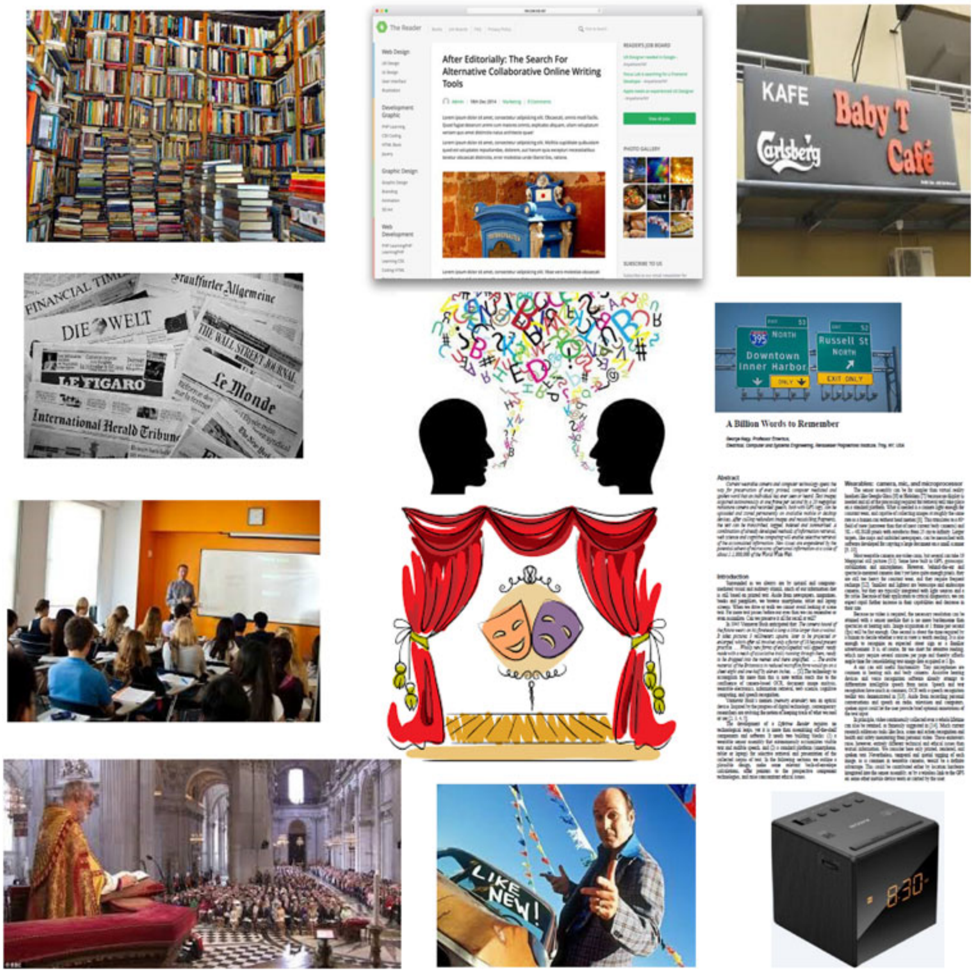

Figure 1. Some examples of what we have seen or heard: books, signs, lectures, sermons, sales pitches. . .

the instructions for relighting the furnace pilot and programming the thermostat. Now where did I see that advertisement for universal calibration charts for wide-angle cameras?

A personal Wikipedia of the papers we have read, the slides that we have seen, and the speakers we have heard over the past decades will let us pose as authorities on whatever turned us on enough to study it. Others may want to retrieve elusive blogs, newsfeeds, webinars, wikis, podcasts, genealogical resources, poems that they had enjoyed, letters they wish they had kept, and PDF or DOC files that they never transferred from an earlier laptop.

The smaller the collection, the more relevant are the associations and the easier it is to find a specific item. In fact, purposive forgetting of unimportant material has been proposed in the context of lifelogging.[4] A small, personal collection of material that we have been exposed to will not only speed up the search for some items but also let us find verbal memories that lack any other digital incarnation.

The Reader requires no insurmountable technological leaps, yet it is much more than an assembly of off-the-shelf components and software. Figure 2 shows its two communicating modules: (1) a wearable sensor assembly and (2) text-processing software on a standard platform (smartphone, tablet, or laptop). The sensor module autonomously accumulates images of visible text and records audible speech. The text processor filters, analyzes, and transcribes the collected text-image and speech-audio data, and retrieves selected segments on demand. Below, I outline design alternatives, present some relevant back-of-envelope calculations, offer pointers to the component technologies, and raise concomitant ethical issues.

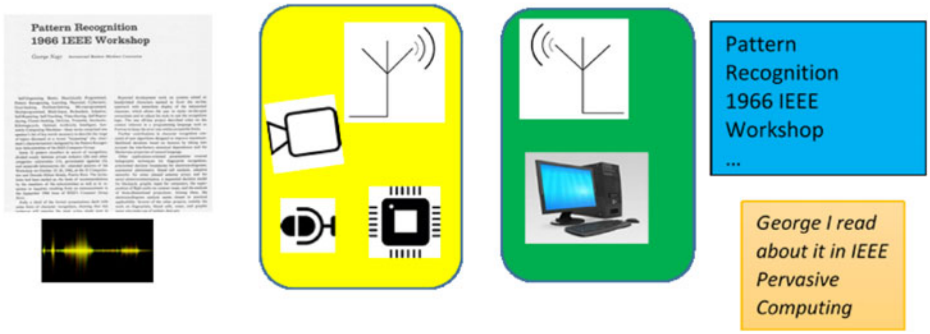

Figure 2. The wearable sensor module collects text images and audio. It transmits them intermittently to the text-processor module, which transcribes and encodes the text data for selective display on demand.

How is the Lifetime Reader situated among other machine reading paradigms? The oldest and best established technology is the transcription of *scanned* printed material that culminated with the fifteen million digitized Google Books. *Camera-based OCR* has two flavors: transcribing selective photos of the text of interest, such as posters, menus, historical tablets, or explanatory labels in museums, and continuous video text transcription (e.g., highway signs on a trip). *Scene text* reading requires the segmentation of legible text from background: buildings, artifacts in shop windows, and clutter on a desk.

Lifeloggers are usually more interested in nontext images that can be used to reconstruct both unique and repetitive visual contexts of the device wearer's experience,[5] and even the faces of people they have met.[6] Wearable cameras may be augmented by audio, GPS, accelerometers, light meters, and distance finders. Ayumu is a lifelogging approach to reading.[3] Unlike the Lifetime Reader, it proposes to capture only text purposively read at arm's length; therefore, it has more modest camera and image-processing requirements. To extend battery life, it is equipped with additional low-energy sensors that trigger the camera only when the user is actually reading. Google Cloud Video services strive to unite all of these threads. Photographs and video submitted to the service are labeled, geo-located, culled for text, and analyzed for offensive material and for related web content (https://cloud.google.com/vision/). Mainly text images can be translated, parsed, and searched. The modules for object and text classification make use of recent advances in deep learning.

WEARABLE TEXT ACQUISITION

Social prejudices against earbuds and hearing aids are disappearing as those against spectacles ("four eyes") did long ago. However, people may still feel uncomfortable with a camera and a mic pointing at them. Even if progress in miniaturization renders them almost invisible and everybody understands that only text images are retained, in some situations it may be desirable to remove the camera. Such interruptions in recording will lose only fragments of conversation. There are, of course, many venues where camera use is already barred. Brian Gaff raises unresolved legal issues of data ownership, privacy, and intellectual property with wearable cameras.[7]

In principle, even video can be continuously collected and retained over a whole lifetime, as famously suggested by Gordon Bell. Leapfrogging science fiction like *The Entire Life of You* episode in *Black Mirror*, lifelogging is already a burgeoning field of research.[8] Computer vision addresses tasks like face, scene, and action recognition and health and safety monitoring from a personal video. These endeavors raise, however, entirely different technical and ethical issues than textual information. Nevertheless, temporal and spatial tagging, as is common in wearable cameras, facilitates text retrieval too. Tagging could be accomplished either by location hardware integrated into the sensor assembly, or by a wireless link to the GPS and cellular network access on some other mobile device worn or carried by the user.

The sensor assembly can be far simpler than virtual reality headsets like Google Glass or HoloLens because no display is needed. Although some proposals for logging reading habits are based on eye-tracking,[9] the actual text cannot be recovered from eye movement. Reviving an idea of artificial intelligence pioneer Herbert Simon, George Buscher *et al*. focus on attention to determine *those pieces* *of information that we think are most relevant, interesting, or useful to us in our current situation.*[9] Simpler than eye tracking, electrooculography has already been incorporated into spectacles and used to measure reading speed by detecting saccades. This measure of the potential difference between the cornea and the retina may prove a useful addition to the Lifetime Reader for tagging attentively read text. But importance can be at least approximated also by how long a given text image is retained in view.

CAMERA AND MICROPHONE

The camera must be light and unobtrusive enough for constant wear. The human body offers relatively little real estate for locating a camera with a field of view similar to that of the eyes. Unless people are willing to tolerate almond-sized devices affixed to their forehead or nose, the least burdensome solution is modified spectacles like the J!NS MEME prototype[10] and KJR DVR290 in Figure 3. In time, the hardware could be miniaturized enough to be embedded in a neutral contact lens or in a fake freckle.

The camera collects images of the text at roughly the same level of detail as a human.[11] A 30-cm wide newspaper subtends 35 degrees at the normal reading distance of ~50 cm. Therefore, a camera with a 60° field of view (narrower than that of most body cameras), with autofocus from 25 cm to infinity, can capture whatever we are able to read without undue effort. Larger targets, like maps and posters that we read with head motion, can be mosaicked with software developed for copying or digitizing a large document on a small scanner. Similar techniques are used for stitching together panoramic photographs. Material destined for optical character recognition is typically scanned bilevel at ~12 lines per mm (300 dots per inch), so 20 Megapixels should suffice. New smartphones have up to (and often under-utilized) 36 Megapixel sensor chips. With fast lenses, automatic sensitivity setting, and optical stabilization, they can capture blur-free legible page images in a wide range of light.

Most wearable cameras are video cams, but several can also take 16 Megapixel still pictures (http://www.toptenreviews.com/electronics/photo-video/best-wearable-cameras/). Some have built-in satellite or cell phone positioning, gyroscopic stabilization, and microphones. However, behind-the-ear and spectacle-mounted cameras do not yet have quite enough pixels, they are still too heavy for constant wear, and they require frequent recharge. Smallest and lightest are borescope and endoscope cameras, but they are typically integrated with light sources that need a fat cable. A microcamera weighing less than 10 g (about the weight of an AAA battery), built around a cellphone sensor chip for observing mouse brain activity, was recently commercialized by Stanford University researchers.

We estimate, based on normal reading speed, that it takes less than one second to decide whether some text in view is worth reading or to recognize an expected street sign or a familiar advertisement. Therefore, image acquisition at 1 frame per second is fast enough. Attentive reading requires up to several minutes per page, which leaves ample time for filtering and compressing recent text images.

All continuous wearable video recording requires sufficient battery life for recharge only at night. None of the wearable cameras is designed specifically for reading. Microsoft's SenseCam, with which lifeloggers have achieved impressive organization of iconic memory, records wide-angle images at VGA resolution (https://www.microsoft.com/en-us/research/project/sensecam/#). SenseCam

Figure 3. Left: KJB DVR290, 1280 × 720 pixels, 312 g. Right: J!NS MEME prototype, smart glasses.

advertises 24 h of battery life at two 5 MP frames per *minute,* while Narrative Clip claims two days at the same frame rate. The Weldex spy camera (still too large at 25 mm × 25 mm × 16 mm) consumes 0.72 W, so a 17 g lithium thionyl chloride battery that can store 9 Wh would last almost a whole day. The DVR-290 (see Figure 3), concealed in a spectacle frame, claims 15 h of a 30 fps low-resolution video. The most efficient combination of high-resolution consumer camera and (sizeable) battery yields only 2 or 3 h of operation (but much of the energy goes into the viewfinder). These factoids suggest that without a breakthrough in battery technology, we will have to wait about two years to enable Moore's law to satisfy the desired 16 h of capture of 20 MP images at 1 fps with a 10-g camera.

Tiny microphones are common in hearing aids and body cameras. Assistive hearing devices and voice recognition software can already differentiate intelligible speech from noise. Speech and character recognition use similar algorithms; character recognition with a speech toolkit was demonstrated by Patrick Schone *et al.* Aside from recording personal conversations and speech on radio, television, and computers, audio input gives the user the option of providing brief oral annotations of the text-image input. The speech audio, which preserves prosody, inflection, and accent, can also be retained without excessive increase in storage requirements. The transcript, geolocation, and temporal tagging will still be used for retrieval.

The microprocessor integrated with the sensors detects text in the field of view and audible speech. It compresses and temporarily stores both. Fast image and video compression algorithms are available, but speed is not essential. Compression need not be done in real time because recently acquired images can be filtered and compressed when there is no text in view. The camera can be kept in a low-resolution surveillance mode except when it sees readable text.

STORAGE REQUIREMENTS

How much wearable storage is necessary? On clean text images, a compression ratio of 40:1 is readily achievable with symbol matching. Even keen readers will have text in view during at most half their waking hours and will often dwell on the same text for many seconds.[3] Therefore, an average compression ratio of 100:1 seems conservative. That yields only 17 GB of text-image data per day (see sidebar). The speech audio amounts to much less.

The processor module or cloud service stores encoded text (*image text*) instead of text images. GZIP, Lempel-Ziv, or other dictionary-based methods yield a fivefold compression on normal prose. Current estimates of the entropy of English text are about 1.6 bits per character. As shown in the sidebar, reading or listening at 300 words per minute for eight hours a day would accumulate only 300 KB per day after compression, in contrast to 17 GB for compressed images! However, leafing through dense pages at one page per second, as opposed to reading everything, will raise the 300 word/minute rate by a factor of 240. That is still only ∼0.5% of the image volume.

Storage calculations

Text-image: 4K × 5K pixels × 3 Bytes/RBG pixel × 1 frame/s × 3600 s/h × 8 h/100× compression
= **17 GB/day**

Audio: 4 KB/s × 3600 s/h × 8 h
= **115 MB/day**

Image text: 2 Bytes/char × 5 chars/ word × 300 words/min × 60 min/h × 8 h /5×
= 300 KB/day → 300 KB × 365 days × 100 years
= **10 GB/lifetime**
Audio text is the same:
10 GB / lifetime (but assume that one is either reading *or* listening)

Audio text rates range from 300 B/s for a vocoder to 1.4 MB/s for stereo CD audio books. An estimate of 4 KB/s for readily understood text gives 115 MB/day. Reduced to alphanumeric text via speech recognition, it is again only 300 KB per day.

Text exposure and purposive reading vary according to age, education, employment, and perhaps even gender. This complicates the collection of statistically representative data for experimentation. Considerable data are available from recent competitions on robust camera- and smartphone-based transcription and from the benchmark datasets of the International Association for Pattern Recognition Technical Committee on Reading.

TEXT TRANSCRIPTION AND ENCODING

The software on the mobile or desktop host computer (or, at the user's choice, in a cloud) culls unreadable and repetitive images that were not filtered out by the camera computer, mosaics some frames, performs layout analysis to determine reading order, and then recognizes (OCR) and indexes the text for eventual retrieval. Verbal audio is processed analogously. The only outputs of the host computer are a display for rendering metadata and minimally formatted text (see Figure 4) and an audio channel for text-to-speech (for example, to listen to passages from a long-ago-read book while driving or exercising). Already available language translation and privacy/security (encryption) features can be added at small cost. In contrast, most current research on camera-based OCR addresses real-time output, as required, for example, for translating posted signs.[12]

Camera-based character recognition dates back to at least 1960, as illustrated in Figure 5, which shows the letter "C" being captured for a perceptron by a camera with a 20 × 20 array of photocells in its

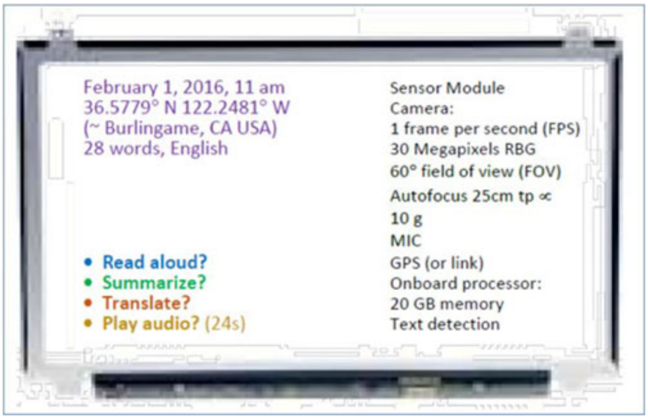

Figure 4. Mock-up of text retrieval. On the left of the screen are capture time and location, and colored control buttons for vocalizing, summarizing, or translating the retrieved text, or playing the original audio. On the right is a barely formatted display of the transcription of the text in the image of a conference slide.

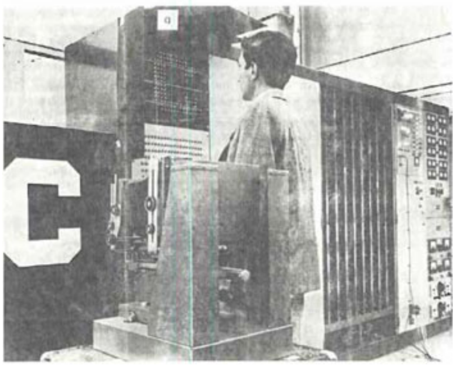

Figure 5. Character recognition in 1960 with a 20 × 20 pixel camera.

image plane. Often-noted differences between scanned and camera captured text are the possibility of severe geometric—affine and perspective—distortion, and contrast variations due to uncontrolled illumination. Although dozens of contrast enhancement and skew detection/removal methods are available, the extent of distortion in camera captured text, especially scene text, requires affine or perspective invariant methods similar to those used in computer vision.[13] Fortunately, most purposively read text will be subject only to modest distortion because we prefer to read in good light and to keep what we read (hardcopy or display) horizontal and perpendicular to our (and the camera's) line of sight.

Because of the variety and unpredictability of the Reader's input, such as single and multicolumn text, bureaucratic forms, tables, comic books, email, posts on social networks, blogs with advertising pop-ups, mathematical formulas, and scene text, layout analysis will be more demanding than is required for the relatively uniform input streams of commercial and historical document digitization.

Algorithms for categorizing the *genre* of the target text (e.g., book, billboard, license plate) help both layout analysis and character recognition. Multiple images of text, captured during attentive reading, can be combined to improve character recognition. Consistency of individual reading, browsing, and travel will avoid frequent resort to script and language recognition. An alternative approach for archival text matches fragments of documents at the image level for retrieval of the entire searchable document from a digital library.

Representative examples of current work include scene-text localization and scene text analysis.[14] The established forum for new results is the biennial international workshop on camera-based document analysis and recognition. Although autonomous, ubiquitous and uninterrupted (i.e., pervasive) text acquisition and transcription does present some new problems, none seems insurmountable.

RETRIEVAL

Browser technologies for retrieval of nonannotated are now well beyond simple keyword search. The puzzle is retrieving vaguely or inaccurately remembered material that one may have browsed in the distant past. The *Remembrance Agent*[15] anticipates some of the prospective solutions, and lifelogging researchers are developing methods for organizing vast, heterogeneous of automatically recorded information, including text.

Unlike the web, a personal collection never has to be re-indexed, because one cannot *unread* something. Initially, there will not be any *PageRank*, but cross-linkages can be automatically constructed using measures of similarity and temporal and spatial proximity. The system will gradually develop and make use of the profile of its single user, i.e., a *personal information model*.[16,17] Ontological tools developed for the semantic web will also play a role in personal collections.

Another set of query tools is available from the library side. These now incorporate all the tools of information retrieval like inverted indices, vector-space models, text tagging, fuzzy clustering, latent semantic indexing, relevance feedback, perfect hashing, signature files, graph algorithms, and pattern matching on the compressed text.

Three factors facilitate retrieval from a personal collection. The first advantage over web search is that there are less data to be searched. Even if one started in a grade school and lived to be a hundred, and we include the text only seen or heard but not read, it is not even one-millionth of the estimated size of World Wide Web. The second advantage is that the list of the top-ranking items displayed in response to a query will already seem familiar, so we can parse it quickly to find the page, passage, or phrase that we sought. (This is why some of us hang on to obsolete but well-thumbed textbooks.) Finally, we will not be bothered by character and speech recognition errors because we are all used to fractured and misspelled prose and because we are unlikely to disseminate *verbatim* what we retrieve.

CONCLUSION

The technology required for the lifetime reader would seem simpler than what has already been demonstrated for self-driving cars, autonomous drones, and smartphones. However, cameras are still too large, format analysis is not robust enough, retrieval from an autonomously collected personal collection has not been demonstrated, and several new ethical issues need deliberate consideration. The major hardware challenge is clearly the development of a 10 g camera module capable of taking 60 000 20 Megapixel images between battery recharges. Extracting enough power from the environment or body heat/motion does not look promising. However, the 20 MP specification could be relaxed with a nonuniform image-sensor array and the average frame rate could be reduced by low-resolution detection of the text.

Software has a more incremental effect. Progress on each item listed in the sidebars will improve the utility of the Lifetime Reader, but current solutions capable of producing only a partial and noisy record may satisfy some early adopters. While some items could be (and, perhaps, will be) the subject of a full-blown paper, here I can only mention them in a sidebar as research opportunities.

The real question is whether occasional retrieval of a dimly remembered read or heard nugget is worth the burden of another wearable. Even if so, the first imperfect prototypes will be worn only intermittently, mostly in an information-rich environment. Early adopters will wear their Lifetime Reader when they currently take their laptops and camera phones: to class, conference presentation, training session, committee meeting, and perhaps even the library. Will it change reading habits as view-on-demand changed TV habits?

A mere forty years ago, folks would leave their home with only their wallet, pocket diary, keys, and watch. Who would have guessed that soon most people would not dream of going out without their music, videos, telephone, email, arcade games, and instant access to credit? That the most popular form of photography would be selfies? That people would count how many steps and breaths they take every day? That telegraphy would have a miraculous revival in the form of texting? That politicians will tweet rather than orate? That keyboards would be shrunk to palm size, and car clocks would be reprogrammed, instead of reset by turning a knob? Perhaps the Lifetime Reader is ready to take its rightful place in this amazing progression.

Potential Research Directions

Image Acquisition

- Text detection in spatial context: at home, at work, in local venues, in transit, abroad.
- Mosaicking required by head and body motion while reading.
- Lazy compression of text images.
- Optional hands-free (via mic) annotation.
- Optional visible (gestural) annotation, e.g., by tracing a phrase on a printed page or computer screen with a designated finger.

Text-Image Analysis

- Perspective-invariant recognition instead of skew removal.
- Reading-order without gaze tracking.
- Duplicate detection from consecutive frames and after (possibly lengthy) interruptions.
- Retention policy for undecipherable and unindexable fragments of text, and for near-duplicates.
- Adaptation to predictable reading material like our daily newspaper, favored magazines, the remaining volumes of the Jack Aubrey series, *Computer*, Python v2.7.6 documentation . . .

Information Retrieval

- Retrieval strategies that mesh with our own mental recall, e.g., “I read it in high school,” or “around Thanksgiving at my sister’s place.”
- Creating and exploiting an evolutionary personal information model.
- Personalization: scripts and languages—reading speed—reading postures—computer display settings—work, leisure, shopping, and napping habits.
- Selective, topic-, time-, or location-specific summarization.
- Mathematical formula and specialized notation (chess, bridge) retrieval.
- Logging queries, responses, and user reactions for improving the system even as one’s own memory fades.

Ethical and Legal Issues

- Security and privacy: what do these mean over a lifetime?
- Copyright: should we have permanent access to anything that we have read?
- Text piracy: would the likely lack of source metadata encourage it?
- What is the legal difference between deliberately acquired information, as with a smartphone or camera, and autonomously acquired information?
- What responsibility does delayed discovery of a crime entail (for instance, reading an airplane seat neighbor’s laptop screen that one glanced at two years ago)? Can the recorded data be subpoenaed?
- What are the social and marketing implications of lifetime text logging?

ACKNOWLEDGMENTS

The author is grateful for the perspicacious suggestions of the referees and for the thoughtful guidance of the EIC.

REFERENCES

1. H. Fujisawa, H. Sako, Y. Okada, and S.-W. Lee, “Information capturing camera and developmental issues,” in *Proc. Int. Conf. Doc. Anal. Recog.*, 1999, pp. 205–208.
2. T. Kimura, R. Huang, S. Uchida, M. Iwamura, S. Omachi, and K. Kise, “The reading-life log—Technologies to recognize texts that we read,” in *Proc. Int. Conf. Doc. Anal. Recog.*, 2013, pp. 91–95.
3. B. Stoddard, K. O’Hanlon, B. Lin, A. Machanvajjhala, and L. P. Cox, “Ayumu: Efficient lifelogging with focused tasks, MobCase,” in *Proc. 8th EAI Int. Conf. Mobile Comput.*, Cambridge, U.K., 2016, pp. 127–137.
4. C. Jilek, H. Maus, S. Schwarz, and A. Dengel, “Managed forgetting, data condensation & preservation in application,” in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.: Adjunct*, Heidelberg, Germany, Sep. 2016, pp. 1046–1053.
5. B. P. Clarkson, “Life patterns: Structure from wearable sensors,” PhD dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 2002.
6. M. Iwamura, K. Kunze, Y. Kato, Y. Utsumi, and K. Kise, “Haven’t we met before? A realistic memory assistance system to remind you of the person in front of you,” in *Proc. Augmented Hum. Int. Conf.*, 2014, Paper 32.
7. B. M. Gaff, “Legal issues with wearable technology,” *IEEE Comput.*, vol. 17, no. 40, pp. 10–12, Sep. 2015.
8. A. R. Doherty *et al.*, “Passively recognising human activities through lifelogging,” *Comput. Hum. Behav.*, vol. 27, no. 5, pp. 1948–1958, Sep. 2011.

9. G. Buscher, A. Dengel, R. Biedert, and L. van Elst, “Attentive documents: Eye tracking as implicit feedback for information retrieval and beyond,” *ACM Trans. Interact. Intell. Syst.*, vol. 1, no. 2, Jan. 2012, Art. no. 9.
10. K. Kunze *et al.*, “Quantifying reading habits – Counting how many words you read,” in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Osaka, Japan, Sep. 2015, pp. 87–96.
11. A. Betancourt, P. Morerio, C. S. Regazzoni, and M. Rauterberg, “The evolution of first person vision methods: A survey,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 5, pp. 744–760, May 2015.
12. T. Kobayashi, M. Iwamura, T. Matsuda, and K. Kise, “An anytime algorithm for camera-based character recognition,” in *Proc. Int. Conf. Doc. Anal. Recog.*, 2013, pp. 1140–1144.
13. Y. Takezawa, M. Hagesawa, and S. Tabbone, “Camera-captured document image perspective distortion correction using vanishing point detection based on radon transform,” in *Proc. 23rd Int. Conf. Pattern Recog.*, Cancun, Mexico, 2016, pp. 3968–3974.
14. Q. Ye and D. Doermann, “Text detection and recognition in imagery: A survey,” *IEEE Trans. Pattern Recog. Anal.*, vol. 37, no. 7, pp. 1480–1500, Jul. 2015.
15. B. J. Rhodes and T. Starner, “A continuously running automated information retrieval system,” in *Proc. 1st Int. Conf. Pract. Appl. Intell. Agents Multi Agent Technol.*, 1996, pp. 487–495.
16. H. Fujisawa, A. Hatakeyama, and J. Higashino, “A personal universal filing system based on the concept-relation model,” in *Proc. 1st Int. Conf. Expert Database Syst.*, Charleston, SC, USA, 1986, pp. 31–44.
17. C. Jilek, H. Maus, S. Schwarz, and A. Dengel, “Diary generation from personal information models to support contextual remembering and reminiscence,” in *Proc. IEEE Int. Conf. Multimedia Expo Workshops*, Torino, Italy, Jul. 2015, pp. 1–6.

ABOUT THE AUTHOR

George Nagy is a Professor Emeritus with Rensselaer Polytechnic Institute. His current research interests include document image processing, optical character recognition, and information extraction from web tables and books. He received the Ph.D. degree in electrical engineering in 1962 from Cornell University’s Cognitive Systems Research Program on neural networks. He is a Life Fellow of IEEE. Contact him at nagy@ecse.rpi.edu.