Have A Say Over What You See: Evaluating Interactive Compression Techniques

Simon Tucker Steve Whittaker
University of Sheffield
United Kingdom
{s.tucker,s.whittaker}@sheffield.ac.uk

ABSTRACT
We all encounter many documents on a daily basis that we do not have time to process in their entirety. Nevertheless, we lack good tools to rapidly skim and identify key information from within such documents. This paper develops and evaluates Interactive Compression (IC) techniques that allow users to dynamically configure the amount of information they view in a document, e.g. by automatically removing unimportant information from view (Excision) or by making important information more salient (Highlighting). We explore IC techniques in the context of meeting transcripts that are typically unstructured - making it difficult to isolate relevant regions and extract key information. We demonstrate the superiority of IC compared with an unmodified text control. In contrast to traditional summaries, our results show extensive use of interactive, as opposed to fixed compression level, summarization. They also show the value of word- as opposed to utterance-based compression. There are also trade-offs between different IC designs. Excision allows users to scan documents faster than Highlighting but at the expense of overlooking relevant sections of the document.

Author Keywords
Interactive compression, scanning, reading, summary, gist, meetings

ACM Classification Keywords
H5.m. Information interfaces and presentation: Miscellaneous

INTRODUCTION
In our daily lives, we are exposed to increasingly large numbers of documents that we do not have time to process in their entirety. This increase is compounded by improvements in tools for information access. Tools for information foraging [7, 23], improvements in search engine quality and novel visualizations [13] are allowing users to better identify documents that are relevant to their needs. With a few notable exceptions [6, 12, 13, 29] less attention has been paid to how users identify key information from within documents, which can be a significant undertaking. And despite the need for more effective document processing tools, we know little about how users might exploit such tools to scan and extract key information from within documents. This paper develops and evaluates UI techniques of Interactive Compression (IC) that allow users to identify critical areas of interest within complex documents - in this case transcripts of meetings.

IC techniques allow users to interactively reconfigure a document to control the amount of information they view, as well as how that information is presented and contextualized. Early systems provided simple controls that allowed users to interactively excise unimportant regions of text or to de-emphasize these by manipulating their font size or color [16, 18]. Other work looked at improving reading of documents using contextual views and animated typography [31] and related approaches have explored the use of presentation/navigation methods such as fish-eye [10], focus and context [4, 14] or zoomable [2, 5] document views, that allow users to control the level of granularity and focus of the viewed document. Despite the intuitive appeal of IC techniques, with some exceptions [4, 7, 11], we lack information about both the overall utility of IC and the trade-offs between different IC designs.

In principle, IC techniques should allow users to directly...
focus attention on key regions of a document, by promoting those regions while simultaneously demoting unimportant information. However there are various reasons why IC techniques may not be effective in practice. Eye gaze studies show that when reading regular unmodified text, users already possess highly effective strategies for focusing on important regions and ignoring less crucial information [27]. IC techniques may not therefore offer benefits over these pre-existing reading strategies. Worse still, over-zealous compression may lead users to inadvertently demote and overlook relevant information. Or over-compression may de-emphasize the contextual information needed to interpret relevant information. Thus, IC does not necessarily guarantee improved ability to extract information. It is therefore crucial when evaluating IC techniques to compare them against unmodified text.

As well as evaluating overall benefits to IC, we also want to evaluate design trade-offs between specific IC techniques. Excision based techniques, remove unimportant information, leaving the user with less text to scan, but with reduced context for the text that remains. Other techniques highlight important information, de-emphasizing unimportant data, but preserving context. With Highlighting, users therefore have more text to scan, but can see at a glance the context for highlighted regions. A second aim was therefore to examine trade-offs between Excision and Highlighting techniques.

Finally our approach to IC is user configurable allowing users to actively choose a level of compression that fits their information processing needs. For a quick overview they might look at a document highly compressed, but decompress it when they wanted to drill down on details. We wanted to pit our interactive approach against an alternative approach taken in traditional summary research. Typical approaches to summarization assume that there is a single, automatically computable optimal length for a summary [19]. We wanted to evaluate this optimality assumption. Our interactive approach instead assumes there will not be an optimal summary. Instead we anticipate that the level of compression will depend on the user and their particular information needs and interests. Our overall aim is to identify effective design principles for IC systems, by addressing the following specific research questions.

- **Comparison with Unmodified Text**: Are IC tools faster and more accurate than existing user strategies for information extraction, given that users are well practiced with unmodified text? And do users prefer IC techniques over unmodified text?
- **Efficiency**: Which IC techniques allow most efficient and fastest scanning? Excision removes information, leaving less text to scan. However text with portions removed may be harder to understand - which may decrease scanning speed. Highlighting, in contrast, leaves users with more text to scan but the context for the emphasized regions can easily be seen.
- **Accuracy**: Are users likely to make more errors in identifying relevant information with Excision techniques that remove potentially important information? Or are they more accurate with Highlighting where de-emphasized material is always visible?

- **Granularity**: It is possible to compress texts at the word or utterance level. Is it better to remove/de-emphasize words or utterances - in terms of efficiency and accuracy? Word removal/de-emphasis may produce locally discontinuous text, whereas utterance manipulation/removal may produce more locally coherent texts but that have longer ‘missing’ segments of de-emphasized information.

- **Interactivity**: Is there a necessary requirement for interaction when compressing the text? Do users want bespoke levels of compression or are there general optimal compression levels?

To investigate these questions we conducted a two part study. We begin by conducting an informal evaluation of six different IC techniques to obtain subjective user feedback about different compression approaches. We conducted this informal study to focus our more formal evaluation, because there are multiple possible IC methods in this complex design space. Formally evaluating many systems would be extremely costly and time-consuming.

From this informal evaluation we identify the two most successful techniques (Word Excision and Word Highlighting). We then evaluate these more rigorously in a second lab study, comparing these IC techniques with a baseline of unmodified text. Note that our aim in both cases is exploratory and evaluative - we are interested in how users interact using IC techniques, as well as whether the techniques allow users to locate and extract information more efficiently than plain text. We collect behavioral data as well as user reactions to IC tools.

We examine these research questions in the context of extracting information from meeting records. There has been recent interest in the collection and analysis of multimedia
meeting records [1, 6, 8]. Typical meeting browsers present users with meeting transcripts derived by applying automatic speech recognition (ASR) to the audio recording [17]. Such transcripts are typically long, cryptic, dense and poorly structured - making it difficult for users to scan and extract information. Meetings are therefore a potentially promising practical domain for examining the benefits of IC in helping users extract information from complex records.

The structure of this document is as follows. We begin by presenting design principles and a general architecture for producing Interactive Compression. We then describe our two evaluation studies, and their implications.

**IC PRINCIPLES**

We adhered to 4 design principles in our implementation of IC:

- **User Control:** The level of compression should be under user control, allowing users to interactively adjust the amount of information shown in the display.

- **Signalling Importance:** IC should manipulate the salience of important information, either by manipulating the presentation of important information or demoting unimportant information.

- **Ease of Navigation:** IC should not only present views indicating important regions of the source document, but the views should also support straightforward navigation to those regions.

- **Providing Context:** IC should allow users to easily recover source context, allowing them to rapidly focus on background information related to important areas.

Applying these principles, we designed a general architecture (see Figure 2) for IC interfaces. We start with a Document Collection (a set of transcripts generated from meeting recordings) and a Target Document. We perform an Importance Analysis of these documents using text processing methods to rank the importance of all the semantic units that make up the text of each document. The Target Document and associated importance rankings are then passed to one of a set of Interactive Compression Techniques which then renders the document on screen. Different IC techniques take the same Target Document and importance scores, but render the document in different ways.

Users specify their desired level of compression manually (i.e. the amount of information that they want to see) and user requested Compression Changes are passed back to the Interactive Compression Technique which then re-renders the document on screen in real-time. Thus if the user requests a greater level of detail (see Figure 1), we recompute the importance threshold so that information which had previously been de-emphasized is now visible to the user. To preserve the user’s **attentional focus**, the re-rendering process ensures that the line at the center of the screen doesn’t move - thereby allowing users to track specific parts of the document as the compression level changes. We describe these processes in detail below.

**Transcripts**

The transcript collection was taken from the AMI corpus [20]. This corpus consists of 100 hours of multimedia meeting recordings (including audio, video, slide capture, whiteboard capture and a record of personal notes). We used records for 153 meetings in total. The meetings were semi-scripted, where the participants were assigned specific roles for a collaborative task of designing new types of TV remote control (See [20] for details). In addition to the raw recordings there are a large number of human and automatic annotations available (including word transcripts and utterance segmentation). To analyze the transcripts we primarily made use of word transcriptions, utterance segmentation and speaker segmentation.

**Word and utterance importance scores**

There are numerous features that might be extracted from the text to compute the importance of different textual units. For efficiency and simplicity we use a standard information retrieval technique derived from TF*IDF measures [15]. While other techniques [19, 21, 24] have been proposed to compute importance, there is no clear consensus about their utility [22]. TF*IDF allows us to compute the importance of each individual word, and each utterance faster than real-time. TF*IDF measures word importance in a given document as the product of (a) how frequently the word appears in that document (TF) and (b) how infrequently the word occurs in the other documents in the corpus (IDF). Thus an important word in a given document is one which occurs frequently in that document, but infrequently in the overall corpus. We compute the word importance as,

\[
imp_{td} = \frac{\log(count_{td} + 1)}{\log(length_d)} \times \log\left(\frac{N}{N_t}\right),
\]

**Figure 3. Two examples of Excision.** The left panel shows Excision at the word level, the right panel shows Utterance Excision, both at a compression level of 50%
where \( count_d_t \) is the frequency with which term \( t \) appears in document \( d \), \( length_d \) is the number of unique terms in document \( d \), \( N \) is the number of documents and \( N_t \) is the number of documents which contain the term \( t \). Before computing TF*IDF we first exclude stop words (such as the, of, and) using a standard stopword list. Utterance importance is derived directly from word scores; utterance importance is the sum of the TF*IDF score for each non-stop word in the utterance divided by the number of constituent non-stop words in the utterance.

**IC Techniques**

The IC techniques take (a) the importance scores and (b) the user-specified compression level and then render the document on the screen. Compression level is expressed as a percentage, so that a 10% compression level presents 10% of the original transcript and hides or de-emphasizes the remainder.

We apply two main visual transformations (though we also describe variants of these). We either remove unimportant elements (Excision), or alter their visibility (Highlighting). These transformations can be applied at both the word or utterance level. This makes a total of four different IC techniques: Word Highlighting, Utterance Highlighting, Word Excision and Utterance Excision. As well as these four basic techniques we also explore two novel techniques which are described below.

**Excision**

The two Excision techniques remove unimportant *words* or unimportant *utterances* from the transcript. Different techniques are used for utterances and words to indicate where information has been removed. For Word Excision, unimportant words are replaced by a single period (see Figure 3, left hand side). Thus the transcript is shortened but the user can see both the number of removed words and where they were removed from. For Utterance Excision, excised utterances are removed from the text and replaced with whitespace to indicate their removal (see Figure 3, right hand side). Where removal of multiple contiguous utterances has meant that a whole speaker turn has been removed, this is indicated in the transcript with an ellipsis (i.e. three dots). Note that a single ellipsis may represent multiple ‘missing’ speaker turns.

**Highlighting**

The third and fourth techniques Highlight important information. Unlike Excision, Highlighting (see Figure 4) preserves the whole source text, but makes important words or utterances more salient by manipulating text contrast using grayscale. The color mapping causes important units to be colored black, whereas unimportant units are colored on a continuous gray scale, according to their level of importance. With Highlighting the user is therefore directly aware of the context of important units, since Highlighting ensures that the entire source is always visible, albeit faintly for unimportant units.

**Keyword Context and Fisheye Techniques**

We also explore two novel techniques - one which combines Excision and Highlighting and another which modifies the presentation method described in Figure 2 for a more complex fisheye representation.

In our first technique, called Keyword Context, we employ Utterance Excision but instead of removing unimportant utterances completely we present a single keyword from each of those utterances but these are grayed out to de-emphasize them (see Figure 5). Thus the user is aware of where utterances have been removed and is given a small hint about the content of the removed utterances. Keywords are chosen as

Figure 4. Highlighting: The left panel shows Highlighting applied at the Word level, the right panel shows Highlighting applied at the Utterance level, both at a compression rate of 20%
the most important word in the utterance, as measured by the TF*IDF score described above.

In the second novel technique we approximate a classic fisheye [9] view in text. The screen is split into five regions - two periphery views, two contextual views and a verbose document view. The verbose document view contains a short section (5 sentences) of unmodified text. Above and below this, the contextual view shows the most important utterances from a longer section (5 sentences from 20 sentences) of the document above and below the verbose text. At the top and at the bottom of the screen the periphery views show the keywords from each of the remaining utterances presented in the order in which they appear in the document. The user navigates through the document by selecting utterances from the contextual view or keywords from the periphery which shifts the focus of the entire view to the selected point and updates the other views accordingly (see Figure 5).

EXPLORATORY EVALUATION
We carried out an exploratory evaluation of the six techniques to explore the design space, by collecting subjective responses from users about the efficiency, utility and drawbacks of the different techniques.

Meetings and questions
A single meeting (IS1008c) was selected from the AMI hub corpus [20] with hand transcripts and lexical boundaries also taken from this corpus. From previous work identifying user requirements for meeting recall [30], we generated 8 representative questions about what was discussed in the meeting. Questions were piloted with 5 users showing that all questions could be answered using the interface. There were 2 questions for each of the 6 interfaces. The questions explored the ability of users to locate information at specific points (e.g. How much money was spent on X?) and over larger regions of the text (e.g. What did person X contribute to the meeting?).

Procedure
Seven English speakers participated. They were first shown an example of each IC technique and the general procedure of using the interface was explained. We then gave them a question and observed them, encouraging them to ‘think aloud’ about how they were using the IC technique to answer the question. At the end of the experiment, we also asked participants to explicitly compare and contrast the interfaces and to rate their two favorite and two least favorite interfaces for each type of question (specific or region based). The order of presentation of the questions was fixed for each user in order to minimize the effects of users seeing answers to future questions whilst answering the current question. The order of presentation of the interfaces was randomized for each user.

Results
We gave a score of +1 to each technique chosen as first or second favorite, and a score of -1 each time it was chosen as least or second least favorite for each question. Techniques that weren’t chosen in the top or bottom two ranks were given a score of 0. Table 1 shows that the most preferred techniques overall were Word Highlighting and Word Excision. Utterance Excision and Utterance Highlighting were not generally liked as indicated by their negative scores. The scores also indicate that users were ambivalent about the Keyword Context display and that opinions were divided on the usefulness of the Fisheye view. Users stated that Highlighting is more efficient because it preserves document structure and allows them to view local context without having to decompress the document. Users also reported finding it easier to recover the original context with Word-as opposed to Utterance-based techniques. They state this is because with Word-based techniques the ‘gaps’ between visible important regions are relatively small. In contrast, Utterance techniques sometimes omit large amounts of context - making it hard for users to infer what is in the missing text. These assertions are supported by a Wilcoxon test comparing subjective scores for IC type for overall ratings, showing that Word were preferred to Utterance manipulations ($z = -2.412, p < 0.05$).

EXPERIMENT
The results of the first study showed that users had a preference for Word over Utterance techniques and indicated that there was also a general preference for Highlighting. However, whilst we now have a good understanding of user preferences for IC interfaces, a limitation of the exploratory study was that we did not collect behavioral data about the strategies that users employed when interacting with IC documents, or about objective differences in performance in using IC interfaces. For example, we were unable to accurately measure how quickly or accurately users scanned documents with IC interfaces, how they manipulated the compression levels to locate information or how they dealt with situations where the IC caused them to overlook relevant information. Finally we didn’t know whether our UIs performed better than unmodified text - for which users already have sophisticated strategies.

Our follow up study addressed these limitations. We compared our two most preferred techniques (Word Excision and Word Highlighting) with a control - a static, unmodified text transcript. We collected behavioral data about strategies for information extraction, including (a) the level of compres-

<table>
<thead>
<tr>
<th>Technique</th>
<th>Positive Score</th>
<th>Negative Score</th>
<th>Overall Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Highlighting</td>
<td>+7</td>
<td>-0</td>
<td>+7</td>
</tr>
<tr>
<td>Word Excision</td>
<td>+5</td>
<td>-2</td>
<td>+3</td>
</tr>
<tr>
<td>Keyword Context</td>
<td>+2</td>
<td>-0</td>
<td>+2</td>
</tr>
<tr>
<td>Utterance Highlighting</td>
<td>+3</td>
<td>-2</td>
<td>+1</td>
</tr>
<tr>
<td>Fisheye</td>
<td>+4</td>
<td>-6</td>
<td>-2</td>
</tr>
<tr>
<td>Utterance Excision</td>
<td>+0</td>
<td>-8</td>
<td>-8</td>
</tr>
</tbody>
</table>

Table 1. Subjective rankings of different IC techniques. Positive and negative scores are shown as well as the overall rating. Note that some users failed to give two preferences and so the scores don’t sum to zero.
sion that users applied to different regions of the text; (b) which regions users scanned (c) how long they scanned each region and (d) the number of times that a given region of text was scanned. To collect this data, we modified the task from the exploratory evaluation. In that study, we used general questions to probe browsing behavior, but we found high variability in the strategies people used to answer each question, depending on exactly how they interpreted the question. To reduce this variability, in study 2 we asked users to answer specific factual questions that required them to locate information in a particular region of the document, e.g. What material was chosen for the remote control? When users found the region they were asked to manually highlight the phrase that allowing them to answer the question and then press a button to record their finding. Questions were chosen so there was only one region relevant to each question. While this does not exhaust all the tasks a user might want to carry out when accessing a meeting record, it nevertheless offered us a simple, controlled way to collect accurate behavioral data associated with the use and efficiency of our compression techniques. For this task we were therefore able to reliably determine (a) success: whether users found the relevant region; (b) efficiency: how quickly they did it.

Materials
The data used in the experiments was again taken from the AMI corpus [20]. 12 meetings were selected. For each meeting we selected a question for users to answer that required them to locate a relevant target region. Transcripts were an average of 980 lines long, and the mean line number of the target region was 580 (the regions were chosen to occur within the central half of the document). We did this because the beginning and end of meetings are predictable (introduction and summing up) and we did not want users to exploit real world knowledge to guess where target information was likely to be. To ensure that compression could be used to help locate the target region, we confirmed that 75% of the region was visible at the 25% compression level. A pilot indicated that the task was of comparable difficulty for the different meeting transcripts.

Procedure
Twelve English speakers participated in the study. We first explained the procedure and a brief training phase allowed them to experiment with the different IC techniques. During training, users were encouraged to use both the compression slider and manual highlighting operations (to indicate when they had located the target region). In the experiment users were presented with the transcript and the corresponding question simultaneously. Both the order of presentation of interfaces and questions were randomized for each user. Users received a small reward for participating.

For each user we logged the line that was central on the screen and the compression level every time the user interacted with the interface. We also noted whether they successfully selected text that including the target. From these measures we were able to infer:

- The time taken to locate the target region.
- How quickly the user scanned the document (measured as the number of screen lines scrolled per second).
- How many times the user ‘missed’ the target region (counted when the user scrolled past the location of the target phrase).
- The different compression levels the user chose for viewing the document.
- The time spent at each compression level.

After completing the task, users were asked to rank the 3 techniques in order of preference and also provide detailed qualitative comparative feedback about the advantages and disadvantages of each of the techniques.

Hypotheses
We tested five main hypotheses.

H1: Efficiency of IC
Overall, users will locate target regions faster with IC (whether this is Highlighting or Excision) than unmodified text, because interactive techniques allow them to better control what they read and hence allow them to better focus their attention. We expect the greater control to compensate for potential loss of context with IC techniques.

H2: Scanning rate with IC
Users will scan documents faster with IC interfaces than unmodified documents. We also expect users to scan faster
with Excision than Highlighting as there is less information to view with this technique.

**H3: Robustness of IC**

Users will overshoot target regions more often with IC than with unmodified text, because interactive techniques de-emphasize potentially relevant information. For the same reasons, they will pass through the document more times with IC techniques. We also expect that users will overshoot more often with Excision than Highlighting, and additionally pass through the document more times - because Excision removes potentially relevant information from view, and makes the remaining text harder to interpret.

**H4: User Preferences**

We expected an overall preference for IC over unmodified text, because of its greater efficiency but also, as with our first experiment, that users would prefer Highlighting over Excision because it preserves document context, reducing the likelihood of error.

**H5: Preferred Levels of Compression**

Traditional approaches to summarization presuppose an optimal level of compression, we should therefore expect to see a preferred level both within and across users, for both Interactive Compression techniques. There should also be consistency between users on the number of levels they choose.

**Results**

We first present some general results. On average it took users 2.8 minutes to locate the target region. On 63% of occasions they ‘overshot’ the target and had to pass through the document multiple times to locate it. On 20% percent of occasions they failed to locate the target region.

**User behavior in the three conditions**

User behavior was very different in the 3 conditions. Figure 6 shows how different IC techniques led to different scanning behaviors, with Interactive Compression leading to faster scan rates, but inducing more overshoots. With unmodified text (left hand side of Figure 6), users tended to scroll through the document in increments, selecting a window of text, scanning it, and scrolling to the next window, as shown by the incremental step-like character of the graph. Sharper gradient on the graph compared with unmodified text). Highlighting in contrast leads to much faster scan rates (center Figure 6). This may be because they are able to focus better on important text, and ignore irrelevant information. Note too, however, that this user overshot the target but scrolled back to the start of the text and was able to locate it successfully on the second pass. Finally, with Excision, (right hand side of Figure 6) users were able to scan the text relatively rapidly, at high levels of compression, but this again leads them to overshoot the target, and pass through the document a second time. Note that with both instances of IC, although users overshoot the target and pass through the document multiple times, the faster scan rates afforded by IC mean that they still seem to locate the target faster than with unmodified text.

To evaluate H2 and H3 we carried out a 3 (IC Technique) X 12 (Document) MANOVA with Scan Rate, Success, Probability of Overshooting and Number of Document Passes as dependent variables. H1 was tested in a separate ANOVA because it required the removal of data from unsuccessful trials. The other hypotheses were tested in the MANOVA.

**Efficiency**

To evaluate H1 we carried out a 3 (IC Technique) X 12 (Document) ANOVA with solution time as a dependent variable. The trials in which the users failed to locate the target region were removed from this analysis because there was no solution time on these trials.

There was an overall effect of interactive compression technique ($F_{(2,77)} = 3.776, p < 0.05$). As expected, Tukey planned comparisons show that users located target regions quicker with Interactive Compression techniques that with unmodified text ($p < 0.01$), although as Figure 7 (left hand side) shows there was no difference between Excision and Highlighting for solution time ($p > 0.05$).

**Scan Rates**

For scan rate (see Figure 7, right hand side), we found an effect for interactive compression technique ($F_{(2,108)} = 25.319, p < 0.01$) with Tukey planned comparisons indicating that, as predicted, users scanned documents faster with IC ($p < 0.01$) and that users were faster with Excision than Highlighting ($p < 0.01$).

**Robustness**

The results for robustness are shown in Figure 8. Contrary to our expectations, users were able to locate the target region equally successfully with each of the interfaces.
(F(2,108) = 1.546, p > 0.2). However, as predicted, (see Figure 8, right hand side) we found a significant effect of interactive compression technique on the propensity for the user to overshoot (F(2,108) = 7.969, p < 0.01). Tukey planned comparisons confirmed that users were more likely to overshoot with Excision than Highlighting (p < 0.01).

We also found as predicted a main effect for the number of document passes (F(2,108) = 13.817, p < 0.01). Tukey planned comparisons confirmed that users made more passes with IC (p < 0.01), and specifically more passes were made with Excision than with Highlighting (p < 0.01).

**User Preferences**

To assess user preferences we scored technique rankings such that the top ranked technique was given 2 points, the middle 1 point and the bottom ranked technique 0 points. Contrary to our prediction, the two IC methods were equally rated (18 and 16 points for Highlighting and Excision respectively), although as predicted both IC techniques were preferred to unmodified text (2 points).

**Preferred level of compression**

To examine compression levels we started by looking at how many different levels each person used to view a document at. To ensure that the compression level was used to actively view the document, and the user wasn’t in transition between levels, we only analyzed levels where the user spent more than 5 continuous seconds. In addition, we removed from the analysis the end points of full and zero compression since the user begins browsing at the former level and the latter gives the user maximum context and so both were used frequently. To further simplify this analysis we categorized the remaining data into 10 equally sized bins.

Figure 9, (right hand side) shows the distribution of these compression levels across all trials. A (User X IC technique) with level as dependent variable ANOVA showed there is no overall preferred level of compression between users (F(11,840) = 0.668, p > 0.7).

We also examined the number of distinct compression levels used in each trial (see the left side of Figure 9). There was no preferred number of levels used by subjects for either interactive compression technique which was confirmed with an ANOVA main effect for subject with number of levels as dependent variable (F(11,72) = 2.489, p < 0.01).

**User comments and behavior**

In addition to observing user behaviors, we analyzed spontaneous comments made during the study as well as post-hoc user comparisons of the advantages and disadvantages of the techniques. For unmodified text, users noted the advantage of unmodified text over the compression techniques, “All information is present at all times” However they often commented that unmodified text was slower to process than the IC techniques: “[It’s] too slow to search - I had to skim through all words, not just the useful ones” For Excision, users noted that documents can be scanned rapidly and the target region found quickly, but that this was dependent on key terms being currently visible: “Very easy to see keywords. Whole document can be scanned in seconds.”

Figure 9. Plots indicating how users controlled compression level. The left panel shows the mean number of compression levels used for Highlighting and Excision. The right panel shows the absolute levels of compression used (the compression scale has been placed into ten equally space bins with the end points removed)
Despite these benefits, some users expressed misgivings about Excision, because of a fear of overlooking important information:

“I often got the sense [with Excision] that I was missing a lot.”

With Excision people also described how they would use a multi-pass technique interactively - modifying the compression level until they found what was needed:

“if the word is not found immediately when the text is sparse, then [I scroll] over the text more and more with different slider settings.”

With Highlighting, many users appreciated that the technique preserves the context of the important words.

“It allowed you to pick out the important words but also to see them in their context.”

Others noted that this fine detail with Highlighting compromised efficiency.

“so [I] still have to wade through it all, even though it can be done faster.”

Our ‘think aloud’ procedure and observations revealed two other interesting browsing strategies. The first was a tendency for IC users to switch between high compression global views and low compression local views. Particularly with Excision, users would choose a high level of compression, scan the document rapidly until they found a relevant region, and then decompress to scrutinize details. Once their detailed analysis was over, they would recompress to their original global level and iterate the process.

The second observation was that scanning was asymmetric. After decompressing, users occasionally failed to notice highly relevant material that was above the center line of their screen. In contrast they usually spotted relevant material occurring below that line. This suggests that people tended to choose a point of interest after decompression (usually the central paragraph on the screen), but to only scan forward of this. This behavior may be related to traditional reading practices.

DISCUSSION

Our results show the utility of IC in helping users rapidly extract information from long unstructured documents. Even though users already have effective strategies for identifying key regions of unmodified text, we demonstrated that IC techniques allow them to extract information more efficiently. Furthermore, users were able to do this even at the expense of dealing with the occasional errors that IC creates. Even though IC users overshot key information more often, IC allowed users to scan text faster. Even when users needed multiple passes through a document to locate key information, they were still quicker to locate that information than with unmodified text, and they did this without compromising overall success levels. In addition, users prefer IC to unmodified text. And although we used meeting transcripts throughout the experiments, it should be noted that our techniques are general and easily applicable to any kind of textual document.

Our findings validate the principle of interactive user contro; in contrast to traditional summary approaches. These approaches generally assume users want an automatically computed short, static, fixed length representation of a long document [19]. This was not confirmed by our data. For this domain at least, we found that users rarely chose consistent levels of compression when browsing different documents. Instead the choice of compression level seemed to be dictated by contextual preferences. The exact level of compression selected for detailed processing seemed to depend on the user’s current understanding of the document, and its structure. For example, users might rapidly scan the entire document at high compression to gain an overview and locate potential areas of interest. On later passes they might decompress the document and revisit these for more detailed scrutiny. This finding provides strong support for interactive approaches [3] that allow users to specify and control the exact amount of information they see on a moment to moment basis. The user’s view should be based on their current understanding and information needs, rather than a precompiled static summary.

We also identified key differences between IC techniques that have important design implications. Excision allows the user to scan and process the document much faster since there is less text to read. However, because relevant information may be elided or vital context omitted, users are more likely to miss key information. In contrast, Highlighting retains complete context. Users are therefore less likely to miss information with Highlighting, but at a cost of a slower scanning speed. The specific choice of IC algorithm should therefore depend on situational constraints. If the user task demands accuracy, then Highlighting is preferred, because it reduces the risk of overshooting errors. If efficiency is critical, then Excision should be chosen because it supports faster scanning.

Three further design implications follow from our findings. First, we have some indications of the benefits of word-based techniques. Much of the traditional summary literature has explored techniques that excise utterances, but our users preferred word based techniques. As they noted, word-based techniques meant that the ‘gaps’ between important elements tended to be smaller, making it easier to infer the contents of those elided materials. Here we found that word based techniques performed better than unmodified text. Similar results have been found for speech browsing, where word based techniques were also effective[26]. Second, our observational data suggest ways for improving the design of IC browsers. We noted that users tended to switch between global and local views of a specific document, as they moved from high level overview to analysis of details. A simple modification to our browsers might be to provide a toggle to allow users to move between such global and local views more seamlessly. Finally, we observed that users tend to browse asymmetrically, sometimes failing to note important information that was behind their current attentional focus. We might therefore experiment with hyperbolic methods such as fish-eye to present greater context for materials after, rather than before, users’ attentional focus.
There are other algorithmic ways that we might refine our current IC techniques. We currently use the TF-IDF measure of word importance for its simplicity and computational efficiency. However we also plan to explore alternative ways to compute importance measures, for example incorporating prosodic information, or more complex text processing to determine important regions of text [24]. We also want to look at how our techniques fare when we use ASR based meeting transcripts instead of the human generated transcripts we used here. Although other research suggests that speech browsing tools are resilient to automatic transcription errors [25, 28], we intend to test this assumption and to also apply IC methods outside the meeting domain, on other classes of document.

There are also theoretical implications to this study. Our findings have considerable overlap with recent work on Information Foraging [23], which has examined how users allocate attention in complex document spaces. Information Foraging has documented user strategies for identifying regions of high information value, ignoring regions of lower utility, as well as the trade-offs for allocating attention in this way. It has also studied the impact of technologies such as Scatter/Gather [9] browsing on users’ ability to reduce the costs of finding such high value regions. We plan to further explore how Foraging theory might be applied to our findings as well as exploring different forms of information extraction tasks.

In sum, we have developed and evaluated techniques that address an important issue for users in an era of information overload, namely extracting and locating key information from within documents. We have implemented several IC techniques and shown how and why they provide benefits. In future work we intend to extend our techniques, as well as our theoretical understanding of how to address this increasingly important problem.

REFERENCES