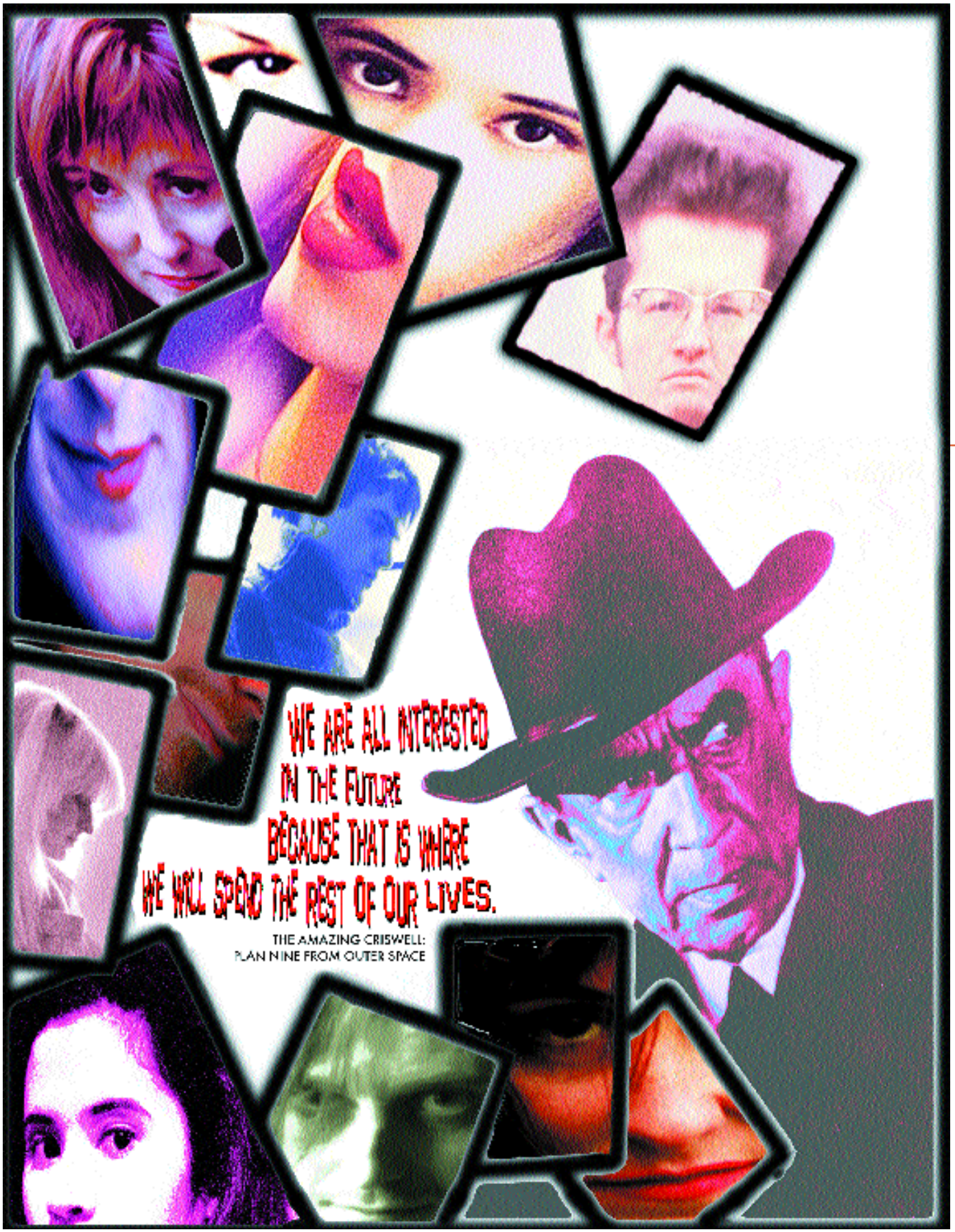
M A R C D A V I S
www.marcdavis.me

P U B L I C A T I O N S
info@marcdavis.me

Garage Cinema and the Future of Media Technology

Bibliographic Reference:

Marc Davis. "Garage Cinema and the Future of Media Technology." *Communications of the ACM (50th Anniversary Edition Invited Article)* 40 (2 1997): 42-48.



RALPH ALFONSO

Marc Davis

Garage Cinema and the Future of Media Technology

ALTHOUGH WE ARE LIVING IN THE "COMPUTER AGE," THE FULL IMPLICATIONS of computational ideas have not been realized in our century. We are at the early apparatus phase of computational development—the profound ideas of computation have not yet affected all other fields of human inquiry, especially our thinking about media. As computational ideas transform our thinking about media, new apparatuses and new ideas will emerge that will change our relationships to media and to each other. The ways we create, communicate, and play will become computationally revisioned, transforming us in the process.

Motion Pictures and Computation

The twentieth century saw the invention and development of two fundamental, new technologies for creating and manipulating representations of the world: motion pictures and computation. Motion pictures gave us the ability to capture and construct sequences of moving images that enabled the creation of a new language of storytelling and visual experience. Computation provides a method for constructing universal machines which, by manipulating representations of processes and objects, can create new processes and objects, and even new machines. The deep integration of computation and motion pictures has not yet occurred, but the implications of their deeper integration over the next 50 years will have profound technological, linguistic, and social effects. This essay traces part of the history and future of computational motion pictures as well as the cultural factors this technology will draw on and foster.

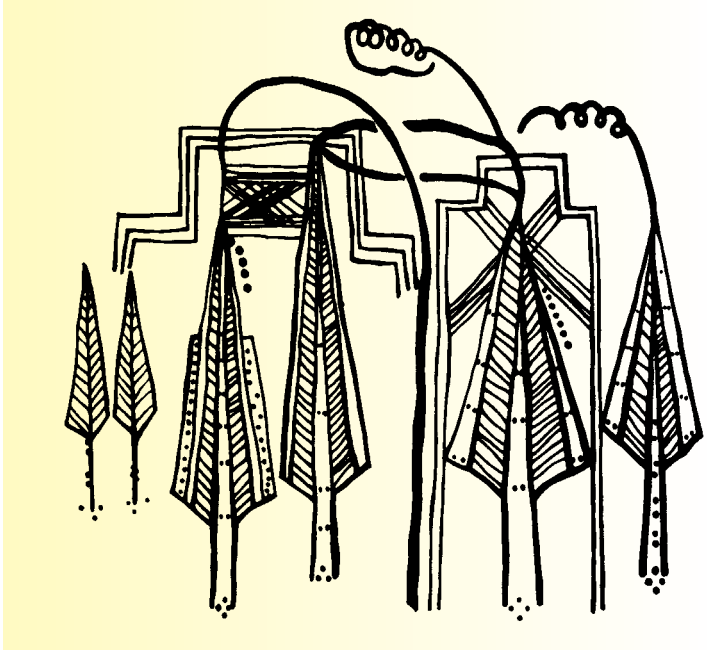
What I am interested in and what I think will happen has implications that operate on the scale of centuries: changes in the forms and possibilities of language, communication, and human expression. We are on the verge of a monumental change—like the invention of writing—that will arise out of the still-evolving transformations of the television, camcorders, and computers. What the next 50 years of computational motion pictures will bring is a fundamental change in the possibilities of "written" language and communication, and I am not talking about email. We can begin by looking back at the his-

tory of writing in order to understand the future of media technology.

Semasiography and the Future of Media Technology

Commonly, writing is understood to represent speech. Yet there exist systems of inscription that do not record and transmit speech. Geoffrey Sampson offers a taxonomy of writing systems that is helpful in thinking about what writing records, transmits, and enables to be constructed [9]. In the first two levels of his taxonomy he divides writing systems into *semasiographic* and *glottographic*. This distinction focuses on what a writing system represents: semasiographic writing represents "meanings" (from the Greek *semasi* "meaning"); glottographic writing represents "sounds" (from the Greek *glotto* "tongue" or "language"). Most of what people commonly think of as writing is glottographic—a notational system for recording and reproducing human speech. Glottographic writing can be understood as a sort of primitive tape-recording system that selects certain salient features of speech in order to enable the reader to reproduce the recorded speech.

Semasiographic writing has very different forms of organization that resemble the conventions of the visual arts more than those of spoken language. An excellent example of semasiographic writing is given in [9], from which the image appearing here has been reproduced. The image, entitled "Yukaghir Epistle," depicts a letter sent by a young woman of the Yukaghir tribe of northeastern Siberia to her estranged boyfriend (try interpreting the image before looking at its explanation in the caption, which appears on the following page).



Speech is not being recorded and reproduced in the exchange of this document, but the image functions as a coherent system of graphs the writer and reader can use to communicate with one another. Conventions of line, layout, size, and symbols govern the syntax and semantics of this writing system, yet this semasiography has no single "translation" into human speech. It is a semiotic system of a different order with its own principles of organization and use, its own history and future. One way to think about semasiographic writing systems is to imagine what it would mean to "translate" this epistle into a short movie.

There are many other examples of semasiography: the numerous semasiographic writing systems of the North American Indian tribes [5]; the elaborate system of Baroque allegorical painting with multiple levels of rhetorical device and mythological allusion, which has its roots in the rhetorical art of memory [10]; Blissymbolics, which, arising out of the fascination with a universal language that has possessed the West since the 16th and 17th centuries, is the most thorough example of an attempt to create an international semasiography [1]; and, in our period, the growing number of international icons used in traffic signs, equipment instructions, and packaging [7].

Most writers about the transition from orality to written media and from written media to computational media have traditionally assumed that written media encompass glottographic, *but not semasiographic writing*. This is in large part due to the lack of a history of the development of semasiographic writing systems. One could explain the lack of such a history in two ways: either there has been no significant development in semasiographic writing systems over the last 10,000 years; or the classification of semasiographic writing systems as the "visual arts" has hidden the existence of such a development from the historian's view. I believe the answer lies between these positions. For important technological reasons there has been very little possibility of the development of image-making technologies into semasiographic writing systems until the twentieth century. In traditional semasiographic writing systems the production of the images themselves requires considerable skill and time. A semasiographic system using tens of thousands of images is much more difficult to

write than a glottographic system like our alphabet in which one only has to learn to produce 26 simple characters. On the other hand, a semasiographic system can be much easier to *read* than a glottographic system. Semasiographic writing can engender a more immediate affective response (even for those who may not have mastered the language) and can create more compact "documents" than glottographic writing systems (if a picture is worth a thousand words, a movie can be worth many more).

The invention of motion pictures (including television and camcorders) is the most significant development in semasiographic writing since we first learned to scrape stone in order to form crude pictograms. Throughout the twentieth century we have been surrounded by advances in the display, transmission, and communicative language of moving images. Today we are constantly *reading* images and image sequences, but we *write* very little. Future computational semasiographic technology has the possibility of initiating a revolution in the development of semasiographic writing systems by making it far easier to *write* semasiographic motion pictures. We now stand on the verge of a major series of developments in semasiographic writing systems—with the integration of motion pictures and computation—such that they may come to rival glottographic writing systems in accessibility, expressiveness, and use value.

Computational Ideas and Media Technology. Although most writers focus on the invention of the printing press or desktop publishing as significant recent developments in writing technology, the most significant advance in the development of writing technology for both glottographic and semasiographic writing systems (even more so than the invention of motion pictures), is the idea of computation itself. Alphabetic writing's analysis of speech into composable components and print's mechanization and uniformity of alphabetic writing paved the way for computation's transformation of these technologies into an entirely new phenomenon. With computation, we not only can write about things and processes, but we can describe them in a language that enables us to create variants and entirely new things and processes from these descriptions. With computation, writing becomes a technology of representing, manipulating, and creating things and processes. With computational writing, writing is no longer merely a digital sampler, but an information processor and synthesizer.

Computational writing creates a class of texts that not only can be read, but that can *write*. With computation, we can now use writing to create textual machines that can manipulate symbols, affect the world, and ultimately shape ourselves. This is because of the idea of the universal machine: computational machines can simulate any machine. They can even accept new inputs to create new computational machines. In the history of writing systems, computation is the height of the development of glottographic writing systems that capture temporal phenomena (speech) and transform them into spatial representations (writing). What remains to be seen is how the invention of computational writing may shape the future of semasiographic writing systems.

Movies as a Mother Tongue. Walter Ong has argued that with computer technology we are on the verge of returning to a form of "secondary orality" or as Marshall McLuhan (informing Ong's vision) would say that we are becoming a "global village" in which speech and sound regain ascendancy over image and text [6, 8]. I would argue, however,

FIGURE DESCRIPTION: THE FOLLOWING DESCRIPTION OF THE YUKAGHIR EPISTLE IS FROM [9], © 1985 BY GEOFFREY SAMPSON. "THE CONIFER-SHAPED OBJECTS... ARE PEOPLE. THE SECOND FROM THE RIGHT IS THE WRITER (THE ROW OF DOTS REPRESENTS PLAITED HAIR AND THUS SHOWS THAT SHE IS A WOMAN); THE NEXT ONE LEFTWARDS, THE RECIPIENT OF THE LETTER, WAS PREVIOUSLY HER LOVER, BUT HAS NOW GONE OFF TO LIVE WITH A RUSSIAN WOMAN (PLAITED HAIR, TOGETHER WITH A SKIRT WITH PANNIERS DISTINGUISHING RUSSIAN FROM YUKAGHIR COSTUME). THE RUSSIAN WOMAN, NATURALLY, HAS BROKEN UP THE RELATIONSHIP BETWEEN WRITER AND ADDRESSEE (LINE FROM THE HEAD OF THE RUSSIAN WOMAN CUTTING THROUGH THE LINES JOINING THE TWO YUKAGHIR); NEVERTHELESS, THE NEW *MÉNAGE* IS STORMY (CRISS-CROSS LINES LINKING THE TWO). THE WRITER IS UNHAPPY (CROSSED LINES) ALONE IN HER HOUSE (THE RECTANGULAR ENCLOSING STRUCTURE), AND SHE IS STILL THINKING OF THE ADDRESSEE (CURLY TENDRIL REACHING TOWARDS HIM). ON THE OTHER HAND THE ADDRESSEE SHOULD BEAR IN MIND THAT THERE IS ANOTHER YOUNG MAN AT HOME (FAR RIGHT) SENDING A TENDRIL TOWARDS HER. IF THE ADDRESSEE WANTS TO ACT ON THIS MESSAGE, HE HAD BETTER HURRY BEFORE HIS NEW HOUSEHOLD HAS CHILDREN (TWO SMALL CONIFERS ON THE LEFT)." [9, P. 29]

that we are on the verge of a transformation that is not merely another phase in the development of glottographic writing technology, or a technological return to an oral culture, but the taking up of a technological and cultural line of the development of visual communication technology that has largely remained dormant over the past 10,000 years: the development of a computational semasiographic writing enabling us to use computational video as a mother tongue.

To gain a picture of what I am thinking about, let us engage in a thought experiment. Dolphins have the ability to send out sonar signals and to receive images of their environment. For them, this form of "vision" is their primary sense. Imagine that dolphin speech, which has proven so indecipherable to humans, is not based on the structural and semiotic principles of human orality (which glottographic writing seeks to transcode), but represents entirely other principles of communication, those which semasiographic writing formalizes and visualizes. Imagine that dolphin speech is the exchange of sonar reflections, in other words, that dolphins have the ability to send sonar reflections as well as to receive them. This would mean that dolphins communicate images as their primary form of "speech." One must then ask, at what rate are these images sent? If we imagine a minimal rate of 24 images per second, then dolphins would not send images to each other, but movies.[1] What would a conversation look like between dolphins? Or more to the point here, between humans who, using computational prosthetics, could write computational semasiographic video from an early age? We are on the brink of finding answers to this question.

Toward a Language of Computational Media. Computational media hold the promise of a new growth of semasiographic writing systems combining the power of computational writing and motion pictures. The most challenging questions lie in conceptualizing a more fundamental integration and hybridization of computational functionality and motion pictures than we have today. How can video be used functionally, how can we program it and program with it? How can we use its ability to record and construct events in the world in a way analogous to how computational writing captures processes as objects and uses them to build new processes? These are the open questions that will occupy the next decades and fundamentally change the ways humans communicate, learn, and create.

Garage Cinema and the Next 50 Years of Computing

Over the next 50 years we will witness an explosion of access to and production and distribution of video by communities that could not earlier afford to produce video in their homes, schools, and offices. Just as desktop publishing gave consumers the power of the printing press on their desks (but it took the Internet to make everyone a publisher since without it the distribution channel was lacking), and digital audio samplers gave birth to a whole new genre and population of music makers, computational video technology will enable these and new communities to make video a part of their daily communication. In the spirit of garage bands, I think of this new population of motion picture producers as practitioners of "Garage Cinema."[2] It is what scratch, slash, rap, home video, and "a TV, two VCRs, and a cable" will become. These are the people who in the next century will be running a TV station/movie studio out of their garages. There are already communities that engage in making Garage Cinema. With a TV, two VCRs, and a cable, fans of various television programs have for years been making their own movies out of found materials. This artistic and social practice has been studied by Henry Jenkins and reveals the ways in which media fans customize the audio-visual materials of popular culture into new motion picture artifacts that meet their own community's needs [4].

Today, current cultural practices of repurposing popular media give us a glimpse of how people might use computational media in their daily lives if video sequences could be quickly and easily assembled, retrieved, processed, and transmitted like dolphins sending and receiving their sonar "movies" or like the conversations of people raised to use computational video as a mother tongue. Looking at fan video-making practice one can piece together a vision of a Garage Cinema in which user selected/produced content is interwoven with the expressive repertoire of materials drawn from popular culture (e.g., movies, TV, news, cartoons) to enable people to create partici-

[1] In conversation with Marvin Minsky, he mentioned that dolphin images may actually be three-dimensional because the sonar reflections contain within them both range information for distance as well as Doppler shift information indicating motion vectors. Concatenating these "images" together would result in a type of holographic dolphin cinema.

[2] A term coined by Michael Johnson, a Ph.D. graduate from the MIT Media Lab, who is now at Pixar.

patory communities around their experience of media by creating artifacts that express personal and shared desires not being satisfied by mainstream media. For example, female fans of *Star Trek* currently explore the subtext of the relationship between Captain Kirk and Mr. Spock in music videos made from their favorite songs and parts of *Star Trek* episodes. Other examples hint at a cultural landscape of ubiquitous participatory video: *America's Funniest Home Videos*, the Rodney King videotape, and video karaoke bars. These are just foretastes of the much more fundamental cultural and technological shift the growth of Garage Cinema will bring about.

Computational semasiographic technology will alter the relations of production and consumption informing the writing, reading, and distribution of motion pictures. Just as 500 years ago very few people could write or publish written texts, today very few people can make and distribute movies or television programs. Over the next 50 years, technologies enabling the popularization and personalization of motion picture production and distribution will bring about far-reaching cultural changes just as the invention of the printing press did centuries ago. As with the invention of writing (or even language) itself, the deep integration of motion pictures and computation will bring about entirely new possibilities for human expression and communication.

The Future of Garage Cinema. In order for Garage Cinema to be a common daily practice, many technological, social, and legal changes have to occur, but the two major technological challenges that have to be met are the development of tools for accessing content and tools for manipulating content. The main difference between a word processor and a Garage Cinema machine is that with language, if I want to tell you a story about a summer day in Paris in which a little dog stole my hat, I just did (or have begun to). With motion pictures I cannot simply speak or write images as I can sentences. In order to make my movie I have three options:

- I can take my production team to Paris or a studio and with several hundred thousand dollars shoot and edit this story.
- I can wait 20–30 years for photorealistic computer graphics to become real-time and affordable, but I will still miss the visual and philosophical feel of working with recorded video.
- I can access stock footage of Paris, dogs, hats, even appropriate footage of actors and/or myself, and piece together my movie.

For Garage Cinema makers, the first challenge is getting access to video content in order to be able to tell a wider range of stories than they can shoot or synthesize. The second challenge lies in having tools enabling Garage Cinema makers to manipulate video according to its content rather than requiring the specialized skills needed in current motion picture and video production.

In the future, we can imagine a world in which digital media are produced anywhere by anyone and are accessible to anyone anywhere. Media will accrete layers of annotations describing their contents as they move around the globe throughout their life cycles of use and reuse. In the future, annotation, both automatic and semi-automatic, will be fully integrated into the production, archiving, retrieval, and reuse of video and audio data. There will remain many annotations that computers won't be able to automatically encode.

A central challenge for computational media technology is to develop a language of description that both humans and computers can read and write, which will enable the integrated description and creation of video data. In order to overcome the inherent limitations of current keyword-based description and retrieval systems, we need to develop representations capturing the temporal, semantic, and relational content of video data. These representations also need to be convergent and scaleable to a global media archive. We have developed a language for the representation of video content called Media Streams, which addresses these issues [2, 3].

By having a structured representation of video content—meaningful bits about the bits—future annotation, retrieval, and composition technology will enable Garage Cinema makers to access already recorded video and to manipulate video streams according to their contents. With this kind of technology we will have tools that enable users to operate on higher-level content structures as opposed to being stuck with just bits, pixels, or even frames or clips.

Current video editing technology is like word processing with bitmaps. Today there isn't even an ASCII for video data, let alone spell checkers and grammar checkers, and other tools for structuring data according to its content. Editing tools that make use of content representations will complete the answer to the

two major needs of Garage Cinema makers: tools for accessing content and tools for manipulating content.

Media Producers and Users. In the next 50 years, the emergence of a global media archive and tools to access and manipulate this archive according to its contents will enable fundamental changes in the relationships between producers and users of digital media.

The most profound changes will occur at the traditionally "lower end" of video production. Changes in technology will bring about a merging of independent video producers and home video makers into a broad and active market sector. Today people speak of the "New Hollywood" and refer to the merger of Hollywood and Silicon Valley. When the tools and infrastructure are in place to enable cheap and effective home use of video annotation, retrieval, and repurposing tools, the garages of the world will be the sites of the "New New Hollywood" creating hundreds of millions of channels of video content. The conditions of production and use will have changed such that a large group of amateurs and home users will be regularly making video that can compete in the information marketplace of networked computers. The television networks will be supplanted by a situation in which the "Net works." As the PC revolution of the 1980s brought the text and numerical processing power once held by corporations to people's desktops, in the next decades the production and distribution power of Hollywood studios, television networks, and stock footage houses will reside on people's desks and in their garages.

Conclusion: Toward Garage Cinema

My view of the future of media technology is inspired by my desire to get my hands inside that television set I loved, worshipped, and grew up with as a child. If I were growing up in the next century, I would want to make Garage Cinema from video I downloaded from the Internet, taped from television, and recorded with my camcorder. Maybe you do too. It is my contention that you, or your children, will.

Just as we often find it hard to imagine our own civilization before the advent of widespread literacy in the 17th and 18th centuries, in the next century our descendants will find it hard to understand that while everyone watched movies, videos, and TV, so few had the tools to make them. The vision of Garage Cinema attempts to convey the radical changes in practices of production, distribution and use and in our possibilities of language and communication that video representation technology, like Media Streams, will make possible. It may be hard to conceptualize a world in which you engage in a daily practice of making movies from parts of existing ones to communicate and play with others, but your grandchildren will not understand how you ever lived without it. Watching what people are already doing with the primitive tools of camcorders and computers today is inspiring. Imagining what they would be able to do with computational semasiographic video technology initiates the journey toward the other side of a paradigm shift in media technology and human communication that we are about to begin. ◻

Acknowledgments

I would like to thank Brian Williams and Golan Levin for their amazing work in building Media Streams at the MIT Media Lab with me. To them and to David Levitt, Dick Shoup, and Gordon Kotik my thanks for helping me continue this work and its successors at Interval. Also thanks to Golan Levin and Andrew Singer for valuable editorial suggestions.

References

1. Bliss, C.K. *Semantography-Blissymbolics*. 3d Ed., Australia: Semantography-Blissymbolics Publications, Sydney, N.S.W., Australia, 1978.
2. Davis, M. Media Streams: Representing video for retrieval and repurposing. Ph.D. dissertation, Massachusetts Institute of Technology, 1995.
3. Davis, M. Media Streams: An iconic visual language for video representation. In *Readings in Human-Computer Interaction: Toward the Year 2000*, R.M. Baecker, J. Grudin, W.A.S. Buxton, and S. Greenberg, Eds., 854–66. 2nd ed., San Francisco: Morgan Kaufmann Publishers, Inc., 1995. (Also available at http://www.interval.com/papers/mediastreams/).
4. Jenkins, H. Textual poachers: Television fans & participatory culture. *Studies in Culture and Communication*, J. Fiske, Ed. Routledge, New York, 1992.
5. Mallery, G. *Picture-Writing of the American Indians. Vols. 1 and 2*. Dover Publications, New York, 1972.
6. McLuhan, M. *Understanding Media: The Extensions of Man.* McGraw-Hill, NY, 1965.
7. Neurath, O. *International Picture Language.* State Mutual Book and Periodical Service, New York, 1981.
8. Ong, W.J. *Orality and Literacy: The Technologizing of the Word.* T. Hawkes Ed., New Accents, Methuen, London, 1982.
9. Sampson, G. *Writing Systems: A Linguistic Introduction.* Stanford, California: Stanford University Press, 1985.
10. Yates, F.A. *The Art of Memory.* University of Chicago Press, Chicago, 1966.

Marc Davis (davis@interval.com) is a member of the Research Staff at Interval Research Corporation in Palo Alto, CA.