

Will Anyone Really Need a Web Browser in Five Years?

V. Michael Bove, Jr.

MIT Media Laboratory

<http://www.media.mit.edu/~vmb>

Introduction: The Internet as Phenomenon becomes The Internet as Channel

For the past two or three years, “The Internet will kill television,” has been a commonplace. But what does this scenario really imply? Does it mean that our living rooms will go unused, as we spend all our discretionary time in solitary corners of the house hunched over PC keyboards and screens?

Let’s instead examine for a moment the possibility that television kills the Internet. When I suggest that, I don’t mean that the Internet will cease to exist, only that the Internet-as-primary-phenomenon will cease to exist. In other words, when that point is reached, the typical viewer – even of video delivered using the Internet Protocol – will consider the activity in which he or she is engaged as television watching, and would no more dream of saying, “I’m surfing the Net for streaming video” than he or she would think of saying, “I’m consuming electricity.” It’s a goal of my research group and several others that the definition of “television watching” broadens to encompass a number of activities which currently take place on the PC screen, but that at the same time it remains fundamentally a television-like experience. The television screen in most homes occupies an enviable architectural location compared to the PC, but it also has lower screen resolution, is equipped with less intimate interaction means, and often is used by more than one person at a time – commonly with a lower level of attentiveness and a higher level of relaxation than PC users. Thus adding a keyboard and filling the screen with text or icons is probably not the best way to take advantage of the net connectivity that television displays are starting to gain.

Migrating the Service, not the Web Page, into the Environment

What has to happen before the Internet becomes just a medium is the disassociation of the services from the Web browser. Currently, the browser on the PC screen remains the preferred venue for a variety of activities, particularly reading text, browsing images, filling out forms, downloading streaming media and other files, ordering merchandise, sending/reading mail, and participating in chat rooms and other forms of synchronous and asynchronous interpersonal communication. While the generality and extensibility of the browser enabled the development (or in some cases migration) of these services, that very generality makes the typical current Web browser non-ideal for most of them. But the PC has until recently been the only device in the typical home or office with appropriate connectivity and programmability (the latter crucial in the formative days of services).

There is no question that as connectivity and computation spread to other devices, a plethora of Internet-related activities will migrate as well. The concern is that of taking the Web-browser metaphor so seriously that everything else from an e-book to a television program is expected to emulate it in a literal way. Consider the Web-enabled mobile phone: why should a fundamentally audio-style device, which in its main mode of use is held in a position such that any display won’t be in the line of sight, be given net access by means of a tiny text screen? Similarly, we’ve probably all seen the prototype refrigerators or microwave ovens with Internet Explorer on their doors; there may be a variety of excellent reasons for appliances to have net connectivity but it’s more valuable if it actually relates to the device’s main function

in the household, rather than simply taking advantage of the fact that the appliance provides a flat surface on which to hang an LCD screen.

Is it Still a Browser if it Looks Like Something Else?

But besides the inappropriateness of the user interface for non-PC circumstances, what else is wrong with the traditional browser model? I suggest at least the following:

- *Lack of community support:* Many activities performed across the Internet are fundamentally about finding and interacting with a community, and many others (such as shopping) are on-line versions of activities that in the physical world are basically social functions. What kinds of additional support can be provided so that users can find and interact with others easily? How do we better work with a situation in which the ongoing flow of information is among a group of peers rather than mostly one-way between a publisher and a consumer? And how can a Web browser or equivalent support sharing of experiences with others, either in the same physical space or across cyberspace?
- *Lack of context awareness and automatic adaptation:* This point becomes particularly relevant because of the rapid scaling of Internet content and services beyond the size at which an individual (or even a search engine) can maintain a useful model of “what’s out there.” A valuable addition would be the adaptation of one’s view of Web services to correspond to other simultaneous activities, history, or circumstances.

Four prototypes developed by the Object-Based Media Group at the Media Laboratory over the past several years attempt to address these issues in the case of the television display.

The *Vision Television* system (Figure 1) adds a television camera, a microphone, and an Internet back channel to a television receiver. Face-finding software locates viewers around the room and the system connects them with a community of other people watching the same program, who can chat among themselves as if they were in the same room. Given video content that contains provisions for personalization, it can also use information gathered from the faces (number, location, attentiveness, expression, identity) to modify the presentation.

Reflection of Presence (Figure 2) is a distributed multipoint conferencing system in which users are segmented from their backgrounds and composited into a shared space, which appears to the viewer to be an “intelligent” mirror. [1] The system also effectively gives the users access to on-line images, sounds, and video, letting them share and interact with bookmarked information while using the system. Thus it can be seen as an example of a browser that several people can use simultaneously, and indeed one which they inhabit.

We have developed object-tracking and identification software that permits a video editor to make objects clickable, thus letting a viewer see additional information overlays, add items to a “shopping cart,” or jump to other video segments in the same way that Web pages of text link together. *HyperSoap* is a hyperlinked video drama closely patterned after daytime television soap operas (Figure 3), produced with assistance from retailer JCPenney. [2] Furnishings, clothing, and props are all selectable. In an augmented-broadcast scenario, selecting an item provides an information box on the screen; the playback system remembers all objects that have been selected and at the end of the program can provide more interactions based on the selections (for example, using a back channel to take the viewer to an on-line shopping page). If local

storage such as a disk recorder is available, the program can be interrupted and purchasing or other interactions inserted into the program while the incoming stream is being recorded for playback after the break; in this case we found it most effective for the system not to interrupt until an author-specified appropriate point, typically when an actor has finished speaking a set of lines.

HyperSoap was basically a linear presentation with augmentations. To prototype a nonlinear hyperlinked video application such as might be available from a video-on-demand system or a DVD, we created a “web” of video clips with the help of WGBH television and Julia Child. In *Interactive Dinner at Julia’s* Julia Child introduces the viewer to a dinner party at which a succession of courses are served (Figure 4). The viewer simply interested in menu ideas can watch the program with no interactions, but selecting food, drink, or accessory items leads not only to explanatory text overlays (as in *HyperSoap*) but also to illustrated preparation instructions. At the end of a video clip the viewer is returned not to the exact jumping-off point but to a slightly earlier establishing shot to re-establish context, since the viewer may have been to a variety of places in the meantime. In accordance with the “web” idea, we also created a “back button” which appears on the screen when a video link is pursued, showing a small image of the scene to which the viewer will return at the end of the clip; the viewer can return at any time by selecting the button. Additional stacked back buttons appear when following nested video links, so that the viewer can immediately return to any of the previous clips.

Conclusion

Yes, in five years you will still need a Web browser. But instead of just one, on a daily basis you will likely have contact with literally dozens of devices that speak the hypertext transfer protocol, and the ways in which they provide access to content or services will be more appropriately related to the devices’ main function in the world. In the area of television, this means something rather different than a literal carrying-over of the standard content and interaction modes from the PC screen, and ideally it points the way to a whole new generation of services.

Thanks to the many students and colleagues who have worked on this and related research, in particular Stefan Agamanolis and Jonathan Dakss. The projects described in this paper have been supported by the Digital Life Consortium and the Broadcasting Special Interest Group at the MIT Media Laboratory.

References

- [1] S. Agamanolis and V. M. Bove, Jr., “Multilevel Scripting for Responsive Multimedia,” *IEEE Multimedia*, 4:4, October-December 1997, pp. 40-50.
- [2] V. M. Bove, Jr., J. Dakss, S. Agamanolis, E. Chalom, “Adding Hyperlinks to Digital Television,” *SMPTE Journal*, 108, November 1999, pp. 795-801.



Figure 1. The *Vision Television* receiver at the Media Laboratory connects television viewers to a video chat room involving other viewers of the same program. Face-finding software locates people around the room.



Figure 2. In our *Reflection of Presence* multipoint conferencing system, the on-line information shares the space with the users: reaching out toward the “mirror” brings a menu which can be used to call up images, video, and/or audio which become the meeting background.



Figure 3. *HyperSoap* embeds hyperlinks in a television broadcast, allowing the viewer to place items in a Web shopping cart simply by selecting them in the video.



Figure 4. A frame from *Interactive Dinner at Julia's* showing "back" button in upper left. The viewer is here one link away from the main program. Note that items may link to text overlays as well as video clips. (Image copyright WGBH Educational Foundation. Used by permission.)