Information Visualization (2002) 1, 13–19

www.palgrave-journals.com/ivs

COMMENTARY

Rapid, Serial and Visual: a presentation technique with potential

Robert Spence[1]

[1]*Imperial College, London, U.K.*

Correspondence:
Robert Spence, Department of Electrical and Electronic Engineering, Imperial College, Exhibition Road, London SW7 2BT, U.K.
Tel: 0207 594 6259; Fax: 0207 581 4419;
E-mail: r.spence@ic.ac.uk

Abstract
This commentary addresses the recently exploited presentation technique called Rapid Serial Visual Presentation, or RSVP, and identifies a number of applications. It describes a number of different RSVP modes, discusses potentially attractive features and reports the results of experimental investigations. Some of the many issues requiring further research are identified
Information Visualization (2002) **1,** 13–19. DOI: 10.1057/palgrave/ivs/9500008

Keywords: Rapid Serial Visual Presentation; space-time trade-off; information visualization; web navigation; small displays

Riffling in the physical world

You walk into a bookshop, and see a book whose cover attracts your attention. The likelihood is that you'll pick it up and riffle through its pages in order to acquire some feeling for its content, layout, genre and relevance. And all that in the space of a few seconds. This procedure, appropriately known as Rapid Serial Visual Presentation (RSVP) is so useful in the physical world of books, catalogues and other documents that its extension to electronic information spaces is well worth investigation. How often, for example, have we wished we could quickly riffle through the contents of a poorly labelled folder (Figure 1) to recall its content or to search for a mislaid diagram? Or quickly scan through an online catalogue to locate a product of interest?

In this commentary, the current state of application of RSVP to everyday tasks within the context of electronic information spaces is discussed. As with other information visualization techniques the objective of RSVP is to support the efficient creation of a mental model of some data: it therefore joins many other presentation techniques on the 'palette' of the information visualization tool designer to be used, often jointly with other techniques, as appropriate.

Riffling in electronic information space

An early embodiment of the RSVP technique[1,2] in the context of electronic information spaces was the 'carousel-mode' RSVP shown in Figure 2. A mouse-click on a folder icon initiates a flow of images which emerge from one side of the folder (Figure 2a), follow a roughly circular trajectory around it (Figure 2b) and eventually disappear back into the opposite side of the folder (Figure 2c). Each image might be visible for a total of around 200 to 400 ms, so that as many as 50 images might be viewed in as little as 3 or 4 s. The human ability to recognize the presence of a target image in as little as 100 ms or even less[3] is such that carousel-mode RSVP can support the efficient search for a target image. An image may be pictorial, perhaps showing a Monet or Gaugin, or it may be a surrogate for visual or nonvisual content. It could, for example, be one of a number of representative 'key frames' taken from a video, or a label for a section of a catalogue.

Received: 23 August 2001
Revised: 10 December 2001
Accepted: 8 January 2002

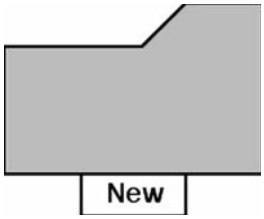

Figure 1 A poorly labelled folder, whose content has been forgotten.

Design dimensions

Familiarity with the carousel mode of RSVP immediately generates a host of questions concerning features such as the effect on performance of visibility time and trajectory. However, although the primary aim is to try to understand the fundamental issues involved, it is helpful first to introduce some of the many trajectories that can be associated with RSVP.

Collage-mode RSVP

Imagine that a travel agent is showing you some photographs of available holiday destinations. To give you some overall feeling for possibilities the agent drops the photographs, one by one but quite rapidly, on to the table between you, spacing them out so that as many as possible can been seen simultaneously before some are inevitably masked by later arrivals. Figure 3 shows an implementation of Collage RSVP[4–6] in the context of an online bookstore. Controls are provided to enable the user to adjust the speed of presentation, to reverse it and, most importantly, to stop it.

Floating-mode RSVP

With Floating-RSVP one sees each image initially in miniature – and perhaps slightly out of focus – in the centre of the display, after which it moves radially in one of a number of possible directions to the edge of the display, expanding in size all the time (Figure 4). The effect is similar to that of driving along a motorway, with various signs approaching and then disappearing from view 'over one's shoulder'.

Shelf-mode RSVP

Inspired by an e-commerce application[7] in which bottles of wine move along a shelf to allow inspection, the RSVP mode illustrated in Figure 5 was proposed. Here, the initial presentation of an image is 'full size' and located at the lower right-hand corner of the display. It remains there for about 500 ms before moving at constant speed, along a linear trajectory, towards the upper left-hand corner of the display, decreasing in size all the time.

Other RSVP modes

It is not difficult to propose other RSVP modes having different trajectories and other features, but the four already described are sufficiently representative to trigger a variety of relevant questions and tentative hypotheses.

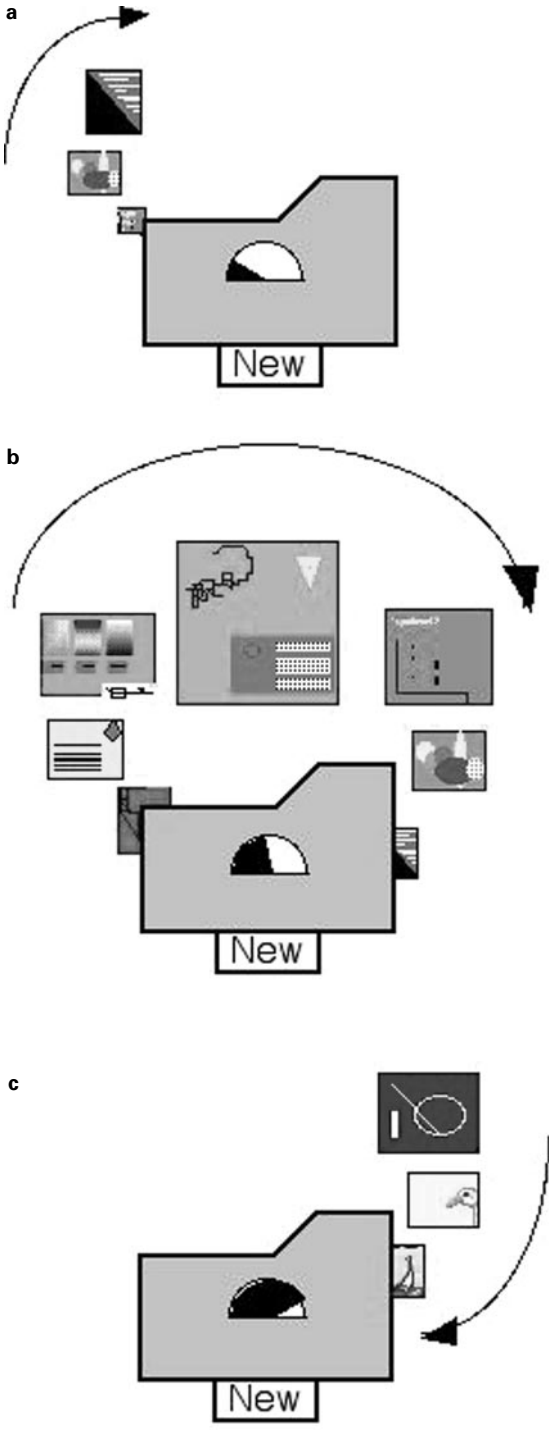

Figure 2 Carousel RSVP. (a) RSVP has started. (b) General appearance. (c) RSVP has almost finished.

First reactions

Together, the four RSVP modes presented above identify some of the major features available to the interaction designer who wishes to exploit this relatively new form of presentation. The carousel-mode, for example, appears

to offer a user the ability to move their gaze momentarily to follow an image whose tentative recognition requires confirmation. Nevertheless, the collage-mode might in addition make more effective use of available display space. A potential advantage of floating-mode RSVP will be appreciated by car drivers who, perhaps despite fog or rain, can initially identify the relevance or otherwise of a road sign at a considerable distance, often as the result of the 'text envelope' of a city name or because an icon is readily recognized by its colour and/or shape. Indeed, in the context of collage- and floating-mode RSVP, Wittenburg *et al.*[4] remark that they allow a user '... to see further and faster into an information space to which they may potentially jump in a single move'. They also point out the potential advantage that, in the application of these two modes, the presentation of outlinks[8] is dynamic rather than static, thereby improving the 'scent' to which the user is exposed and facilitating navigation.[9]

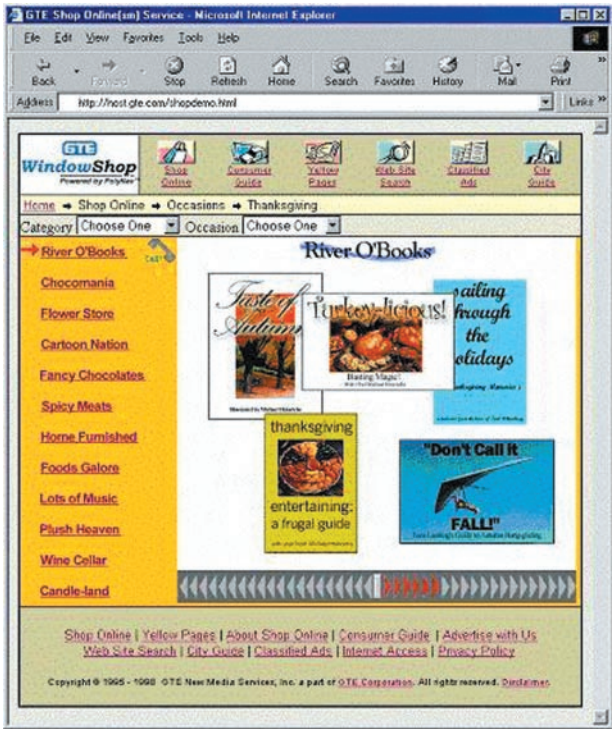

Figure 3 The contents of an online bookstore presented in collage-mode RSVP. The set of arrows just under the presentation allows control of the speed and direction of presentation.

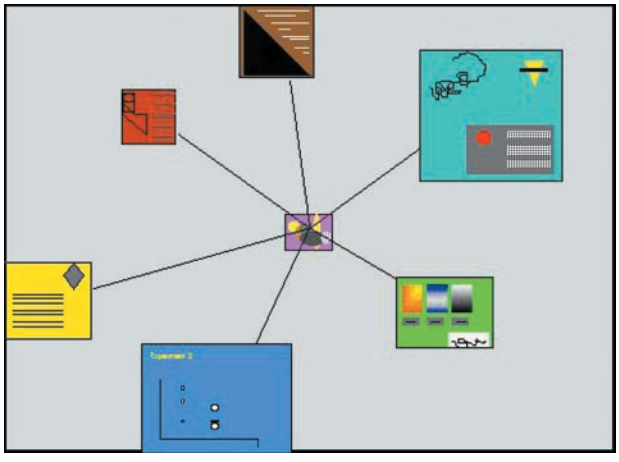

Figure 4 Each image, initially appearing in miniature at the centre, expands as it moves radially. The trajectories, which do not appear in an implementation, are indicated by thin lines.

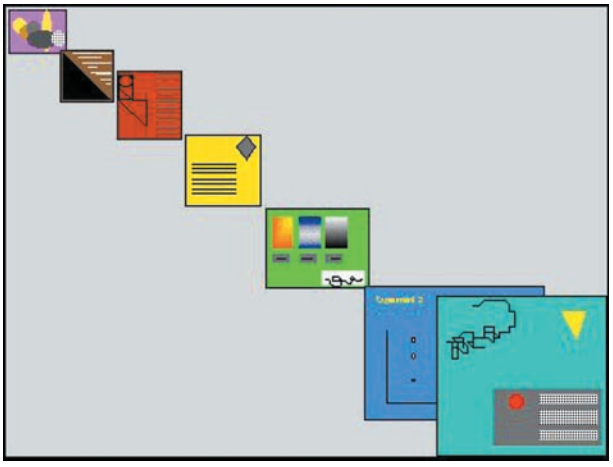

Figure 5 Each image initially appears 'full size' at lower right, dwells there for a short while, then decreases in size as it moves to upper left.

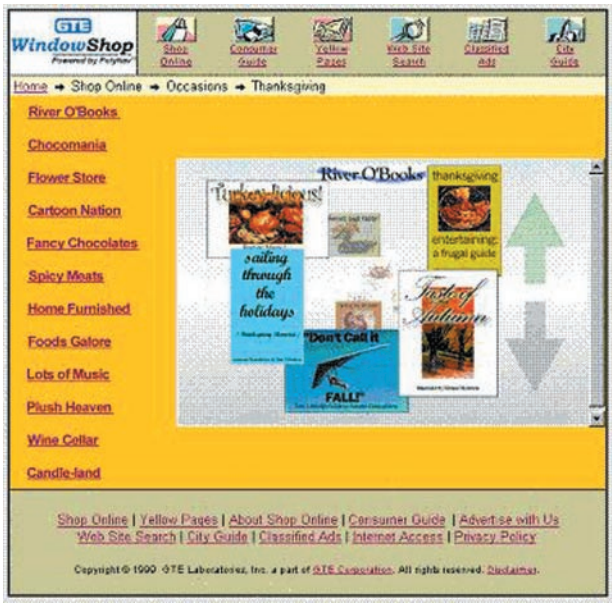

Figure 6 An implementation of floating-mode RSVP for an online bookstore. The arrows on the right allow control of the speed and direction of image flow.

The potential advantage of the shelf mode is also apparent: the user's eye gaze might not often need to move from the location where each image, albeit for a short time, is both stationary and of maximum size.

From even a cursory examination it is clear that the design freedoms offered by RSVP to the interaction designer are so many that a deeper understanding is required, not only of the perceptual and cognitive activities involved but, in addition, the behaviour and attitudes of users.

Experimental studies

Online bookstore

Relatively few experimental studies of RSVP have been undertaken in the context of practical applications. Nevertheless, Wittenburg *et al.*[4–6] implemented collage-mode RSVP to provide the experimental web-shopping interface illustrated in Figure 3. They investigated the acceptability of collage-mode RSVP by offering it as an alternative to two other presentation techniques: one (a 'slide show') involved the temporal display of only one image at a time, always in the same location, while the other was a set of conventional HTML pages. At this point the reader may wish to pause to anticipate, informally and intuitively, the order of user preference.

Users were asked to rate the acceptability of each of the three modes of presentation on a scale from 1 (most favoured) to 5 (least favoured). The 'slide-show' scored best at 1.37 and was ranked significantly higher than the HTML (2.05) and collage (2.58) presentations, a result which may be found surprising in view of the fact that the slide show does not offer the chance of extended gaze fixation to confirm, or otherwise, the identity of an image.

The experiments carried out by Wittenburg *et al.*[4–6] also afforded an opportunity to explore suitable control techniques: mouse rollover was judged to be preferable to radio button selection to halt the display, and also convenient for controlling the speed and direction of presentation (see lower part of Figure 3). Wisely, they advised caution in the interpretation of these early experimental results, since there are many factors in the designs that need refinement. They also devised the floating-mode RSVP whose appearance in the context of an online bookstore is shown in Figure 6.

Video selection

The selection of a video for later viewing obviously requires that some judgement be made of a candidate video in a time considerably shorter than that required to view the entire video! Typically, to support the process of selection, each video is represented by a succession of keyframes chosen to give any viewer a good idea of that video's content.[10] In this context two activities are relevant. One is *gist determination*, the ability to understand the gist or essence of the video from which the key frames are selected, while the other is *object recognition*, the act of identifying, within the set of key frames, an image of interest.

In the context of video selection Tse *et al.*[11] investigated two methods of presentation. They compared an inherently dynamic slide-show presentation with a static presentation ('film strip') of the same keyframes, the thematic order being preserved in each case. Two aspects of human performance were observed, object recognition and gist determination.

Tse *et al.*[11] discovered statistically significant effects, of both presentation mode and rate of frame presentation, upon observed human performance. With slide-show RSVP, for example, no significant deterioration in object recognition performance was detected at display rates up to 8 frames per second, though above this rate performance decreased significantly. Gist determination – a performance not easy to measure – degraded less severely. Notwithstanding the space saved by the dynamic slide show, users generally preferred the static filmstrips. As Tse *et al.*[11] point out, many factors can potentially influence the success of slideshows and filmstrips: the manner in which key frames are selected is one of them.

Keyhole-mode RSVP: a space-time trade-off

The 'slide show' investigated by both Wittenburg *et al.* and Tse *et al.*, and also termed 'Keyhole-mode RSVP', is yet another form of RSVP to be considered alongside the four modes already introduced. Although initially it might not seem to offer some of the advantages associated with the other modes it was, as we have noted, judged to be preferable to collage-mode RSVP and HTML by the subjects involved in Wittenburg *et al.*'s experiments. It was also seen to support a high rate of presentation in the experiments of Tse *et al.*[11] Moreover, Keyhole RSVP does possess one valuable feature that other modes do not share. It was pointed out by Lam and Spence,[12] and commented upon by Wittenburg *et al.*[4–6] and Tse *et al.*,[11] that keyhole-mode RSVP offers *a trade-off between space and time*. Such a trade-off is especially valuable in situations where space is severely limited, as with mobile telephone displays and PDAs.

Exploitation of the space-time trade-off

A specific example will illustrate the power of the space-time trade-off. Effective navigation of the World Wide Web and many other information spaces requires the ability to view not only a current page (Figure 7) but, in addition, two other sources of information.

One source is the collection of visible representations of outlinks, a source that enables an informed decision to be made about the next destination: Wittenburg *et al.*[4–6] have commented upon the value of scent in this context. Second, in view of the frequent need to backtrack, a visible – and selectable – representation of 'footprints' (those pages previously visited) can also ease navigation. Field and Apperley,[13] for example, have pointed out that a potential benefit of selective retreat

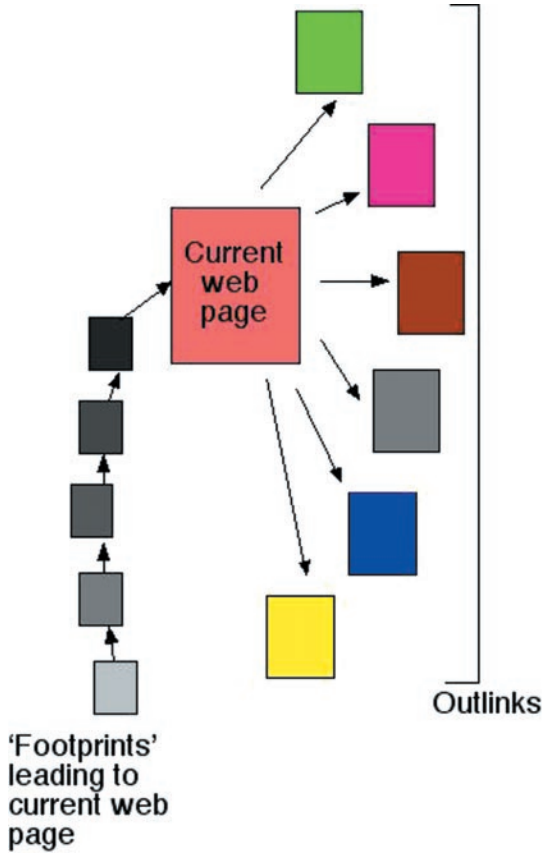

Figure 7 Images supportive of navigation in information space.

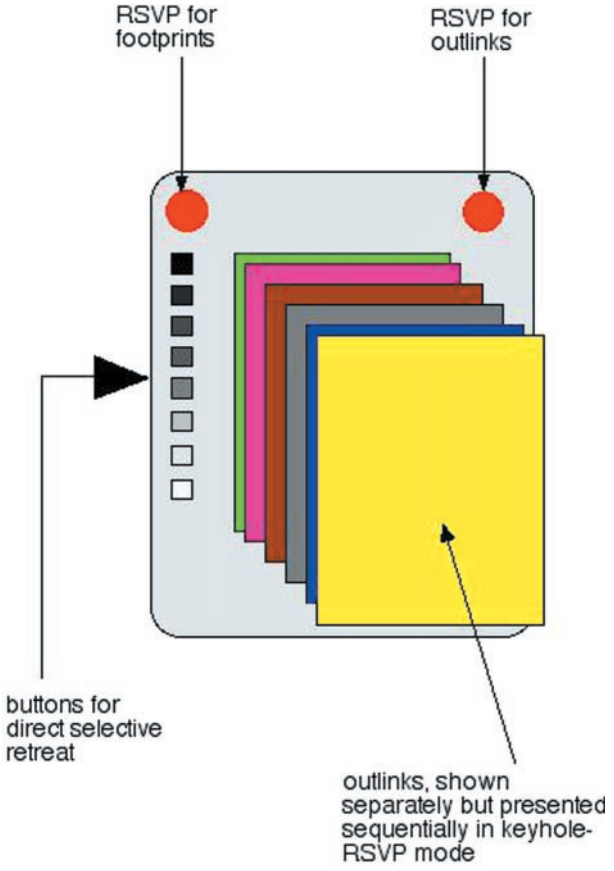

Figure 8 Scheme exploiting keyhole-mode RSVP to allow navigation of information space from a display of limited size.

is a consequent enhancement of the user's mental model of the information space. All these requirements can be satisfied by keyhole-mode RSVP in a manner illustrated in Figure 8.[14,15] One control generates, in keyhole mode, the display of appropriately designed images of outlinks, and in such a way that the presentation can be halted and an outlink selected. Another control similarly handles the footprints.

To some extent the space-time trade-off has been exploited by Wittenburg *et al.*[4–6] with their collage display, in the context of which they remark that their technique 'combines temporal and spatial resources'. Tse *et al.*[11] also pointed out the space advantage of keyhole-mode RSVP.

Human performance with keyhole-mode RSVP

Quite independently of any space-saving advantages, experiments carried out by deBruijn and Spence[16] and deBruijn *et al.*[17] yielded interesting comparisons between carousel- and keyhole-mode RSVP.

Subjects were able to look at a target image for as long as they wished. They were then asked to say if that target was in a seven-image carousel display of a set of 20 images. The exercise was then repeated, with a different time period assigned to the location of an image in each of the seven locations. It was found that, with presentation times as low as 100 ms (when each image would be visible for a total of 700 ms), the presence or otherwise of a target was reported with high accuracy. With the keyhole mode, the presence or absence of the target image was reported accurately even when each image was visible for only 40 ms; with one subject the exposure time was as low as 20 ms. Thus, performance in image recognition required only about 40 ms exposure in keyhole mode compared with about 700 ms in carousel mode. One hypothesis put forward to explain the keyhole-mode performance suggested that, with keyhole RSVP, the user's eye-gaze can only usefully be fixated in one location.

As suspected, the nature of the image influenced subject performance. The experiment was conducted with three different types of image, the extremes of which were subjectively categorised by the experimenters as 'most distinctive' and 'least distinctive'. Generally, longer visibility times were required as the nature of the images varied from 'most distinctive' to 'least distinctive'.

Unanswered questions

While experimental evidence has demonstrated the potential of RSVP it has also highlighted the need for

research that will provide the interaction designer with guidance. Therefore, some of the principal features requiring attention are discussed.

RSVP modes

It is possible to imagine more RSVP modes to add to the five discussed here. Also, features such as the trajectory of the images, their variation in size during presentation, the number concurrently visible, the visibility time and many other features remain to be explored. In doing so it is highly likely that the task being performed by the user, the display device employed and any transmission bandwidth limitations will all exert a strong influence upon the suitability of any RSVP mode for a given application.

Controls

Wittenburg *et al*.[6] have already reported the design iteration that followed studies of the speed and start/stop controls of their RSVP implementations. There may well be a close relationship between the controls and both the task being performed and the value of a residual display following RSVP. For the general application of RSVP to folders and documents, an attractive alternative to the control superimposed on a folder (Figure 2) might be a Magic Lens,[18] available from a menu, whose positioning over a folder or document would give rise to an RSVP display. Such a lens could be provided with the 'elapsed' time display shown in Figure 2 as well as other appropriate controls and indicators.

Eye gaze

It is recognized that a major determinant of image recognition, and with it the efficiency and acceptability of RSVP, may be the pattern of eye movements that a user is required to make. Fortunately, gaze tracking technology has developed sufficiently to ease this type of study. Typically, the output from either a head-mounted or stand-alone gaze detector can provide the experimenter with a trajectory of eye gaze with saccades and fixations identified. Experiments with the collage-, floating-, shelf- and carousel-modes of RSVP[19] have led to useful schematic representations of these modes and initial evidence concerning human behaviour under these conditions. Much more research is needed.

Cognitive mechanisms

Research into the RSVP of both text and images has a long history (see, for example, Coltheart[20] and Intraub[21]) and we do not offer a comprehensive survey here. Rather, two concepts of considerable interest and relevance are remarked upon here.

One concept relevant in the present context is that of Conceptual Short-Term Memory (CSTM), originally proposed by Potter[22] and discussed most recently in Potter.[23] To summarize, Potter points out that we are *continually* browsing – we look at our table, the waiter across the restaurant, the lighting fixture, the wall and so on – but that some internal mechanism rapidly determines whether what we see (or hear) should be forgotten or remembered. The parameters of the CSTM phenomenon would appear to be compatible with the fact that RSVP rates of around 10 per second are found to approach the limit of comprehension. A consideration of CSTM may well prove useful in the derivation of design guidelines.

The 'Floating RSVP' mode poses particularly interesting problems of a perceptual and cognitive nature. The similarity of this mode to the observation of motorway signs while travelling raises the question as to whether significant features of an image can be detected early. If so, two important consequences could follow. First, if an image is deemed to be of no interest, the user could immediately and usefully direct their attention elsewhere; second, the bandwidth needed for transmission might usefully be reduced. Relevant research is being pursued.[24]

Navigation in information space

Experimental evidence would also be useful concerning the later implicit recall of images encountered during RSVP, to establish the extent to which RSVP browsing contributes to an internal model that is subsequently of value in the activity of navigation. Indeed, little is known about the way in which RSVP affects a user's navigation of information space, and research is needed. If the effective encoding of 'scent'[9] is conducive to navigation, then the remark of Wittenburg *et al*.[4] to the effect that 'the presentation of outlinks is dynamic rather than passive, thereby improving the scent' suggests that RSVP may have the potential to enhance navigation.

Landmarks

It is widely appreciated that the provision of landmarks can significantly support navigation in both physical and electronic information spaces. The question naturally arises as to whether landmarks can be useful in the context of RSVP: 'Just after the yellow image is the one you're looking for'. The fact that such 'landmark' information might well be of no value is due to a phenomenon called Attentional Blink.[25,26] Briefly, if a subject trying to identify a target image in a sequence of rapidly presented images is told that the target will appear just after (say) the yellow image, then the attention paid to that yellow image will prevent any subsequent target image being identified for a period that could be as long as 350 ms.[20,26] The effect is the same as if the subject were to blink just after recognizing the yellow image. Thus, the use of landmarks in the context of keyhole-mode RSVP must be treated with caution.

However, in contrast to keyhole-mode RSVP, designers of floating-mode RSVP applications might reach a different set of conclusions regarding landmarks. Wittenburg *et al*.[4] mention being inspired by the navigation practices of seagoing Micronesians in their designs. Historical approaches to navigation by the Caroline Islanders make use of relative landmarks in fixing their position. In parti-

cular, the relative positions of islands to the side of their course (visible or not) are used to determine progress towards their destination. In a floating-mode RSVP context, relative landmarks could perhaps be used similarly.

Content and resolution

Experimental evidence already suggests, and not surprisingly so, that the nature of an image can influence its recognition within RSVP. Detailed knowledge of this influence would be of considerable value in the design of applications such as online shopping.

Conclusions

There is little doubt that the various modes of RSVP offer a powerful presentation technique. While relevant research into RSVP has been proceeding for many years, its potential to inform design is now beginning to be realized and provides a rich source of interest to those concerned with information visualization, whether psychologist, interaction designer or entrepreneur.

Acknowledgments

Discussions with Dr. Oscar deBruijn and Dr. Kent Wittenburg are very gratefully acknowledged. Kent Wittenburg's provision of images is especially appreciated. Collaboration with Kevin Lam, Kenny Tong, Ping Lee and Min Leung led to useful insight. Discussions with Henke-Jelle Reitsma, Nieske van Woerden and Alistair Jackson provided valuable stimulation.

References

1 Spence R. A content explorer. *Information Engineering Report 98/08*. Department of Electrical and Electronic Engineering: Imperial College, London, 1998.
2 Spence R. *Information Visualization.* Addison-Wesley: Harlow, UK, 2001.
3 Healey CG, Booth KS, Enns JT. High-speed visual estimation using preattentive processing. *Transactions on Computer-Human Interaction* 1995; **3:** 107–135.
4 Wittenburg K, Ali-Ahmad W, LaLiberte D, Lanning T. *Rapid-fire image previews for information navigation.* Proceedings of the Working Conference on Advanced Visual Interfaces (AVI'1998) (Aquila, Italy, 1998), ACM Press: New York, 76–82.
5 Wittenburg K, Nicol J, Paschetto J, Martin C. *Browsing with dynamic key frame collages in web-based entertainment video services.* Proceedings of the IEEE International Conference on Multimedia Computing and Systems (Florence, Italy, 1999), IEEE: Piscataway, NJ, USA, 913–918.
6 Wittenburg K, Chiyoda C, Heinrichs M, Lanning T. *Browsing though rapid-fire imaging: requirements and industry initiatives.* Proceedings of Electronic Imaging 2000 (San Jose, 2000), SPIE: Bellingham, USA, 48–56.
7 Witkowski M, Arafa Y, deBruijn O. *Evaluating user reaction to character agent mediated displays using eye-tracking technology.* Proceedings of AISB'01 Symposium on Information Agents for Electronic Commerce (York, UK, 2001), SSAISB: Brighton, UK, 89–98. ISBN 1 902956205.
8 Furnas GW. *Effective view navigation.* Proceedings CHI'97 (Atlanta, USA, 1997), ACM: New York, 367–374.
9 Pirolli P. *Computational models of information scent-following in a very large browsable text collection.* Proceedings of CHI'97 (Atlanta, USA, 1997), ACM: New York, 3–10.
10 Komlodi A, Marchionini G. *Key frame preview techniques for video browsing.* Proceedings of Digital Libraries'98 (Pittsburgh, PA, USA, 1998), ACM Press: New York, 118–125.
11 Tse T, Marchionini G, Ding W, Slaughter L, Komlodi A. *Dynamic key frame presentation techniques for augmented video browsing.* Proceedings of the Working Conference on Advanced Visual Interfaces (AVI'1998) (Atlanta, USA, 1998), ACM Press: New York, 185–194.
12 Lam K, Spence R. *Image browsing: a space-time trade-off.* Proceedings of INTERACT'97 (Sydney, Australia, 1997), Chapman and Hall: London, 611–612.
13 Field GE, Apperley MD. Context and selective retreat in hierarchical menu structures. *Behaviour and Information Technology* 1990; **9:** 133–146.
14 deBruijn O, Spence R, Lee P. RSVP offers a useful trade-off for information presentation design: a discussion and experimental results. *Information Engineering Report 00/01.* Department of Electrical and Electronic Engineering: Imperial College, London, 2000.
15 deBruijn O, Spence R, Tong CH. *Movement in the Web.* Interactive video poster, Companion Proceedings CHI' 2001 (2001), ACM Press, 209–210.
16 de Bruijn O, Spence R. *Rapid Serial Visual Presentation: A space-time trade-off in information presentation.* Proceedings of the Working Conference on Advanced Visual Interfaces (AVI'2000) (Palermo, Italy, 2000), ACM Press: New York, 189–192.
17 deBruijn O, Spence R, Tong CH. *Movement in the Web.* Imperial College Television Studio, Production 1428(A). Also published in the CHI'2001 Video Proceedings, 2000.
18 Stone M, Fishkin K, Bier E. *The movable filter as a user interface tool.* Proceedings CHI'94 (Boston, USA, 1994), ACM: New York, 306–312.
19 deBruijn O, Spence R. Patterns of eye gaze during rapid serial visual presentation. *Information Engineering Report 01/03.* Department of Electrical and Electronic Engineering: Imperial College, London, 2001.
20 Coltheart V. (Ed). *Fleeting Memories: Cognition of Brief Visual Stimuli.* Cambridge, MA: MIT Press, 1999.
21 Intraub H. Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Perception & Performance* 1980; **6:** 1–12.
22 Potter M. Very short-term conceptual memory. *Memory and Cognition* 1993; **21:** 156–161.
23 Potter M. Understanding sentences and scenes: the role of conceptual short-term memory. In: Coltheart V (Ed.). *Fleeting Memories: Cognition of Brief Visual Stimuli.* Cambridge, MA: MIT Press, 1999; 13–46.
24 Fiser J, Subramaniam S, Biederman I. Size tuning in the absence of spatial frequency tuning in object recognition. *Vision Research* 2001; **41:** 1931–1950.
25 Raymond JE, Shapiro KL, Arnell KM. Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance* 1992; **18:** 849–860.
26 Coltheart V. For an online demonstration of attentional blink, see http://currawong.psy.mq.edu.au/~alice/fleeting/about.html (accessed 1999).