

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TikTok Inc.
Petitioner

v.

NTECH Properties, Inc.
Patent Owner

IPR2024-01338
U.S. Patent No. 8,875,185

“Method and Apparatus for Efficient, Entertaining Information Delivery”

DECLARATION OF KEVIN ALMEROOTH, PH.D. IN SUPPORT OF
PETITION FOR *INTER PARTES* REVIEW OF
U.S. PATENT NO. 8,875,185

TABLE OF CONTENTS

I.	Introduction.....	1
A.	Qualifications	1
1.	Educational Background.....	1
2.	Career	2
3.	Other Relevant Qualifications.....	10
B.	Previous Expert Witness Testimony	11
C.	Preparation for this Declaration	11
II.	Legal Understandings	12
A.	Claim Construction	13
B.	Anticipation	13
C.	Obviousness.....	14
III.	The '185 Patent.....	15
A.	Specification.....	15
B.	Prosecution History	18
1.	The '597 Provisional (Ex. 1003).....	19
2.	The '819 Application (Ex. 1004).....	19
3.	The '709 Application (Ex. 1005).....	30
4.	The '110 Application (Ex. 1006).....	49
5.	The '873 Application (Ex. 1007).....	50

6.	The '374 Application (Ex. 1008).....	66
C.	Level of Ordinary Skill in the Art.....	81
D.	Claim Construction	83
E.	Challenged Claims	83
IV.	The Prior Art.....	87
A.	State of the Art	87
B.	Marcus156 (Ex. 1009).....	90
C.	Rosen (Ex. 1011).....	91
D.	Sull (Ex. 1010)	92
E.	Hosken (Ex. 1015).....	93
F.	Byers (Ex. 1018).....	94
G.	Ferman (Ex. 1012).....	94
V.	Ground 1: Marcus156 Renders Claims 1-20 Obvious	95
A.	Claim 1	95
1.	Element 1[pre]: “1. A method of providing video programming to a user, comprising the steps of:”.....	95
2.	Element 1[a]: “searching, by one or more computing devices, one or more databases of video clips in response to a user request wherein each video clip comprises a plurality of frames,”	96
3.	Element 1[b]: “selecting, by one or more of said computing devices, a plurality of first video clips responsive to the user request, wherein each first video clip is associated with media content comprising one or more second video clips;”	103

4. Element 1[c]: “selecting, by one or more of said computing devices, a plurality of first transitional clips for association with at least some of the selected first video clips;” 106
5. Element 1[d]: “determining a first temporal sequence for concatenating the selected first video clips and first transitional clips;” 111
6. Element 1[e]: “wherein said selecting, determining and concatenating is performed according to one or more first templates, to define first video programming, wherein said one or more first templates comprise a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and” ..114
7. Element 1[f]: “presenting said first video programming in a first area of a display associated with one of said computing devices;” 126
8. Element 1[g]: “concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display;” 129
9. Element 1[h]: “receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;” 135
10. Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection;” 138
11. Element 1[j]: “selecting, by one or more of said computing devices, a plurality of video clips including the second video clip, and determining a second temporal sequence for concatenating the selected plurality of second video clips, wherein said selecting, determining, and concatenating is

	performed in accordance with one or more first or second templates, to define second video programming;”	142
12.	Element 1[k]: “wherein said one or more first or second templates comprises a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and”	142
13.	Element 1[l]: “presenting said second video programming on a display associated with one or more of said computing devices.”	143
B.	Claim 2: “A method according to claim 1, wherein at least one of said video clips of said first or second video programming is a transitional clip.”	144
C.	Claim 3: “A method according to claim 2, wherein at least one of said video clips of each of said first and second video programming is a transitional clip.”	145
D.	Claim 4: “A method according to claim 3, wherein said transitional clip is a hosting clip consistent with at least one template variable.”	145
E.	Claim 5: “A method according to claim 2, wherein at least one transitional clip is a hosting clip consistent with at least one template variable.”	149
F.	Claim 6: “A method according to claim 2, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.”	149
G.	Claim 7: “A method according to claim 6, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.”	150

- H. Claim 8: “A method according to claim 6, wherein one or more video clips comprise advertising.” 151
- I. Claim 9: “A method according to claim 6, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.” 152
- J. Claim 10: “A method according to claim 6, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.” 153
- K. Claim 11: “A method according to claim 6, wherein said stored associations between said video clips comprise a plurality of media tags.” 154
- L. Claim 12: “A method according to claim 6, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.” 155
- M. Claim 13: “A method according to claim 1, wherein said first and second video programming are each presented as a stream of information to the user.” 156
- N. Claim 14: “A method according to claim 1, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.” 158
- O. Claim 15: “A method according to claim 14, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.” 158
- P. Claim 16: “A method according to claim 1, wherein one or more video clips comprise advertising.” 158

Q.	Claim 17: “A method according to claim 1, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.”	158
R.	Claim 18: “A method according to claim 1, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.”	158
S.	Claim 19: “A method according to claim 1, wherein said stored associations between said video clips comprise a plurality of media tags.”	159
T.	Claim 20: “A method according to claim 1, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.”	159
VI.	Ground 2: Marcus156 and Rosen Render Claims 1-20 Obvious	159
A.	Specific Proposed Modification to Marcus156’s System	159
B.	Motivation to Combine	160
C.	Reasonable Expectation of Success	162
D.	All Claims and Claim Elements Other Than 1[g], 1[h], and 1[i]	165
E.	Rosen Discloses Element 1[g]: “concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display;”	165
F.	Rosen Discloses Element 1[h]: “receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;”	167
G.	Rosen Discloses Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored	

	association between the second video clip and at least one first video clip indicated by said user selection;”	168
VII.	Ground 3: Marcus156 and Sull Render Claims 1-20 Obvious.....	169
A.	Specific Proposed Modification to Marcus156’s System	169
B.	Motivation to Combine	169
C.	Reasonable Expectation of Success	170
D.	All Claims and Claim Elements Other Than 1[g], 1[h], and 1[i]	171
E.	Sull Discloses Element 1[g]: “concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display;”	172
F.	Sull Discloses Element 1[h]: “receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;”	173
G.	Sull Discloses Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection;”	174
VIII.	Ground 4: Marcus156, Rosen, and Hosken Render Claims 1-20 Obvious.	175
A.	Specific Proposed Modification to the Combined System of Marcus156 and Rosen	175
B.	Motivation to Combine	176
C.	Reasonable Expectation of Success	177
D.	All Claims and Claim Elements Other Than 1[i].....	178
E.	Hosken Discloses Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored	

	association between the second video clip and at least one first video clip indicated by said user selection;”	178
IX.	Ground 5: Marcus156, Sull, and Hosken Render Claims 1-20 Obvious	180
X.	Ground 6: Marcus156, Rosen, Hosken, and Byers Render Claims 8 and 16 Obvious.....	180
	A. Specific Proposed Modification to the Combined System of Marcus156, Rosen, and Hosken Using Byers.....	180
	B. Motivation to Combine	181
	C. Reasonable Expectation of Success	182
	D. Byers Discloses and Renders Obvious “One or More Video Clips Comprise Advertising”.....	183
XI.	Ground 7: Marcus156, Sull, Hosken, and Byers Render Claims 8 and 16 Obvious.....	184
XII.	Ground 8: Ferman Renders Claims 1-7, 9-15, and 17-20 Obvious.....	184
	A. Claim 1	185
	1. Element 1[pre]: “1. A method of providing video programming to a user, comprising the steps of:”	185
	2. Element 1[a]: “searching, by one or more computing devices, one or more databases of video clips in response to a user request wherein each video clip comprises a plurality of frames,”	185
	3. Element 1[b]: “selecting, by one or more of said computing devices, a plurality of first video clips responsive to the user request, wherein each first video clip is associated with media content comprising one or more second video clips;”	190
	4. Element 1[c]: “selecting, by one or more of said computing devices, a plurality of first transitional clips for association with at least some of the selected first video clips;”	193

5. Element 1[d]: “determining a first temporal sequence for concatenating the selected first video clips and first transitional clips;” 197
6. Element 1[e]: “wherein said selecting, determining and concatenating is performed according to one or more first templates, to define first video programming, wherein said one or more first templates comprise a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and” ..199
7. Element 1[f]: “presenting said first video programming in a first area of a display associated with one of said computing devices;” 208
8. Element 1[g]: “concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display;” 211
9. Element 1[h]: “receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;” 211
10. Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection;” 212
11. Element 1[j]: “selecting, by one or more of said computing devices, a plurality of video clips including the second video clip, and determining a second temporal sequence for concatenating the selected plurality of second video clips, wherein said selecting, determining, and concatenating is performed in accordance with one or more first or second templates, to define second video programming;” 214
12. Element 1[k]: “wherein said one or more first or second templates comprises a plurality of variables corresponding to

	(a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and”	215
13.	Element 1[l]: “presenting said second video programming on a display associated with one or more of said computing devices.”	216
B.	Claim 2: “A method according to claim 1, wherein at least one of said video clips of said first or second video programming is a transitional clip.”	216
C.	Claim 3: “A method according to claim 2, wherein at least one of said video clips of each of said first and second video programming is a transitional clip.”	217
D.	Claim 4: “A method according to claim 3, wherein said transitional clip is a hosting clip consistent with at least one template variable.”	217
E.	Claim 5: “A method according to claim 2, wherein at least one transitional clip is a hosting clip consistent with at least one template variable.”	219
F.	Claim 6: “A method according to claim 2, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.”	219
G.	Claim 7: “A method according to claim 6, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.”	222
H.	Claim 9: “A method according to claim 6, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.”	224
I.	Claim 10: “A method according to claim 6, wherein the step of selecting a second video clip based on said association with at least	

	one first video clip is performed according to at least one of the first or second templates.”	227
J.	Claim 11: “A method according to claim 6, wherein said stored associations between said video clips comprise a plurality of media tags.”	228
K.	Claim 12: “A method according to claim 6, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.”	229
L.	Claim 13: “A method according to claim 1, wherein said first and second video programming are each presented as a stream of information to the user.”	234
M.	Claim 14: “A method according to claim 1, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.”	235
N.	Claim 15: “A method according to claim 14, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.”	235
O.	Claim 17: “A method according to claim 1, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.”	235
P.	Claim 18: “A method according to claim 1, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.”	236
Q.	Claim 19: “A method according to claim 1, wherein said stored associations between said video clips comprise a plurality of media tags.”	236

R.	Claim 20: “A method according to claim 1, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.”	236
XIII.	Ground 9: Ferman and Byers Render Claims 8 and 16 Obvious	236
A.	Specific Proposed Modification to Ferman.....	236
B.	Motivation to Combine	237
C.	Reasonable Expectation of Success	238
D.	Claims 8 and 16: Ferman and Byers Render Obvious “One or More Video Clips Comprise Advertising”	239
XIV.	No Secondary Considerations of Non-Obviousness	240
XV.	Conclusion	241

LIST OF EXHIBITS

Exhibit	Description
Ex. 1001	U.S. Patent No. 8,875,185 (“the ’185 patent”)
Ex. 1002	Declaration of Dr. Kevin Almeroth (“Almeroth”)
Ex. 1003	U.S. Provisional Patent Application No. 60/482,597 (“the ’597 provisional”)
Ex. 1004	File History of U.S. Patent Application No. 10/874,819 (now abandoned) (“the ’819 application”)
Ex. 1005	File History of U.S. Patent No. 8,875,185 (U.S. Patent Application No. 13/195,709) (certain prior art omitted) (“the ’709 application”)
Ex. 1006	File History of U.S. Patent No. 9,317,597 (U.S. Patent Application No. 14/494,110) (“the ’110 application”)
Ex. 1007	File History of U.S. Patent No. 11,184,686 (U.S. Patent Application No. 15/131,873) (“the ’873 application”)
Ex. 1008	File History of U.S. Patent Application No. 17/529,374 (currently pending) (“the ’374 application”)
Ex. 1009	U.S. Patent No. 6,032,156 (“Marcus156”)
Ex. 1010	U.S. Patent Publication No. 2002/0069218 (“Sull”)
Ex. 1011	U.S. Patent Publication No. 2004/0201610 (“Rosen”)
Ex. 1012	U.S. Patent Publication No. 2002/0059584 (“Ferman”)
Ex. 1013	International Patent Application Publication No. WO 01/06398 (“Hosken”)
Ex. 1014	U.S. Patent No. 7,334,249 (“Byers”)
Ex. 1015	List of Recent Expert Witness Engagements of Dr. Kevin Almeroth
Ex. 1016	Curriculum Vitae of Dr. Kevin Almeroth

IPR of USP 8,875,185
Decl. of Dr. Kevin Almeroth

Exhibit	Description
Ex. 1017	Complaint for Patent Infringement, <i>NTECH Properties, Inc. v. ByteDance Ltd. et al</i> , Case No. 2:24-cv-130 (E.D. Tex. Feb. 23, 2024), ECF No. 1
Ex. 1018	Amended Complaint for Patent Infringement, <i>NTECH Properties, Inc. v. ByteDance Ltd. et al</i> , Case No. 2:24-cv-130 (E.D. Tex. May 24, 2024), ECF No. 6

I, Dr. Kevin Almeroth, declare as follows:

I. Introduction

1. I have been retained by Quinn Emanuel Urquhart & Sullivan, LLP on behalf of the Petitioner TikTok Inc. (“Petitioner”) as an independent expert in this *inter partes* review (this “Proceeding”) before the Patent Trial and Appeal Board of the United States Patent and Trademark Office (the “Board”) to review claims 1-20 (“the challenged claims”) of U.S. Patent No. 8,875,185 (“the ’185 patent”). I have been asked by the Petitioner to assist in evaluating the claims and the disclosure of the ’185 patent.

A. Qualifications

2. Petitioner’s Exhibit 1016 is a true and correct copy of my current CV, which describes my education, patents and publications, employment and research history, and professional activities and awards.

1. Educational Background

3. I hold three degrees from the Georgia Institute of Technology: (1) a Bachelor of Science degree in Information and Computer Science (with minors in Economics, Technical Communication, and American Literature) earned in June 1992; (2) a Master of Science degree in Computer Science (with specialization in Networking and Systems) earned in June 1994; and (3) a Doctor of Philosophy (Ph.D.) degree in Computer Science (Dissertation Title: Networking and System

Support for the Efficient, Scalable Delivery of Services in Interactive Multimedia System, minor in Telecommunications Public Policy) earned in June 1997. During my education, I have taken a wide variety of courses as demonstrated by my minor. My undergraduate degree also included a number of courses more typical of a degree in electrical engineering including digital logic, signal processing, and telecommunications theory.

2. Career

4. I am currently a Professor Emeritus in the Department of Computer Science at the University of California, Santa Barbara (UCSB). While active at UCSB, I held faculty appointments and was a founding member of the Computer Engineering (CE) Program, Media Arts and Technology (MAT) Program, and the Technology Management Program (TMP). I also served as the Associate Director of the Center for Information Technology and Society (CITS) from 1999 to 2012. I have been a faculty member at UCSB since July 1997.

5. One of the major concentrations of my research has been the delivery of multimedia content and data between computing devices, including various network architectures. In my research, I have studied large-scale content delivery systems, and the use of servers located in a variety of geographic locations to provide scalable delivery to hundreds or thousands of users simultaneously. I have also studied smaller-scale content delivery systems in which content is exchanged

between individual computers and portable devices. My work has emphasized the exchange of content more efficiently across computer networks, including the scalable delivery of content to many users, mobile computing, satellite networking, delivering content to mobile devices, and network support for data delivery in wireless networks.

6. In 1992, the initial focus of my research was on the provision of interactive functions (e.g., VCR-style functions like pause, rewind, and fast-forward) for near video-on-demand systems in cable systems; in particular, how to aggregate requests for movies at a cable head-end and then how to satisfy a multitude of requests using one audio/video stream broadcast to multiple receivers simultaneously. This research has continually evolved and resulted in the development of techniques to scalably deliver on-demand content, including audio, video, web documents, and other types of data, through the Internet and over other types of networks, including over cable systems, broadband telephone lines, and satellite links.

7. An important component of my research has been investigating the challenges of communicating multimedia content, including video, between computers and across networks including the Internet. Although the early Internet was used mostly for text-based, non-real time applications, the interest in sharing multimedia content, such as video, quickly developed. Multimedia-based

applications ranged from downloading content to a device to streaming multimedia content to be instantly used. One of the challenges was that multimedia content is typically larger than text-only content, but there are also opportunities to use different delivery techniques since multimedia content is more resilient to errors. I have worked on a variety of research problems and used a number of systems that were developed to deliver multimedia content to users. One content-delivery method I have researched is the one-to-many communication facility called “multicast,” first deployed as the Multicast Backbone, a virtual overlay network supporting one-to-many communication. Multicast is one technique that can be used on the Internet to provide streaming media support for complex applications like video-on-demand, distance learning, distributed collaboration, distributed games, and large-scale wireless communication. The delivery of media through multicast often involves using Internet infrastructure, devices and protocols, including protocols for routing and TCP/IP.

8. Starting in 1997, I worked on a project to integrate the streaming media capabilities of the Internet together with the interactivity of the web. I developed a project called the Interactive Multimedia Jukebox (IMJ). Users would visit a web page and select content to view. The content would then be scheduled on one of a number of channels, including delivery to students in Georgia Tech dorms delivered

via the campus cable plant. The content of each channel was delivered using multicast communication.

9. In the IMJ, the number of channels varied depending on the capabilities of the server including the available bandwidth of its connection to the Internet. If one of the channels was idle, the requesting user would be able to watch their selection immediately. If all channels were streaming previously selected content, the user's selection would be queued on the channel with the shortest wait time. In the meantime, the user would see what content was currently playing on other channels, and because of the use of multicast, would be able to join one of the existing channels and watch the content at the point it was currently being transmitted.

10. The IMJ service combined the interactivity of the web with the streaming capabilities of the Internet to create a jukebox-like service. It supported true Video-on-Demand when capacity allowed, but scaled to any number of users based on queuing requested programs. As part of the project, we obtained permission from Turner Broadcasting to transmit cartoons and other short-subject content. We also connected the IMJ into the Georgia Tech campus cable television network so that students in their dorms could use the web to request content and then view that content on one of the campus's public access channels.

11. More recently, I have also studied issues concerning how users choose content, especially when considering the price of that content. My research has examined how dynamic content pricing can be used to control system load. By raising prices when systems start to become overloaded (*i.e.*, when all available resources are fully utilized) and reducing prices when system capacity is readily available, users' capacity to pay as well as their willingness can be used as factors in stabilizing the response time of a system. This capability is particularly useful in systems where content is downloaded or streamed on-demand to users.

12. Protecting networks, including their operation and content, has been an underlying theme of my research almost since the beginning of my research career. Starting in 2000, I have been involved in several projects that specifically address security, network protection, and firewalls. After significant background work, a team on which I was a member successfully submitted a \$4.3M grant proposal to the Army Research Office (ARO) at the Department of Defense to propose and develop a high-speed intrusion detection system. Key aspects of the system included associating streams of packets and analyzing them for viruses and other malware. Once the grant was awarded, we spent several years developing and meeting the milestones of the project. A number of my students worked on related projects and published papers on topics ranging from intrusion detection to developing advanced techniques to be incorporated into firewalls. I have also used firewalls, including

their associated malware detection features, in developing techniques for the classroom to ensure that students are not distracted by online content.

13. Recent work ties some of the various threads of my past research together. I have investigated content delivery in online social networks and proposed reputation management systems in large-scale social networks and marketplaces. On the content delivery side, I have looked at issues of caching and cache placement, especially when content being shared and the cache has geographical relevance. We were able to show that effective caching strategies can greatly improve performance and reduce deployment costs. Our work on reputation systems showed that reputations have economic value, and as such, creates a motivation to manipulate reputations. In response, we developed a variety of solutions to protect the integrity of reputations in online social networks. The techniques we developed for content delivery and reputation management were particularly relevant in peer-to-peer communication.

14. My involvement in the research community extends to leadership positions for several academic journals and conferences. I am the co-chair of the Steering Committee for the ACM Network and System Support for Digital Audio and Video (NOSSDAV) workshop and on the Steering Committees for the International Conference on Network Protocols (ICNP), ACM Sigcomm Workshop on Challenged Networks (CHANTS), and IEEE Global Internet (GI) Symposium. I

IPR of USP 8,875,185

Decl. of Dr. Kevin Almeroth

have served or am serving on the Editorial Boards of IEEE/ACM Transactions on Networking, IEEE Transactions on Mobile Computing, IEEE Network, ACM Computers in Entertainment, AACE Journal of Interactive Learning Research (JILR), and ACM Computer Communications Review. I have co-chaired a number of conferences and workshops including the IEEE International Conference on Network Protocols (ICNP), IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), International Conference on Communication Systems and Networks (COMSNETS), IFIP/IEEE International Conference on Management of Multimedia Networks and Services (MMNS), the International Workshop On Wireless Network Measurement (WiNMee), ACM Sigcomm Workshop on Challenged Networks (CHANTS), the Network Group Communication (NGC) workshop, and the Global Internet Symposium, and I have served on the program committees for numerous conferences.

15. Furthermore, in the courses I taught at UCSB, a significant portion of my curriculum covered aspects of the Internet and network communication including the physical and data link layers of the Open System Interconnect (OSI) protocol stack, and standardized protocols for communicating across a variety of physical media such as cable systems, telephone lines, wireless, and high-speed Local Area Networks (LANs). The courses I have taught also cover most major topics in Internet communication, including data communication, multimedia

encoding, and mobile application design. My research and courses have covered a range of physical infrastructures for delivering content over networks, including cable, Integrated Services Digital Network (ISDN), Ethernet, Asynchronous Transfer Mode (ATM), fiber, and Digital Subscriber Line (DSL). For a complete list of courses I have taught, see my curriculum vitae (CV).

16. In addition, I co-founded a technology company called Santa Barbara Labs that was working under a sub-contract from the U.S. Air Force to develop very accurate emulation systems for the military's next generation internetwork. Santa Barbara Labs' focus was in developing an emulation platform to test the performance characteristics of the network architecture in the variety of environments in which it was expected to operate, and, in particular, for network services including IPv6, multicast, Quality of Service (QoS), satellite-based communication, and security. Applications for this emulation program included communication of a variety of multimedia-based services, including video conferencing and video-on-demand.

17. In addition to having co-founded a technology company myself, I have worked for, consulted with, and collaborated with companies for nearly 30 years. These companies range from well-established companies to start-ups and include IBM, Hitachi Telecom, Turner Broadcasting System (TBS), Bell South, Digital Fountain, RealNetworks, Intel Research, Cisco Systems, and Lockheed Martin.

18. Additional details about my employment history, fields of expertise, and publications are further included in my CV attached as Exhibit 1017 to this Declaration.

3. Other Relevant Qualifications

19. I am a Member of the Association of Computing Machinery (ACM) and a Fellow of the Institute of Electrical and Electronics Engineers (IEEE).

20. As an important component of my research program, I have been involved in the development of academic research into available technology in the market place. One aspect of this work is my involvement in the Internet Engineering Task Force (IETF). The IETF is a large and open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet. I have been involved in various IETF groups including many content delivery-related working groups like the Audio Video Transport (AVT) group, the MBone Deployment (MBONED) group, Source Specific Multicast (SSM) group, the Inter-Domain Multicast Routing (IDMR) group, the Reliable Multicast Transport (RMT) group, the Protocol Independent Multicast (PIM) group, etc. I have also served as a member of the Multicast Directorate (MADDOGS), which oversaw the standardization of all things related to multicast in the IETF. Finally, I was the Chair of the Internet2 Multicast Working Group for seven years.

21. I am an author or co-author of approximately 200 technical papers, published software systems, IETF Internet Drafts and IETF Request for Comments (RFCs). A complete list of my publications is in my CV, attached to this report as Exhibit 1017.

22. I have been awarded numerous teaching awards, including Computer Science Outstanding Faculty Member (1997-98, 1998-99, 1999-2000, 2004-06, UCSB Spotlight on Excellence Award (2000-01), and UCSB Academic Senate Distinguished Teaching Award (2006-07).

B. Previous Expert Witness Testimony

23. The list of recent matters in which I have testified can be found attached as Exhibit 1015.

C. Preparation for this Declaration

24. In forming my opinions, I have considered the '185 patent specification, including the Abstract, the figures, and the claim language itself, as would have been understood by a person having ordinary skill in the art. I have also reviewed the file history of the '185 patent, the Exhibits that are listed in the list of Exhibits within the Petition, and any other material cited in this declaration.

25. In forming my opinions, I have relied on my personal knowledge and professional experience, and on the documents and information referenced in this declaration.

26. This declaration explains, based on facts and information available to me to date, the subject matter and opinions related to this Proceeding. As such, I am prepared to provide expert testimony regarding opinions formed resulting from my analysis of the issues considered in this declaration if asked about those issues by the Board or by the private parties' attorneys.

27. Additionally, I may discuss my own work, teachings, and knowledge of the state of the art in the relevant time period. I may rely on handbooks, textbooks, technical literature, and the like to demonstrate the state of the art in the relevant period and the evolution of relevant technologies.

28. Throughout this declaration, I refer to specific pages of the '185 patent and other documents. The citations are intended to be exemplary and are not intended to convey that the citations are the only source of evidence to support the propositions for which they are cited.

29. I am being compensated for my time spent on this matter at a rate of \$850 per hour, and my compensation is in no way contingent upon the outcome of this matter or on the opinions I offer. All of the opinions expressed in this declaration are my own.

II. Legal Understandings

30. In this section, I describe my understanding of certain legal standards that I have relied upon in forming my opinions set forth in this declaration. I have

been informed of these legal standards by Petitioner's attorneys. I am not an attorney and I have not thoroughly researched the law on patent invalidity. I am relying only on instructions from Petitioner's attorneys for these legal standards.

A. Claim Construction

31. I have been instructed by counsel that claim construction is a matter of law for the arbiter of law to decide. I understand that in an *inter partes* review, claims are construed using the same claim construction standard that would be used to construe the claim in a civil action.

32. I understand that a patent may include two types of claims, independent claims and dependent claims. An independent claim stands alone and includes only the limitations it recites. A dependent claim can depend on an independent claim or another dependent claim. I understand that a dependent claim includes all the limitations that it recites in addition to the limitations recited in the claim from which it depends.

B. Anticipation

33. I understand that a patent claim is anticipated when a single piece of prior art describes every element of the claimed invention, either expressly or inherently, arranged in the same way as in the claim. For inherent anticipation to be found, it is required that the missing descriptive material is necessarily present in the prior art. I understand that, for the purpose of an *inter partes* review, prior art that

anticipates a claim can include both patents and printed publications from anywhere in the world.

C. Obviousness

34. I understand that a patent claim is unpatentable and invalid if the subject matter of the claim as a whole would have been obvious to a person of ordinary skill in the art of the claimed subject matter as of the time of the invention at issue. I understand that the following factors must be evaluated to determine whether the claimed subject matter is obvious: (1) the scope and content of the prior art; (2) the difference or differences, if any, between each claim of the patent and the prior art; and (3) the level of ordinary skill in the art at the time the patent was filed. Unlike anticipation, which allows consideration of only one item of prior art, I understand that obviousness may be shown by considering more than one item of prior art. Moreover, I have been informed and I understand that the so-called objective indicia of non-obviousness, also known as “secondary considerations,” like the following are also to be considered when assessing obviousness: (1) commercial success; (2) long-felt but unresolved needs; (3) copying of the invention by others in the field; (4) initial expressions of disbelief by experts in the field; (5) failure of others to solve the problem that the inventor solved; and (6) unexpected results. I also understand that evidence of objective indicia of non-obviousness must be commensurate in scope with the claimed subject matter.

III. The '185 Patent

A. Specification

35. I understand that the '185 patent is a pre-AIA patent that claims priority to June 24, 2003 based on the filing of U.S. Provisional Patent Application No. 60/482,597 ("the '597 provisional"). It is titled "Method and Apparatus for Efficient, Entertaining Information Delivery" and lists a sole named inventor, Dwight Marcus. The '185 patent is a continuation of U.S. Patent Application No. 10/874,819 ("the '819 application"), now abandoned. The '185 patent was filed as U.S. Patent Application No. 13/195,709 ("the '709 application") on August 1, 2011 and issued on October 28, 2014.

36. I understand that, as of the date of this Declaration, three other subsequent continuation applications in the same family have been filed: U.S. Patent Application No. 14/494,110 ("the '110 application") which issued as U.S. Patent No. 9,317,597 ("the '597 patent"); U.S. Patent Application No. 15/131,873 ("the '873 application"), which issued as U.S. Patent No. 11,184,686 ("the '686 patent"); and U.S. Patent Application No. 17/529,374 ("the '374 application"), which I understand is currently pending.

37. The specification of the '185 patent is brief – three figures and slightly under four columns of text. In the "Background of the Invention" section, the '185 patent specification attempts to distinguish its claimed invention with three different

prior art video delivery systems: television, video on demand, and the world wide web. '185 patent at 1:26-2:2. According to the specification, television was inadequate because “a user has very little control over what information is presented at a specific time.” *Id.* at 1:47-48. According to the specification, video on demand was problematic because it employed user-unfriendly list-based user interfaces. *Id.* at 1:58-61 (“VOD systems typically do not provide users with an efficient or entertaining interface for requesting information. A user must search through and select from a list of available content.”). According to the specification, the world wide web offered more user control, but was too “inefficient and non-entertaining.” *Id.* at 2:1-2.

38. The '185 patent claims to address these deficiencies by customizing video programming based on additional metadata, which the patent uses various terms for, including “template,” “information about the user,” and “header.” *E.g.*, *id.* at 2:6-18; 2:67-3:8. A computer system uses an algorithm to select video “clips” and assemble them into a stream of video programming that consists of multiple clips. *Id.* at 2:19-21 (“[A] user is presented with a stream of information comprised of clips of information assembled for the user.”); 2:65 (“[I]nformation is partitioned into clips.”); 3:35-36 (“[T]he information is packaged by concatenating clips together to form a stream of data.”).

39. A user may “select a clip,” which may in turn trigger the selection and assembly of a new video stream based on the same metadata that was employed and the user selection. Id. at 2:26-29 (“In one embodiment, selection made by the user are monitored and used in association with templates to determine which clips will be selected and how they will be presented to the user in the stream of information in the future.”).

40. Figure 2 shows a particular user interface for viewing the video stream and making selections:

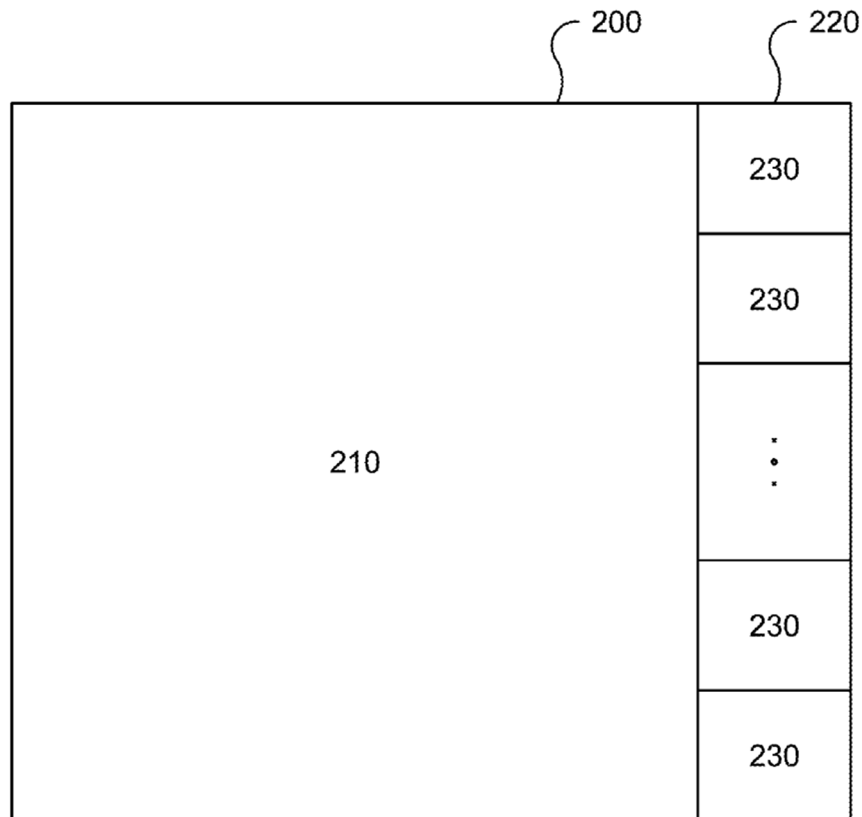


Fig. 2

Id. at Fig. 2; 3:47-48 (“FIG. 2 illustrates an information display in accordance with one embodiment of the present invention.”).

41. In Figure 2, 210 is the “current viewing area” and 220 is a “menu of options area” comprising multiple “mini-display regions” 230. A main video is played in current viewing area 210. After a video finishes playing in 210, it can repeat in 230. *Id.* at 3:58-61 (“In one embodiment, once a clip has displayed in the current viewing area 210, the clip continues to display in one of the mini-display regions 230.”). And rather than a running video in 230, “a representation of the clip” may be displayed instead. *Id.* at 3:63-64. The content of the mini-display regions 230 can change based on user selections. *Id.* at 4:10-19. The content of the current viewing area 210 can also be changed based on a user selection. *Id.* at 4:20-23 (“In one embodiment, the user selects information for display in the current viewing area 210 by selecting a clip either from a stream of clips delivered to the current viewing area 210 or to one of the mini-display regions 230.”).

B. Prosecution History

42. I have reviewed the prosecution histories of the applications in the family of the ’185 patent. Below, I provide a brief overview of some of the major events of the prosecution histories.

1. The '597 Provisional (Ex. 1003)

43. The '597 provisional was filed on June 24, 2003 and included a version of the specification that is similar to the issued '185 patent specification. The '597 provisional did not claim priority to any other patent application. The '597 provisional included draft claims that are different from the issued '185 patent claims. The '597 provisional included an appendix. Ex. 1003 at 23-125. The appendix is described as an attempt “to lay out the novelty of the model in easy-to-read language.” *Id.* at 23. The appendix describes a public use in February 2002 of at least some of the technologies described and claimed in the '185 patent: “The following categories of one embodiment were actually tested in the February 2002 MICRO Program pilot for the system assembled by us and UCSB.” *Id.* at 63.

2. The '819 Application (Ex. 1004)

44. The '819 application was filed on June 23, 2004 and contained a specification that was slightly larger than the '597 provisional. The '819 application claimed priority to the '597 provisional and not to any other patent application. Ex. 1004 at 3, 136, 169. The '819 application also included the same appendix that was included with the '597 provisional. *Id.* at 21-123.

45. The applicant filed information disclosure statements on November 3, 2004 and January 23, 2007, none of which listed any of the prior art raised in this petition.

46. In an office action on January 17, 2008, all pending claims 1-29 were rejected as anticipated by U.S. Patent Publication No. 2003/0163815 (“Begeja”) or obvious in light of Begeja and U.S. Patent Publication No. 2005/0198677 (“Lewis”). *Id.* at 150-69. On July 17, 2008, the applicant amended the claims. *Id.* at 188-203. Among other things, pending claim 1 was amended to add, following the existing claim language “presenting said stream of information to a user,” the additional language “in a first area of a display area; simultaneously with said step of presenting said stream of information, presenting to the user in a second area in the display area one or more frames from recently-displayed clips from the stream of information.” *Id.* at 189. The applicant argued that this addition traversed the examiner’s claim rejections:

The rejection of claim 1 is respectfully traversed on the grounds that Begeja does not teach or render obvious all of the limitations of claim 1 as amended.

Claim 1 has been amended, in substance, to recite that simultaneously with the presentation of the stream of information, frames from previously-presented clips are presented to the user. A first clip is determined wherein the first clip is associated with on [sic] of the previously-presented clips, and the previously-presented clip is selected by the user. The first clip is then delivered to the user.

Disclosure support for the amendment may be found, for example, in the specification, at Paragraphs [00020] to [00022], Fig. 2, and pages 30 and 31 of the Appendix.

Id. at 200.

47. In the next office action on October 24, 2008, the examiner again rejected all pending claims. *Id.* at 210-229. The claims that included the newly-added limitations on simultaneous display of a video and a frame were rejected as obvious in light of Begeja and U.S. Patent Publication No. 2002/0056098 (“White”). *Id.* at 220-221. The examiner noted that Begeja disclosed “simultaneously with said step of presenting said stream of information, presenting to the user in a second area in the display are one or more frames from clips from the stream of information (Paragraph [0066] fig 5; section 505)” but that “Begeja fails to explicitly disclose that the one or more frames are from recently-displayed clips.” *Id.* at 220. White supplied that material and would have been combined with Begeja for convenience and better user interaction:

Nevertheless, in a similar field of endeavor White discloses that the one or more frames are from recently-displayed clips (Paragraphs [0007] [0047] fig 8).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Begeja by specifically providing the elements mentioned above, as taught by White, for the purpose of providing convenience and a better interface interaction which allows users to navigate through the system in an easier and more efficient manner.

Id. at 221.

48. In response, the applicant submitted a request for continued examination on April 30, 2009 and again amended the claims. *Id.* at 241-61. The applicant amended claim 1 and others to add “arranging by the computer a plurality

of clips including the first clip into a second stream of information” and delivering the second stream to the user. *E.g., id.* at 242, 245 (adding “said selection unit being configured to select a plurality of clips including said first clip and said arranging unit being configured to arrange said plurality of clips including said first clip into a second stream of information”). Regarding the new amendments, the applicant stated that “[d]isclosure support may be found, for example, in Paragraph [00025] of the specification” and that “[n]either Begeja nor White teaches displaying one or more frames from recently-displayed clips from a stream of information arranged by a computer in a second area of a display area while presenting the stream of information in a first area of the display area.” *Id.* at 256. The applicant also stated:

The Examiner attempts to combine features from the embodiment of Figs. 4 and 5 of Begeja and the text at Paragraphs [0062] to [0068], with accessing an entire original video file as described in Paragraph [0028] of Begeja. In the embodiment of Figs. 4 and 5, the stream of information may be generated as indicated in Paragraph [0068] if the user does not choose a specific topic or clip. However, there is no indication that the embodiment of Figs. 4 and 5 provides a user with a “more” or “complete” button, as described in Paragraph [0028], for permitting a user to watch an entire original video file.

The Examiner also improperly states that White discloses that one or more frames are from recently-displayed clips. (Office Action, page 12, first paragraph). White discloses a TV recent screen 80 which includes small display screens containing images from other channels prior to being deselected (Paragraph [0049]). Thus, the images on TV recent screen 80 are not images from clips of a stream of information arranged by a computer, as recited in claim 1, but images from channels previously selected and deselected by a user.

In addition, the display of images on TV recent screen 80 of White serves a different purpose from the recited display of clips. White's display solves the problem of a user flipping between channels by entering the channel number or flipping through intermediate channels (Paragraph [0003]). Selection of an image on TV recent screen 80 does not provide an associated clip, but whatever material the selected channel happens to be delivering at that time. Indeed, White teaches updating the small display screens with live pictures or still pictures to reflect current content (Paragraph [0052]). By updating, White may remove any frame from content previously viewed by the user.

Moreover, claim 1 has been amended to recite that selection of the second clip provides a second stream of information including the associated first clip. Begeja teaches providing only a complete video, not a new stream of information. White merely provides an interface for easily re-selecting previously viewed channels (See Paragraph [0004]).

Id. at 257.

49. On May 8, 2009, shortly after submitting these amendments, the applicant also submitted an information disclosure statement. *Id.* at 263-68. Among other references, the information disclosure statement listed two patent publications by named inventor Dwight Marcus, U.S. Patent Publication Nos. 2002/0092019 and 2002/0077902. *Id.* at 265. On the same day, the applicant submitted another information disclosure statement. *Id.* at 270-76. This information disclosure statement listed Marcus156. *Id.* at 270.

50. All pending claims were again rejected in the next office action from May 14, 2009. *Id.* at 283-99. The examiner disagreed with the applicant about whether Begeja disclosed a “more” or “complete” button in the user interface. *Id.* at

285-86. The amended claims were rejected in light of the combination of Begeja and U.S. Patent No. 6,154,771 (“Rangan”). *Id.* at 287-89. Like the White reference, Rangan disclosed a simultaneous display of frames and clips:

However, it is noted that Begeja fails to explicitly disclose that the one or more frames are from recently-displayed clips.

Nevertheless, in a similar field of endeavor Rangan discloses that the one or more frames are from recently-displayed clips (Col. 26 lines 29-35 lines also exhibited on figure 2 and 4).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Begeja by specifically providing the elements mentioned above, as taught by Rangan, for the purpose of providing convenience and a better interface interaction which allows users to navigate through the system in an easier and more efficient manner.

Id. at 288; *see also id.* at 290 (similar).

51. The examiner also submitted both of the applicant’s information disclosure statements from May 8, 2009 and annotated them with “ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /J.M./.” *Id.* at 309-19. Marcus156 was listed. *Id.* at 314.

52. The applicant submitted another set of amendments and remarks on November 30, 2009. *Id.* at 320-37. However, the operative set of claims that included display of a frame and video and the creation of a second video stream were not amended. The applicant instead argued that the rejection in light of Begeja and Rangan was incorrect:

The rejection of claim 1 is respectfully traversed. Neither Begaja nor Rangan, even when combined, discloses a method including both arranging by the computer a plurality of clips into a first stream of information, including a second clip, and arranging by the computer a plurality of clips into a second stream of information that includes a first clip, wherein the first clip is associated with the second clip. In claim 1, two computer-arranged streams of information are presented to the user.

In contrast, in Begaja, pertinent segments are stitched together for continuous delivery to the user (Paragraph [0028]). The segments can be streamed as a single stream of information. The “more” or “complete” button of Paragraph [0028], noted by the Examiner in the Response to Arguments section of the Office Action, merely provides access to portions of an original video file that occurred immediately before or after the extracted segments. There is no disclosure in Begaja of “arranging by the computer a plurality of clips including the first clip into a second stream of information.” There is no arranging in Begaja in response to the “more” or “complete” button, but merely providing of a pre-existing video. Thus, Begaja provides, at most, only one computer-arranged stream of information.

Even if the Examiner's interpretation of Begaja were correct, and the embodiment of Paragraph [0028] could be combined with the page view 400 of Fig. 4 or the content retrieval page of Fig. 5, Begaja would not teach the recited second arranging step of claim 1.

Rangan does not cure this deficiency of Begaja. In Rangan, the thumbnail images 7 4 are of scenes that have been detected automatically, such as by a scene change detection module (col. 27, line 59 to col. 28, line 3), or manually (col. 26, lines 7-13). Video corresponding to thumbnail images is buffered on local storage, and the user clicking on a thumbnail causes a system to start recording from that scene (col. 26, lines 34-36). The selection of the thumbnail does not result in a step of “arranging by the computer a plurality of clips including the first clip into a second stream of information,” but merely recording recently viewed programming.

For at least the foregoing reasons, reconsideration and withdrawal of the rejection of claim 1 are respectfully solicited.

53. In the next office action from February 2, 2010, the examiner maintained the claim rejections. *Id.* at 344-62. The examiner refuted the applicant's arguments:

Applicant's arguments filed 11/30/2009 have been fully considered but they are not persuasive.

Regarding independent claims 1, 2, 10, 11, 19 and 20, applicant argues that Begeja fails to disclose "... arranging by the computer [said] plurality of clips into a first stream of information, including a second clip arranging by the computer a plurality of clips including a first clip into a second stream of information, wherein the first clip is associated with the second clip".

However, the examiner respectfully disagrees with the applicant. Begeja discloses a system for providing electronic clips or eClips from different content sources to the viewers, where the eClips may be stitched together in order to automatically provide a multi source clip experience to the user similar to the system portrayed on figure 5; see abstract and paragraph [0058]. In other words, Begeja discloses arranging a plurality of clips into a first stream, i.e. stream of grouped or stitched plurality of eClips, including a second clip, i.e. an eClip. Furthermore, Begeja clearly discloses a "more" or "complete" button option for permitting users to watch the entire original video file to which an eClip belongs, paragraph [0028]. Moreover, the cited paragraph recites the "more" or "complete" button can be included in the user interface. Hence, Begeja teaches arranging a plurality of clips into a second stream of information that includes a first clip, i.e. full original video, wherein the first clip is associated with the second clip, i.e. full original video is associated to user selected eClip. Therefore, Begeja clearly discloses "... arranging by the computer [said] plurality of clips into a first stream of information, including a second clip arranging by the computer a plurality of clips including a first clip into a second stream of information, wherein the first clip is associated with the second clip".

Id. at 345-46.

54. The applicant submitted amendments and remarks on August 17, 2010. *Id.* at 370-92. In claim 1, the applicant replaced “arranging” with “concatenating” and “stream of information” with “programming.” *Id.* at 371-72. In addition, the applicant introduced the terminology “template” into the claims, adding the language “selecting and concatenating [a plurality of clips], in accordance with a template.” *Id.* at 372. In the remarks, the applicant referenced a telephone interview conducted on August 17, 2010. The applicant wrote:

Telephone Interview

The Applicant thanks Examiners Koenig and Mendoza for the courtesy of a telephonic interview conducted with the undersigned on August 17, 2010.

In the interview, claim 1 and the Begeja reference were discussed. In particular, the undersigned noted that Begeja, in Paragraph [0028], teaches providing a preexisting program, while the claim 1 is directed to providing programming from selected and concatenated clips. Suggestions were made by the Examiners as to claim language that may assist in better distinguishing claim 1 over the prior art of record. No agreement was reached.

Rejection of Claims 1, 10 and 19 under 35 U.S.C. 103(a).

Claims 1, 10 and 19 stand rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Publication No. 2003/0163815 (Begeja) in view of U.S. Patent No. 6,154,771 (Rangan et al.).

Claims 1, 10 and 19 have been amended to recite that clips are selected and concatenated in accordance with a template to provide second programming. Specification support is found, for example, at Paragraph [00015]. In contrast, the “More” button of Begeja, noted in Paragraph [0028], merely results in the delivery of pre-existing programming, and does not involve the selection of clips or the use of a template. Accordingly, even if Begeja and Rangan could properly be

combined as proposed by the Examiner, the combination would not provide each and every limitation of claims 1, 10 and 19 as amended.

For at least the foregoing reasons, reconsideration and withdrawal of the rejections of claims 1, 10 and 19 is respectfully requested.

Id. at 383-84.

55. An interview summary filed on August 19, 2010 noted: “The applicant discussed the invention with respect to arranging a second stream as an original stream. The examiners suggested potential claim language to help overcome the Begeja reference of record, however no resolution was reached. The applicant will file a formal reply.” *Id.* at 400.

56. The next office action from January 20, 2011 maintained the rejection of all pending claims. *Id.* at 403-18. The claims were rejected in light of a new combination of U.S. Patent Publication No. 2002/0019858 (“Kaiser”) and U.S. Patent No. 6,317,784 (“Mackintosh”). *Id.* at 404-07. The examiner wrote:

Claims 1, 10 and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kaiser et al. (Pub No US 2002/0019858) in view of Mackintosh et al. (Patent No US 6,317,784). Hereinafter referenced as Kaiser and Mackintosh, respectively.

Regarding claim 1, Kaiser discloses a method of delivering information (see abstract) comprising:

selecting by a computer a plurality of clips (Paragraphs [0051] [0052] figure 7; personal station playlist search and mapping 715);

concatenating by the computer said plurality of clips into first programming (Paragraphs [0051] [0052] figure 7; personal station playlist building 720);

presenting by the computer said first programming (current playlist) to a user in a display area (Paragraph [0034] figure 1; computing devices 110);

presenting by a computer to the user in the display area recently-displayed clips from the first programming (Paragraphs [0054] [0055] figure 8; historical record of played songs 815);

determining by the computer a first clip (song in new playlist) wherein said first clip is associated with a second clip (song in current playlist) wherein said second clip is one of the presented recently-displayed clips (historical record of played songs 815) and wherein said second clip is selected (get more button) by said user (Paragraphs [0053] [0056] figures 7 and 8; user selects the "get more" button associated with a current song or a previous song from the historical record of played songs 815, in order to build a new playlist of clips 7 45 with the same characteristics as the selected song);

selecting and concatenating, in accordance with a template (media mapping), by the computer a plurality of clips including the first clip into second programming (Paragraphs [0053] [0052] [0056] figures 7 and 8; new playlist);

and delivering by the computer said second programming including said first clip to said user (Paragraphs [0053] [0052] [0056] figures 7 and 8; user experiences new playlist 745).

However, it is noted that Kaiser is silent to explicitly disclose presenting by a computer a first programming to a user in a first area of a display area; simultaneously with said step of presenting said first programming, presenting by a computer to a user in a second area in the display area one or more frames from recently-displayed clips from the first programming.

Nevertheless, in a similar field of endeavor Mackintosh also discloses a clip presenting system (figure 12) which allows the users to obtain more clips related to a selected clip (Col. 15 lines 13-36). Mackintosh further discloses presenting by a computer a first programming (current song in audio stream) to a user in a first area of a display area (Col. 22 line 61 - col. 23 line 6, figure 12; area component 526);

simultaneously with said step of presenting said first programming (current song in audio stream), presenting by a computer to a user in a second area (song history component 525) in the display area one or more frames (255a, 525b, 525c) from recently-displayed clips from the first programming (Col. 23 lines 7-25 figure 12; simultaneously displaying frames 525 of previously displayed songs).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser by specifically providing the elements mentioned above, as taught by Mackintosh, for the predictable result of allowing the users to easily and effectively select clips of content from a screen which presents current and previous clips simultaneously, without the need to force the user to navigate to other screens, allowing the interface to be more user friendly.

Regarding claims 10 and 19, Kaiser and Mackintosh disclose all the limitations of claims 10 and 19; therefore, claims 10 and 19 are rejected for the same reasons stated in claim 1.

Id. at 404-07.

57. Following this office action, the examiner filed a notice of abandonment on September 13, 2011, based on the applicant's failure to timely file a proper reply to the office action. *Id.* at 439-40.

3. The '709 Application (Ex. 1005)

58. The '709 application, which ultimately issued as the challenged '185 patent, was filed as a continuation of the '819 application on August 1, 2011. Ex. 1005 at 105, 116. The sole original claim 1 was similar to the ultimately-issued claim 1.

59. In the first office action from November 16, 2012, claim 1 was rejected as obvious over the combination of Kaiser and Mackintosh. *Id.* at 149-52. The

examiner stated that Mackintosh disclosed the simultaneous display of a video and a frame and that it would have been obvious to combine these teachings with Kaiser:

[I]n a similar field of endeavor Mackintosh also discloses a clip presenting system (figure 12) which allows the users to obtain more clips related to a selected clip (Col. 15 lines 13-36). Mackintosh further discloses presenting by the computer said first video programming (currently selected content) in a first area of a display area (Col. 22 line 61 – col. 23 line 6, figure 12; area component 526. Wherein the broadcast material may be radio content or video programs, col. 5 lines 12-15 and col. 6 lines 25-29);

Simultaneously with said presenting said first video programming (currently selected content), presenting by the computer in a second area (content history component 525) of the display one or more frames (images 525a, 525b, 525c) from recently-displayed video clips from the first video programming (Col. 23 lines 7-25 figure 12; simultaneously displaying frames 525 of previously displayed content);

and receiving by the computer a user selection of one of the displayed frames, the user-selected displayed frame being a frame of a video clip (Col. 3 lines 48-50 figure 12; user may select frames from history bar).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser by specifically providing the elements mentioned above, as taught by Mackintosh, for the predictable result of allowing the users to easily and effectively selects clips of content from a screen which presents current and previous clips simultaneously, without the need to force the user to navigate to other screens, allowing the interface to be more user friendly.

Id. at 151-52.

60. Additionally, the examiner also rejected claim 1 as obvious over the combination of Kaiser and Begeja. *Id.* at 152-56. The examiner stated that Begeja

disclosed the simultaneous display of a video and a frame and that it would have been obvious to combine these teachings with Kaiser:

[I]n a similar field of endeavor Begeja also discloses a clip presenting system (figure 5) which allows the users to obtain more clips related to a selected clip (Paragraph [0042]). Begeja further discloses presenting by the computer said first video programming (current clip 510) in a first area of a display area (Paragraph [0069] figure 5);

simultaneously with said presenting said first video programming (current clip 510), presenting by the computer in a second area of the display one or more frames (section 505 with thumbnails representing each clip) from recently-displayed video clips from the first video programming (Paragraphs [0068] [0069] figure 5; presenting in section 505 thumbnails representing each clip, wherein the user may skip to a previous clip displayed in section 505 with button 525);

and receiving by the computer a user selection of one of the displayed frames, the user-selected displayed frame being a frame of a video clip (Paragraph [0068] [0069] figure 5; user selections using buttons 525-550).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser by specifically providing the elements mentioned above, as taught by Begeja, for the predictable result of allowing the users to easily and effectively select clips of content from a screen which presents current and previous clips simultaneously, without the need to force the user to navigate to other screens, allowing the interface to be more user friendly and presenting media relevant to the interests of the user (Begeja – paragraph [0008]).

Id. at 153-54.

61. In the applicant's May 28, 2013 response to the office action, the applicant argued that "Kaiser fails to disclose a template, or selecting and determining a temporal sequence in accordance with a template, as recited in claim 1." *Id.* at 168. The applicant continued:

In the Office Action, at page 3, it is stated that Kaiser discloses the above-quoted recitation of claim 1 at Paragraphs [0051]-[0052], and Figure 7, and references new playlist building 745. The Office Action includes the phrase media mapping in parentheses after the term template. Applicant understands based [sic] on the providing of the phrase media mapping in parentheses that [sic] the Office is basing the rejection on an equating of the claim term template with the phrase media mapping and [sic] taking the position that the term media mapping is disclosed in Kaiser. However, the [sic] phrase media mapping never appears in Kaiser. For this reason alone, the Section 103(a) rejection of claim 1 should be withdrawn.

However, even if the phrase musical property mapping, which is used in Paragraphs [0051]-[0052] of Kaiser, and not media mapping, was intended by the Office, an equating of musical property mapping with template would [indecipherable]tually incorrect, as explained in detail below.

Paragraphs [0051]-[0052] of Kaiser state (emphasis added):

[0051] FIG. 7 illustrates an embodiment of the present invention whereby a media station, such as a radio station, is tailored to a user through a one-step personalized “Get More” station reprogram technique. For example, this technique could supplement the personal station listening experience described in connection with FIG. 6, or operate upon any playlist currently broadcast to a user based upon an underlying set of preferences, however specified. While listening to a particular song, a user can specify that the current playing song better corresponds to a higher affinity space/state at that time than the current station setting. The user requests to “get more” which instructs the system to find more songs like the current one. In effect, and partly in recognition of a user that doesn’t know exactly how to specify what he or she likes to a tee, but knows what he or she likes when the user hears it, the present embodiment allows a user to specify the properties of the current playing song, by selecting the current playing song, in order to hone the user's preferences. This then captures the musical property mapping of the currently playing song, automatically runs the mapping through the analysis and matching system, and returns a high affinity playlist of songs that replaces the existing station. This process automatically connects the user's high affinity towards the current song

with other songs that have similar mappings and hence a strong likelihood of continuing the user's high affinity state/space.

[0052] Thus, in exemplary detail, at 700, a user finds a computing device having a user interface in accordance with the present invention for accessing any of a variety of types of media, such as music. At 705, the user searches and finds a media entity, such as a song, of high affinity to the user. At 710, the link associated with the song starts the automatic personalization process as described in connection with FIG. 4, although any of the above-described embodiments may be used to build an initial high affinity playlist. At 715, a search of the analysis and matching database is performed to retrieve similarly mapped songs for building into an initial playlist at 720. At 725, the user begins listening to the newly retrieved media from the initial playlist. At 730, the user decides that the current song playing corresponds to a higher affinity state/space than what the playlist offers more generally. Thus, at 735, the user selects a link or other input component to indicate that the user would like to specify his or her preferences more in line with the presently playing song, which begins an automatic personalization process that further hones a playlist to the user's newly specified preference for the presently playing song. At 740, the mappings of the currently playing song are run through the analysis and matching database, to return new media entities for a new playlist at 745. At 750, the user experiences other media with similar psychoacoustic properties as the "Get More" selected song. At 755, the user can opt to prolong the high affinity state/space associated with the newly formed playlist generated as a consequence of the "Get More" selected song. In an alternate embodiment, the mappings represented by the "Get More" song may be used to supplement the mappings represented by the initial playlist, such that the "tweaking" of the playlist is more subtle than generating a brand new playlist.

Thus, in Paragraph [0051], the musical property mapping of a currently playing song is captured. The musical property mapping of a song is described in Paragraph [0043] of Kaiser as:

The specific choice within any of the features can be represented as a mapping along a set of fundamental musical properties that captures a user's psychoacoustic preferences.

Thus, the musical property mapping of a song, as taught in Kaiser, merely describes properties of a single song. Musical property mapping does not disclose any information relating to determining a temporal sequence of songs, using any technique. Furthermore, musical property mapping of Kaiser fails to disclose determining a temporal sequence of songs using a template. The phrase musical property mapping, contrary to the Office Action, thus fails to disclose a template, in accordance with which a temporal sequence may be determined.

Furthermore, Kaiser, in the highlighted portion of Paragraph [0051], merely states that a high affinity playlist of [sic] songs is returned. Kaiser fails to disclose how [sic] the individual songs within this playlist might be arranged in a temporal sequence, whether using a template or otherwise.

Paragraph [0052] of Kaiser is similarly devoid of any indication of how the playlist is arranged in a temporal sequence, whether using a template or otherwise. Paragraph [0052] of Kaiser merely indicates that new media entities are returned for the new playlist. Kaiser fails to disclose how the returned new media entities are arranged in a temporal sequence to generate the new playlist. Kaiser here clearly fails to disclose arranging the returned new media entities in a temporal sequence using a template, as no disclosure is provided whatever as to how the returned new media entities are arranged in a temporal sequence.

Furthermore, Paragraph [0052] of Kaiser vaguely indicates that the mappings represented by the “Get More” song may be used to supplement the mappings represented by the initial playlist, such that the “tweaking” of the playlist is more subtle than generating a brand new playlist, but fails to describe how such a supplement would be implemented. As an example of this failure to describe implementation of the supplement, Kaiser is silent as to how a temporal sequence of songs of a playlist made up of the initial playlist and additional songs returned from the analysis and matching database would be determined. As no disclosure is provided as to how the temporal sequence for the playlist would be determined, it is clear that use of a template to implement the supplemented playlist is not disclosed.

Indeed, the remainder of Kaiser provides no disclosure as to how a temporal sequence for a playlist is determined. For example, Paragraph [0061] merely states:

[0061] This overall high/low-affinity preference profile may then be utilized as a basis for dynamically seeding a playlist generator with likelihood weightings for all potential songs in the roll-up. Songs matching the high affinity state are weighted as more likely to play. Songs matching the low-affinity state are weighted as less likely to play, or these songs are blocked from playing altogether if enough other songs exist to generate a playlist of acceptable length.

This paragraph discloses that songs matching a high affinity state are more likely to be included in the playlist, and songs matching a low affinity state are less likely to be included in the playlist, but provides no information as to determination of a temporal sequence of those songs to be included in the playlist. Clearly, no template is disclosed for determining the temporal sequence.

For at least the foregoing reasons, reconsideration and withdrawal of the Section 103(a) rejection of claim 1 over Kaiser in view of Mackintosh is respectfully requested.

Id. at 168-72. The applicant also argued that the rejection of claim 1 over the combination of Kaiser and Begeja was traversed for similar reasons. *Id.* at 172.

62. The examiner entered a final rejection on May 31, 2013. *Id.* at 180-91.

The examiner maintained the rejections of claim 1 over both the combination of Kaiser and Mackintosh and also the combination of Kaiser and Begeja. Regarding the disclosure of a “temporal sequence” in Kaiser, the examiner wrote:

Applicant’s arguments filed 05/28/2013 have been fully considered but they are not persuasive. Regarding claim 1, applicant argues that Kaiser does not teach “a template or selecting a temporal sequence in accordance with a template”, since the phrase media mapping never appears in the Kaiser reference, remarks pages 6 - 7.

However, the examiner respectfully disagrees with the applicant. Kaiser discloses personal station playlists of media content, wherein the user is able to improve and tailor a playlist to the user's newly specified preferences; paragraphs [0051] [0052] figure 7. Moreover, Kaiser teaches that once the user has pressed the "get more" button 735 in order to obtain more clips related to the currently playing media, mappings of the currently playing media are run through the analysis and matching database, to return new media entities for a new playlist at 745, paragraph [0052].

Kaiser further recites that these pre-defined mappings identify properties representative of each individual media, wherein each mapping allows the analysis and matching system to scan the database for other media entities that have a similar media mapping properties and return other related high affinity media to further tailor the user's playlist (such as for the "get more" user selected media), paragraphs [0043] [0051].

The examiner notes that since the currently claimed template is broad in scope, it may be interpreted as the media mappings taught by Kaiser. If the applicant intended to give the claimed template a different meaning from the examiner's interpretation, the current claim language should be amended to capture and define how the claimed template differs from the teachings of Kaiser.

Therefore, Kaiser clearly teaches "selecting and determining a second temporal sequence for, by the computer and in accordance with a template, a plurality of video clips including the first video clip, to define second video programming", as currently claimed.

Id. at 182-83.

63. The applicant filed a request for continued examination on November 5, 2013, accompanied by amendments and remarks. *Id.* at 196-206. The applicant amended claim 1 in the following manner:

selecting by the computer a plurality of video clips including the first video clip and determining a second temporal sequence for the selected plurality of clips, wherein said selecting and determining is performed by the computer and in accordance with a template, ~~a plurality of video clips including the first video clip,~~ to define second video programming;

wherein the template comprises a plurality of variables corresponding to potential viewing scenarios of two or more selected video clips in a temporal relationship to each other, and

presenting by the computer, responsive to said user selection, said second video programming in the first area of the display.

Id. at 200. The applicant also added new claims 2-10. *Id.* at 200-01. The applicant characterized the amendment as follows: “Claim 1 has been amended. The step of selecting a plurality of clips *and* determining a temporal sequence for the clips ‘is performed in accordance with a template’ and also, ‘the template comprises a plurality of variables corresponding to potential viewing scenarios.’” *Id.* at 202.

64. Regarding the basis of the pending obviousness rejections, the applicant argued that they were traversed with the new claim amendments:

In the previous Response, the Applicant argued that: “Kaiser fails to disclose a template, or selecting and determining a temporal sequence in accordance with a template, as recited in claim 1.” Response of May 28, 2013 at 5. The Examiner equated the claimed “template” with the term “media mapping,” and cited to Kaiser at ¶¶ [0051-52] and Fig. 7. As noted previously, Kaiser actually uses the term “musical property mapping,” which is not a “template” as claimed. The “mapped” representations of fundamental musical properties that are associated with individual songs in Kaiser is used to “further hone” a playlist (the songs that are selected), according to a “high affinity” state. Kaiser, at ¶¶ [0016]; [0031]; [0051-52]. This does not provide a “template” that is used to determine a temporal sequence of selected songs (or other media) according to claim 1. The individual songs in Kaiser’s playlist are not arranged in a temporal sequence (in any particular order in relation to each other) according to the “mapping” data or otherwise.

No “template” as claimed is provided for this purpose. Response, at 7-9. As explained in reference to ¶[0061] of Kaiser, a preference profile can be used (with the mapping data) to determine the likelihood that a song will play, or to block songs altogether. The order of play, i.e. the temporal sequence of songs that actually play, is not provided for in Kaiser, and apparently is random or inconsequential. Id., at 9-10.

The Examiner has not found these arguments to be persuasive, and the rejection from the previous Office Action has been maintained. The Examiner explains that Kaiser improves and tailors a playlist to a user’s newly specified preferences, by returning new selections of media with similar mapping properties, upon request (Fig. 7 and the “get more” button). The claimed “template” is interpreted to broadly encompass this feature of Kaiser. Office Action, at 2.

Thus, the Office Action states (at 3): “The examiner notes that since the currently claimed template is broad in scope, it may be interpreted as the media mappings taught by Kaiser. If the applicant intended to give the claimed template a different meaning from the examiner’s interpretation, the current claim language should be amended to capture and define how the claimed template differs from the teachings of Kaiser.”

Initially, the Applicant notes that the term “template” in Claim 1 is not used in isolation. Even if, standing alone, the word “template” may broadly be thought to include Kaiser’s “media mapping properties” (Office Action, at 2), the claimed “template” must expressly be used for “selecting and determining a second temporal sequence (i.e. “in accordance with a template”). See, original Claim 1. Kaiser does not disclose a “template” or “property mappings” that provides this claimed step, so that the temporal sequence (order)of [sic] selected media clips depends on a template.

Kaiser does provide a playlist, and a playlist conventionally offers media to be played, one selection at a time, in an order from first to last selection. However, Kaiser uses its “mappings” to identify what might be added to or removed from a playlist as a selection, not when any one or more existing selections will be played and in what order or temporal sequence. In other words, the selection process is distinct from the sequencing process. Claim 1 provides selection steps, but in addition, provides steps and features for combining and arranging the selections

into media programming - steps and features not suggested or described by the reference. The invention can be thought of as picking up where Kaiser left off, by assembling previously selected media elements (whether or not they are selected by matching to pre-defined mappings per Kaiser) in an ordered or temporal sequence that is not random, undefined, inconsequential or unspecified; it is done “in accordance with a template.” Respectfully, Claim 1, as previously presented, does distinguish over Kaiser and its “mappings,” when the steps of the claim, and the language of the claim, is taken in context and as a whole. Thus, Applicant believes the rejection should be withdrawn.

Additionally, to advance prosecution, and for the sake of clarity, Claim 1 has been amended. The amendment highlights that selecting a plurality of clips and determining a temporal sequence for the clips “is performed in accordance with a template” and also, “the template comprises a plurality of variables corresponding to potential viewing scenarios.” This amendment is supported by original Claim 1, by the Specification at ¶¶ [0007-0008]; [0016]; [0018]; and the Appendix, at p. 3 (viewing scenarios).

As discussed above, the “property mapping” in Kaiser concerns fundamental musical properties that are associated with individual songs, so that songs having similar properties can be matched to each other and proposed for a playlist. The invention provides additional steps and features not found in Kaiser, where video clips are selected according to a template that comprises variables corresponding to potential viewing scenarios. A “template” broadly means a pattern or a model, and in computing broadly means rules or parameters by which variables can be processed to provide actual arguments or results. Here, the claimed “template” provides a selection and a temporal sequence of clips according to different viewing scenarios, governed by corresponding variables, and not just by matching similar clips to each other from a database of fundamental musical properties. The nature of the claimed template, and its use, is different from the “property mapping” in Kaiser.

Mackintosh does not supply the deficiencies in Kaiser. Accordingly, Applicant respectfully requests the Examiner to reconsider and withdraw the rejection. Claim 1, particularly as amended, is not obvious over Kaiser and Mackintosh.

Id. at 203-05. The applicant said these arguments also applied to the combination of Kaiser and Begeja. *Id.* at 205-06.

65. In the next office action from December 27, 2013, the examiner rejected the amended and new claims 1-10 as obvious over the combination of Kaiser, Marcus156, and Begeja. *Id.* at 217-26. Kaiser continued to be the base reference. Regarding the additional contributions of Marcus156 and Begeja, the examiner wrote:

However, it is noted that Kaiser is silent to explicitly disclose that the template comprises a plurality of variables corresponding to potential viewing scenarios of two or more selected video clips in a temporal relationship to each other.

Nevertheless, in a similar field of endeavor Marcus also discloses that the template comprises a plurality of variables corresponding to potential viewing scenarios of two or more selected video clips in a temporal relationship to each other (Col. 9 lines 9-21, col. 13 lines 45-51, col. 14-13-29 figures 5 and 6; utilizing various templates 505 to assemble video clips based on a temporal organization for the media elements, implementing parameters such as aesthetic features, psychographic characteristics, etc.).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser by specifically providing the elements mentioned above, as taught by Marcus, for the predictable result of automatically creating and selecting visually smooth transitions between adjoining clips in order to properly cover a requested subject matter in an appropriate sequence (Marcus - col. 1 lines 39-47).

However, it is noted that Kaiser and Marcus are silent to explicitly disclose presenting by the computer said first video programming in a first area of a display; simultaneously with said presenting said first video programming, presenting by the computer in a second area of the display one or more frames from recently-displayed video clips from

the first video programming; and receiving by the computer a user selection of one of the displayed frames, the user-selected displayed frame being a frame of a video clip[.]

Nevertheless, in a similar field of endeavor Begeja also discloses a clip presenting system (figure 5) which allows the users to obtain more clips related to a selected clip (Paragraph [0042]). Begeja further discloses presenting by the computer said first video programming (current clip 510) in a first area of a display area (Paragraph [0069] figure 5);

simultaneously with said presenting said first video programming (current clip 510), presenting by the computer in a second area of the display one or more frames (section 505 with thumbnails representing each clip) from recently-displayed video clips from the first video programming (Paragraphs [0068] [0069] figure 5; presenting in section 505 thumbnails representing each clip, wherein the user may skip to a previous clip displayed in section 505 with button 525);

and receiving by the computer a user selection of one of the displayed frames, the user-selected displayed frame being a frame of a video clip (Paragraph [0068] [0069] figure 5; user selections using buttons 525-550).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser and Marcus by specifically providing the elements mentioned above, as taught by Begeja, for the predictable result of allowing the users to easily and effectively select clips of content from a screen which presents current and previous clips simultaneously, without the need to force the user to navigate to other screens, allowing the interface to be more user friendly and presenting media relevant to the interests of the user (Begeja - paragraph [0008]).

Id. at 219-21.

66. Regarding pending claim 3, which included the language “wherein said first and second video programming are each presented as a stream of information

to the user,” which is also specified in issued claim 13 of the ’185 patent, the applicant wrote:

Kaiser and Marcus are silent to explicitly disclose that said first and second video programming are each presented as a stream of information to the user.

Nevertheless, in a similar field of endeavor Begeja discloses that said first and second video programming are each presented as a stream of information to the user (Paragraph [0030]; segments may be streamed to the user).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser and Marcus by specifically providing the elements mentioned above, as taught by Begeja, for the predictable result of implementing a means of providing an easy-to-use delivery methodology for the user, and as a means of conserving bandwidth (Begeja - paragraph [0030]).

Id. at 222.

67. Regarding pending claim 5, which included the language “wherein the template variables comprise sequencing limitations for incorporating video clips at specific locations in the second media programming,” which is similar to language contained in issued claims 6 and 14 of the ’185 patent, the applicant wrote:

Regarding claim 5, Kaiser, Marcus and Begeja disclose a method according to claim 1; however, Kaiser is silent to explicitly disclose that the template variables comprise sequencing limitations for incorporating video clips at specific locations in the second media programming.

Nevertheless, in a similar field of endeavor Marcus discloses that the template variables comprise sequencing limitations for incorporating video clips at specific locations in the second media programming (Col. 9 lines 9-21, col. 13 lines 45-51, col. 14-13-29 figures 5 and 6; utilizing various templates 505 to assemble video clips based on a temporal

organization for the media elements, implementing parameters such as aesthetic features, psychographic characteristics, etc. For example, it may not be possible for aesthetic reasons to place a certain clip immediately before or after a certain other clip, col. 14 lines 27-29).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser by specifically providing the elements mentioned above, as taught by Marcus, for the predictable result of automatically creating and selecting visually smooth transitions between adjoining clips in order to properly cover a requested subject matter in an appropriate sequence (Marcus - col. 1 lines 39-47).

Id. at 222-23.

68. Regarding pending claim 6, which included the language “wherein the sequencing limitations comprise at least one of length requirements and style requirements,” which is similar to language contained in issued claims 1 and 7 of the ’185 patent, the applicant wrote:

Regarding claim 6, Kaiser, Marcus and Begeja disclose a method according to claim 5; however, Kaiser is silent to explicitly disclose that the sequencing limitations comprise at least one of length requirements and style requirements.

Nevertheless, in a similar field of endeavor Marcus discloses that the sequencing limitations comprise at least one of length requirements and style requirements (Col. 9 lines 9-21, col. 13 lines 45-51, col. 14-13-29 figures 5 and 6; utilizing various templates 505 to assemble video clips based on a temporal organization for the media elements, implementing parameters such as aesthetic features, psychographic characteristics, etc. For example, it may not be possible for aesthetic reasons to place a certain clip immediately before or after a certain other clip, col. 14 lines 27-29).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser by specifically providing the elements mentioned above, as taught by Marcus, for the

predictable result of automatically creating and selecting visually smooth transitions between adjoining clips in order to properly cover a requested subject matter in an appropriate sequence (Marcus - col. 1 lines 39-47).

Id. at 223-24.

69. Regarding pending claim 9, which included the language “wherein one or more video clips comprise advertising,” which is the same language contained in issued claims 8 and 16 of the ’185 patent, the applicant wrote:

Regarding claim 9, Kaiser, Marcus and Begeja disclose a method according to claim 1; Kaiser discloses presenting the one or more video clips (see claim 1). However, Kaiser is silent to explicitly disclose that the one or more video clips comprise advertising.

Nevertheless, in a similar field of endeavor Marcus discloses that the one or more video clips comprise advertising (Col. 3 lines 52-55; assets like promotional segments, advertisements, news flashes). Moreover, Begeja also teaches that the one or more video clips comprise advertising (Paragraph [0026]).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser by specifically providing that the one or more video clips comprises advertising, for the predictable result of implementing an automated and user tailored playlist of content, including ads, that is more likely to be enjoyed by the viewer, hence increasing exposure time of the viewer to the successfully advertised products.

Id. at 224-25. Marcus156 was listed in a Notice of References Cited accompanying the office action. *Id.* at 227.

70. The applicant filed responsive amendments and remarks on March 17, 2014. *Id.* at 232-42. Pending claims 1-10 were canceled and new claims 11-30

(similar to issued claims 1-20) were added. *Id.* at 233-36. The applicant summarized its view of the nature of the invention represented by the new claims:

The invention as now claimed provides a computer-implemented method for a user to search and browse videos in a customized manner, make a selection from the customized presentation, and receive further customized video programming based on the selection, according to an association between a second video clip and a first video clip that is indicated by the user's selection. This association can comprise media tags, which correspond to categories of information including context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history. First and second video clips, comprising first and second video programming, are selected, determined and concatenated according to templates comprising information about the user, viewing scenarios of two or more selected video clips in a temporal relationship to each other, and style information. Thus, templates and video associations are used in concert to provide a user with at least two sets of video programming that are related to each other and customized.

In certain embodiments a first video clip represents a summary of a second video clip. Sequencing limitations can also be applied to require that certain clips be presented together, in a particular order or priority, at specific locations, or within a particular time. The video clips can also comprise transitional clips, e.g. hosting clips, that can be placed between other video clips, for example in connection with smoothly concatenating the video programming according to template requirements. Template variables comprising user information include at least one of demographic information, psychographic information, and a viewing history.

The original Specification, including the accompanying Appendix, provide support for the new claims. See, for example, pages 80-81 of the Appendix, regarding clip types and media tags.

Id. at 237.

71. Regarding the previous set of rejections and the concept of “temporally arranged media programming” in the combination of Kaiser and Marcus¹⁵⁶, the applicant wrote:

Previously the Applicants made arguments to distinguish Kaiser from the claimed invention, and in particular endeavored to explain that a “playlist” is not temporally arranged media programming as claimed. The Office Action does not comment on the Applicant’s previous arguments, however the previous rejections based on Kaiser in view of Macintosh or Begeja, have been withdrawn. The pending rejection attempts to import the concept of temporal programming from Marcus into the playlists of Kaiser. This is misplaced, for the following reasons.

First, the Examiner continues to rely on the “playlists” and “mappings” of Kaiser as disclosing media programming and templates of claim 1. See, Kaiser at ¶¶ [0051-52] and Fig. 7. Kaiser’s playlist identifies a set of video clips that are to be presented to a user, one at a time, according to musical properties that in some way suggests that the listed clips, though disconnected and not necessarily related to each other, can be placed together in a group. The “high affinity” criteria used by Kaiser to “further hone” a playlist is based on artistic attributes, in a system that does not contemplate programming having a temporal content and a sequence-sensitive context, nor provide templates that account for and manage the delivery of media elements as coherent programming according to temporal requirements. The individual songs (or clips) in Kaiser are not arranged in a temporal sequence - they have no particular order in relation to each other. A playing order arises only coincidentally, because each clip stands alone, and only one clip at a time is played. Thus, the order of play, i.e. the temporal sequence of songs that actually play, is not provided for in Kaiser, and apparently is random or inconsequential. The Applicants believe that a person of ordinary skill in the art would not apply temporal considerations to a playlist that has none, and therefore would not borrow from Marcus to do so. Any such extrapolation would require impermissible hindsight from the present invention.

In the last Response, it was noted that Kaiser uses “mappings” to identify what media might be added to or removed from a playlist as a selection, not whether such media actually should be presented or displayed, and if so, when a media element (e.g. a video clip) will be played, and in what order or temporal sequence, with respect to other media elements making up the media programming. In other words, the selection process is distinct from the sequencing process. New claim 11 provides selection steps, but in addition, provides steps and features for combining and arranging (determining and concatenating) the selections into media programming according to sequencing requirements, and according to other criteria. These features are not provided by Kaiser. As claimed, video clips are selected, and then the video clips to be selected are sequenced and concatenated according to template variables - steps and features not suggested or described by the reference.

The Office Action attempts to bolster Kaiser by pointing to the temporal considerations in Marcus. However, a person of ordinary skill in the art would not replace the “what to play” criteria of Kaiser with a “when to play” concept from Marcus. Concepts of “what” and “when” are plainly quite different and implicate different considerations with different requirements and results.

Furthermore, the artisan would not add a temporal concept from Marcus to the playlist in Kaiser, because the playlist and its independent clips have no temporal basis upon which to work. Respectfully, new claim 1 is not obvious over Kaiser and Marcus, because a person of ordinary skill in the art would not combine these references to achieve the invention. Marcus furthermore does not provide the currently claimed combination of steps nor make them obvious. Marcus provides a method and system for creating seamless temporally ordered video programming from clips that are not required to be part of the same original programming nor have a preexisting temporal order with respect to each other. Marcus does not go further, and provide a method and system of the invention, and its multiplicity of branched and inter-related programming. The invention, unlike Marcus, provides customized first video programming as a basis from which to select a wide range of additional second video programming that is also customized, and which adheres to template rules by which this different video programming is related or consistent, and also in each case is

presented in an appropriate temporal order and in a consistent style, and according to user information. The present invention improves upon Marcus in a material and non obvious way.

Id. at 239-41.

72. Next, a notice of allowance issued on April 14, 2014. *Id.* at 253-61. The examiner wrote that “[r]egarding claim 11, the prior art of record (individually or in combination) fails to explicitly or implicitly suggest the features of claim 11” and listed the full text of claim 11 save for the preamble. *Id.* at 258-59. “Regarding claims 12 – 30, they depend from allowable claim 11. Therefore, claims 12 – 30 are also held allowable.” *Id.* at 259.

73. The applicant submitted a request for continued examination and an accompanying information disclosure statement on May 29, 2014. *Id.* at 274-86.

74. Another notice of allowance issued on June 24, 2014. *Id.* at 294-303.

75. The ’709 application issued as the ’185 patent on October 28, 2014. *Id.* at 328.

4. The ’110 Application (Ex. 1006)

76. The ’110 application was filed as a continuation of the ’709 application on September 23, 2014. Ex. 1006 at 1-8, 11. The initial set of claims was similar to the claims that ultimately issued. *Id.* at 12-15.

77. In the first office action from May 4, 2015, all pending claims were rejected on double patenting grounds in light of the issued claims of the ’185 patent:

Claims 2-21 are rejected on the ground of nonstatutory double patenting as being unpatentable over claim 1 of U.S. Patent No. 8,875,185. Although the claims at issue are not identical, they are not patentably distinct from each other because at least independent claim 2 of the instant application is a broader version similar to the subject matter scope captured by claim 1 of U.S. Patent No. 8,875,185.

Id. at 68. An examiner-initiated interview summary noted that the examiner had advised the applicant to file a terminal disclaimer. *Id.* at 70. The applicant submitted a terminal disclaimer on November 4, 2015. *Id.* at 79-88. A notice of allowance issued on December 16, 2015. *Id.* at 93-103.

78. In an office action accepting the terminal disclaimer and making a typographical examiner's amendment, the examiner also included a section entitled "Citation of Pertinent Prior Art" and wrote that "The prior art made of record and not relied upon is considered pertinent to the applicant's disclosure: . . . Marcus (Patent No US 6,032,156) – utilizing templates to assemble video clips in a play list, Col. 9 lines 9-21, col. 13 lines 45-51 figures 5 and 6." *Id.* at 100-01. Marcus156 was also included in a notice of references cited. *Id.* at 105. The pending claims of the '110 application issued as the '597 patent on April 19, 2016. *Id.* at 121.

5. The '873 Application (Ex. 1007)

79. The '110 application was filed as a continuation of the '110 application on April 18, 2016. Ex. 1007 at 1-8, 12. The application included a single claim 1, similar to issued claim 1 of the '185 patent. *Id.* at 23.

80. The initial office action rejected claim 1 on double patenting grounds in light of the '185 patent and the '597 patent. *Id.* at 58-61. Additionally, the claim was rejected as obvious in light of the combination of U.S. Patent Publication No. 2002/0019858 ("Kaiser") and U.S. Patent No. 6,317,784 ("Mackintosh"). *Id.* at 61-64. Claim 1 was also rejected as obvious in light of the combination of Kaiser and Begeja. *Id.* at 64-66. The examiner stated that both Mackintosh and Begeja disclosed simultaneous display of a video and a frame, as well as user selection of the frame. *Id.* at 61-66. Regarding obviousness in light of Kaiser and Mackintosh, the examiner wrote:

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kaiser by specifically providing the elements mentioned above, as taught by Mackintosh, for the predictable result of allowing the users to easily and effectively selects clips of content from a screen which presents current and previous clips simultaneously, without the need to force the user to navigate to other screens, allowing the interface to be more user friendly.

Id. at 63-64. The examiner supplied a similar motivation to combine Kaiser and Begeja. *Id.* at 66.

81. In response, the applicant canceled pending claim 1 and added different claims 2 and 3. *Id.* at 78-82. The new claims were significantly shorter and directed to partitioning an information display into two or three areas and simultaneously displaying multiple videos in the display areas. *Id.* at 80. Regarding the obviousness rejection of previously-pending claim 1 in light of Kaiser and Mackintosh, the

applicant wrote: “While Mackintosh displays multiple areas, they display static images of previous audio clips, via album covers, and previous advertising. None of these are ‘played’ as these are only historical information and not the clips themselves. The present claims maintain the previous videos playing in the second and third viewing areas, as the next clip moves into the primary viewing area.” *Id.* at 81. The applicant did not attempt to distinguish Begeja.

82. In the next office action from October 6, 2017, the examiner maintained the rejection of pending claims 2 and 3. *Id.* at 89-99. Claim 2 was rejected “on the ground of nonstatutory double patenting as being unpatentable over claim 1 of U.S. Patent No. 9,317,597 in view of Begeja et al. (Pub No US 2004/0025180).” *Id.* at 93. The examiner noted that “Begeja discloses playing, in the second viewing area of the information display, the primary video clip (Paragraphs [0068] [0092] figure 5; video clips in section 505 may be represented as thumbnails, video thumbscreens or video files).” *Id.* The examiner argued that there was a “predictable result of enriching a user interface video clip representations in a manner that is more visually pleasing to the user, effectively increasing customer satisfaction.” *Id.* at 94. Moreover, pending claims 2 and 3 were rejected as obvious in light of the combination of Mackintosh and Begeja. *Id.* at 94-97.

83. The applicant next amended claim 2 to replace “displaying” with “simultaneously playing.” *Id.* at 109. The applicant argued that “while Mackintosh

displays multiple areas, only one of the ‘viewing areas’ is being played at any one moment.” *Id.* at 110-11. Similarly, the applicant argued that Begeja only played one video at a time, not two. *Id.* at 111.

84. The examiner maintained the rejection of pending claims 2 and 3 on the basis that both Mackintosh and Begeja disclosed playing two videos simultaneously. *Id.* at 123-33. As part of making this argument, the examiner again noted that both references disclosed displaying a playing video in one display area and a frame in another. *Id.* at 125-26.

85. The applicant a response with amendments and remarks on August 16, 2018. The applicant amended pending claim 2 as follows:

2. (Currently Amended) A method of providing video programming, comprising:

providing, by a computer, an information display and partitioning the information display to comprise:

a first viewing area;

a second viewing area; and

playing, in the first viewing area of the information display, a primary video clip at a first point in time;

playing, in the first viewing area of the information display, a secondary video clip at a second point in time, where the second point in time occurs after the first point in time; and

playing, in the second viewing area of the information display, the primary video clip ~~when the first viewing area is simultaneously playing a secondary video clip~~ at the second point in time.

Id. at 140. The applicant argued that the combination of Mackintosh and Begeja did not disclose the amended claim language, which the applicant summarized as “playing two video clips in two viewing areas at the same time.” *Id.* at 141-42.

86. In the next office action on December 13, 2018, the examiner maintained the rejection of all pending claims 2-3. *Id.* at 146-58. The examiner argued that both Mackintosh and Begeja each disclosed “playing two video clips in two viewing areas at the same time.” *Id.* at 148-50.

87. In the applicant’s response from March 13, 2019, pending claims 2 and 3 were both significantly amended:

2. (Currently Amended) A method of providing video programming, comprising:

providing, by a computer, an information display and partitioning the information display to comprise:

a first viewing area;

a ~~second plurality of mini-display viewing areas~~; and

playing, in the first viewing area ~~of the information display~~, a primary video clip at a first point in time;

replaying in a loop, in the mini-display areas, a plurality of clips historically played in the main viewing area;

playing, in the first viewing area ~~of the information display~~, a secondary video clip at a second point in time, where the second point in time occurs after the first point in time; ~~and~~

replacing, in the mini-display areas, a least recently displayed clip of the plurality of clips with the primary video clip ~~playing, in the second viewing area of the information display, the primary video clip at the second point in time; and~~

replaying the primary video clip in the loop in the mini-display areas until the primary video clip is replaced or selected.

3. (Currently Amended) The method of claim 2, wherein the plurality of mini-display areas include a second viewing area and a third viewing area, and wherein the method further comprising comprises the steps of:

~~partitioning the information display to comprise a third viewing area;~~

playing the primary video clip in the second viewing area at the second point in time;

playing the primary video clip in the third viewing area when the secondary video clip completes displaying in the first viewing area; and

playing the secondary video clip in the second viewing area when the primary video clip is displayed in the third viewing area.

Id. at 171. The applicant argued that the prior art of record did not disclose the newly-added limitations. *Id.* at 168-69.

88. The examiner issued an advisory action on March 22, 2019 that the applicant's March 19, 2022 response was filed after a final rejection but was not accompanied by a notice of appeal or request for continue application. *Id.* at 173-74.

89. The applicant filed a request for continued examination on April 15, with new amendments and remarks different from the March 19, 2019 amendments. *Id.* at 176-86. The applicant amended pending claims 2 and 3 as follows:

2. (Currently Amended) A method of providing video programming, comprising:

providing, by a computer, an information display and partitioning the information display to comprise:

a first viewing area;

a ~~second~~ second display viewing area; and

playing, in the first viewing area ~~of the information display~~, a primary video clip at a first point in time;

presenting, in the second display area, one or more frames from one or more video clips historically played in the main viewing area;

playing, in the first viewing area ~~of the information display~~, a secondary video clip at a second point in time, where the second point in time occurs after the first point in time; and

replacing, in the second display area, a frame of a least recently displayed video clip of the one or more video clips with a frame of the primary video clip ~~playing, in the second viewing area of the information display, the primary video clip at the second point in time~~

3. (Currently Amended) The method of claim 2, wherein the second display area includes a second viewing area and a third viewing area, and wherein the method further comprising ~~comprises~~ the steps of:

~~partitioning the information display to comprise a third viewing area;~~

playing the primary video clip in the second viewing area at the second point in time;

playing the primary video clip in the third viewing area when the secondary video clip completes displaying in the first viewing area; and

playing the secondary video clip in the second viewing area when the primary video clip is displayed in the third viewing area.

Id. at 182. The applicant also added new claims 4-8. *Id.* at 182-83. The applicant argued that the newly-amended claims were not disclosed or rendered obvious by the combination of Mackintosh and Begeja. *Id.* at 184-85.

90. In the next office action from June 6, 2019, the examiner rejected pending claims 2-8. *Id.* at 192-205. The examiner argued that Mackintosh disclosed the new claim limitations in pending independent claims 2 and 5. *Id.* at 194-96, 199-204. Additionally, regarding Begeja, the examiner stated:

In this instance, Begeja discloses that it is known in the art to provide a representation of video clips, e.g. section 505, as thumbnails as shown on figure 5, or alternatively as video thumbscreens or video files; paragraphs [0068] [0092] figure 5, along with a main video section 510 being played.

Enriching a user interface video clip representations in a manner that is more visually pleasing to the user, effectively increasing customer satisfaction. Therefore, the combination of Mackintosh in view of Begeja clearly discloses the features of claims 2 and 5, as presented in the current office action.

Id. at 196. The examiner also issued a variety of double patenting rejections. *Id.* at 196-99.

91. The applicant submitted amendments and remarks on September 6, 2019. *Id.* at 211-19. Pending independent claims 2 and 5 were significantly amended to add numerous new limitations. *Id.* at 212-14. The applicant paraphrased the new claim language and contended that the currently-cited prior art did not disclose the new limitations. *Id.* at 216-19.

92. All pending claims 2-10 were rejected in the next office action on December 9, 2019. *Id.* at 223-41. The examiner issued a variety of double patenting

rejections. *Id.* at 225-28. The examiner also rejected claims 2, 5, 9, and 10 as lacking written description:

Claims 2, 5, 9 and 10 are rejected under 35 U.S.C. 112(a) or 35 U.S.C. 112 (pre-AIA), first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor or a joint inventor, or for pre-AIA the inventor(s), at the time the application was filed, had possession of the claimed invention.

Regarding claim 2, the applicant currently claims “receiving, by a processor of a media delivery server, login information of a user from the user’s device” in lines 2 and 3. However, the disclosure of the instant application as originally filed fails to explicitly or implicitly point out to any type of a user login.

Regarding claims 2 and 5, the applicant currently claims “a unique unicast” in lines 9, 11, 16 of claim 2 and lines 6 and 7 of claim 5. The examiner notes that the disclosure of the instant application as originally filed fails to explicitly or implicitly disclose the claimed “a unique unicast” content distribution scheme.

Regarding claims 2 and 5, the applicant currently claims “a user’s click” in line 22 of claim 2 and line 13 of claim 5. However, the examiner notes that the disclosure of the instant application as originally filed fails to explicitly or implicitly disclose the claimed clicking mechanism.

Regarding claim 9, the applicant currently claims that “visual exemplary representation includes at least one of an image ... text information, and an audio file”. The disclosure teaches that the mini-displays 230, i.e. visual exemplary representation of played media clips, in figure 2 may be a clip replays in a loop; paragraph [0021].

However, the examiner notes that the originally filed disclosure fails to explicitly or implicitly disclose that the “visual exemplary representation (e.g. mini-displays 230) include at least one of an image ... text information, and an audio file” as claimed.

Regarding claim 10, the applicant currently claims that “the predefined rule is based on at least one of the following: contextual understanding of the user, and time of day”. The disclosure teaches collecting user information such as the user’s demographic data and geographic data (figure 1); however, it is silent to explicitly or implicitly disclose collecting “contextual understanding of the user, and time of day” as claimed.

Id. at 228-30.

93. The examiner rejected pending claims 2 and 4-10 as obvious over the combination of Kaiser, Mackintosh, and U.S. Patent No. 6,199,076 (“Logan”). *Id.* at 230-37.

94. On May 11, 2020, the applicant again significantly amended the pending claims to add limitations about “a first class of non-monetizing media clips and a second class of monetizing media clips” and “short-form teasers of advertisements.” *Id.* at 251-53. The applicant also added new claims 11-17. *Id.* at 254-55. The applicant amended claim to remove the terms that the examiner has contended lacked written description. *Id.* at 256. The applicant argued that the prior art of record failed to teach the limitations of the newly-amended and newly-added claims. *Id.* at 256-60.

95. The examiner issued an advisory action on May 20, 2020 that the applicant’s May 11, 2020 response was filed after a final rejection but was not accompanied by a notice of appeal or request for continue application. *Id.* at 268-70.

96. The applicant submitted a request for continued examination on June 9, 2020 and referenced the amendments and remarks previously filed on May 11, 2020. *Id.* at 271-73.

97. In the next office action on June 25, 2020, the examiner rejected all pending claims 2-8 and 10-17. *Id.* at 280-97. Claims 5 and 13 received double patenting rejections. *Id.* at 284-86. Claims 2-4 and 10-17 were rejected as lacking written description:

Claims 2 and 13-17 are rejected under 35 U.S.C. 112(a) or 35 U.S.C. 112 (pre-AIA), first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor or a joint inventor, or for pre-AIA the inventor(s), at the time the application was filed, had possession of the claimed invention.

Regarding claim 2, the applicant currently claims “searching ... in near-real time”. The original disclosure is silent to explicitly or implicitly disclose searching “in near-real time”.

Regarding claim 2, the applicant currently claims “searching ... a viewing history database for user's profile after receiving the user's request”. The original disclosure is silent to explicitly or implicitly disclose the sequence of events in the order that was claimed. While paragraph [0017] and figure 1 teach collecting user information regarding demographic and prior history of interaction with the system; they is nothing in the specification that points to searching the collected user profile information “after receiving the user's request”. The examiner suggests deleting the claimed “after receiving the user's request” feature.

Regarding claim 2, the applicant currently claims “wherein revenue generated from each monetizing media clip is shown as a credit in the user's profile”. The original disclosure is silent to explicitly or

implicitly disclose any sort of credit assigned to a user profile, even less showing said credit in the user's profile.

The examiner notes that dependent claims 3, 4 and 10-12 are also rejected under 35 U.S.C. 112(a) or 35 U.S.C. 112 (pre-AIA), first paragraph, as failing to comply with the written description requirement; since dependent claims 3, 4 and 10-12 also incorporate all the rejected new matter elements as rejected for independent claim 2.

Regarding claim 13, the applicant claims "a collection of user tags or headers defining information about each user". The examiner notes that the disclosure of the instant application as originally filed fails to explicitly or implicitly disclose the claimed.

Regarding claim 13, the applicant claims "compare each user's profile data with the stored collection of user tags or headers for each candidate media clip". The examiner notes that the disclosure of the instant application as originally filed fails to explicitly or implicitly disclose the claimed.

Regarding claim 13, the applicant claims "media clips with tag values appropriate for each user's profile". As noted before, the disclosure as originally filed fails to explicitly or implicitly disclose "tag values" related to a user's profile.

Regarding claim 13, the applicant claims that "each user's click or equivalent gesture serves to initiate concatenation deemed appropriate by a template or equivalent assembly rules for delivery to each user". The examiner notes that the disclosure of the instant application as originally filed fails to explicitly or implicitly disclose the claimed user's clicking mechanism.

Regarding claims 13-16, the applicant currently claims "a user's click" and "media clips which are clickable". The examiner notes that the disclosure of the instant application as originally filed fails to explicitly or implicitly disclose the claimed user's clicking mechanism.

Regarding claim 15, the applicant claims that "as each user clicks on a successively-delivered media clip, the processor returns to the user progressively longform media clips of the successively-delivered media clip, or an assembly of a plurality of concatenated media clips associated with the successively-delivered media clip". The examiner

notes that the disclosure of the instant application as originally filed fails to explicitly or implicitly disclose the claimed feature.

Regarding claim 16, the applicant claims that after selecting a clip receiving payment for “for a subsequent delivery of each successively-delivered media clip or an assembly of a plurality of concatenated media clips associated with each successively-delivered media clip”. However, the examiner notes that while the specification discloses that the clips may be advertisements or pay to view clips, there is nothing that supports receiving payment for “each successively-delivered media clip or an assembly of a plurality of concatenated media clips associated with each successively-delivered media clip” as claimed. At best, the disclosure as originally filed teaches that the user pays to view certain clips; paragraph [0025].

Regarding claim 17, the applicant that the media clips are “tagged” as having remunerative values. However, the examiner notes that while the specification discloses that the clips may be advertisements or pay to view clips, there is nothing that supports that these clips are “tagged” as claimed.

Id. at 287-90.

98. Claims 13-15 were rejected as obvious over the combination of Marcus156, Mackintosh, and Kaiser. *Id.* at 290-96. Regarding the disclosures of Marcus156, the examiner wrote:

Regarding claim 13, Marcus discloses a system of providing media distribution, comprising:

a non-transitory storage medium configured to store (Figure 1):

a collection of user tags or headers with defining information about each user (Col. 3 lines 1-30 figures 1 and 2; content tags 120 including viewer profiles 130);

and a plurality of candidate media clips (Figure 1; clips 80);

a processor configured to: compare each user's profile data with the stored collection of user tags or headers for each candidate media clip

(Col. 10 lines 9-21, col. 14 lines 8-17; using viewer database information to create a customized programming);

select a plurality of media clips with tag values appropriate for each user's profile data (Col. 9 lines 23-45 figure 6; read clip identification information, e.g. tag values, 605 and pull clips from library 610);

assemble the selected media clips (Col. 9 lines 23-45 figure 6; assemble clips into customized programming 615);

send the assembled media clips to a device of each user (Col. 9 lines 9-21, col. 13 lines 45-51, col. 14-13-29 figures 5 and 6; utilizing various templates 505 to assemble customized video clips programming to distribute to the user).

Id. at 290-91.

99. Regarding disclosure and obviousness of “each user’s click or equivalent gesture serves to initiate concatenation deemed appropriate by a template or equivalent assembly rules for delivery to each user,” the examiner wrote:

However, it is noted that Marcus and Mackintosh are silent to explicitly disclose each user's click or equivalent gesture serves to initiate concatenation deemed appropriate by a template or equivalent assembly rules for delivery to each user.

Nevertheless, in a similar field of endeavor Kaiser discloses disclose each user’s click or equivalent gesture serves to initiate concatenation deemed appropriate by a template or equivalent assembly rules for delivery to each user (Paragraphs [0053] [0052] [0056] figures 7 and 8; user selects the “get more” button associated with a current song or a previous song from the historical record of played songs 815, in order to build a new playlist of clips 745 with the same characteristics as the selected song).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Marcus and Mackintosh by specifically providing the elements mentioned above, as taught by Kaiser, for the predictable result of allowing the user to further redefine

their generated playlists with the highest affinity possible to their preferences (Kaiser - paragraph [0015]).

Id. at 292-93.

100. Claims 16 and 17 were also rejected as obvious over the combination of Marcus156, Mackintosh, Kaiser, and U.S. Patent Publication No. 2003/0040917 (“Fiedler”). *Id.* at 295-96.

101. The applicant submitted amendments and remarks on October 23, 2020. *Id.* at 305-14. Claim 2 was amended to replace “teasers of advertisements” with “monetizing media.” *Id.* at 306. Claim 5 was significantly amended to add numerous new limitations. *Id.* at 307-08. Additionally, the applicant amended claims 13-14 (*id.* at 309-310) and argued that claims 13-17, as amended, were not rendered obvious by the prior art of record (*id.* at 312-13).

102. In the next office action from January 28, 2021, all pending claims 2-5 and 10-17 were rejected. *Id.* at 320-39. For claims 2 and 13-17, the examiner noted that the objected claim language was not removed, so the written description rejections were maintained. *Id.* at 322. Regarding the disclosure of Marcus156, the examiner wrote:

Regarding claim 13, applicant argues that Marcus fails to teach the features direct to “assemble the selected media clips by a template or equivalent organizing logic governing ordering and assembly or each user's playlist in order to approximate human decisions”; remarks pages 8-9.

However, the examiner respectfully disagrees with the applicant. In this case Marcus clearly discloses utilizing various templates 505 to assemble customized video clips programming to distribute to the user; Col. 9 lines 9-21, col. 13 lines 45-51, col. 14-13-29 figures 5 and 6. Wherein temporal organizational templates 315/505 may be based on target audiences demographics utilizing to create the content clip playlist; col. 3 line 66 - col. 4 line 7, col. 8 lines 16-43; figures 3 and 5. Therefore, Marcus clearly discloses “assemble the selected media clips by a template or equivalent organizing logic governing ordering and assembly or each user's playlist in order to approximate human decisions”, as claimed.

Id. at 323. The rejection of claims 13-15 over the combination of Marcus156, Mackintosh, and Kaiser was maintained. *Id.* at 332-36. The rejection of claims 16 and 17 over the combination of Marcus156, Mackintosh, Kaiser, and Fiedler was also maintained. *Id.* at 337-38.

103. The applicant filed a response, including amendments and remarks, on June 25, 2021. *Id.* at 360-67. Pending claim 2 was amended to remove the limitations “in near-real time,” “after receiving the user’s request,” and “wherein revenue generated from each monetizing media clip is shown as a credit in the user’s profile.” *Id.* at 361. Pending claims 13-17 were canceled. *Id.* at 364.

104. A notice of allowance issued on July 8, 2021. *Id.* at 373-75. Pending claims 2-5 and 10-12 were allowed. *Id.* at 377-80. The ’873 application issued as U.S. Patent No. 11,184,686 on November 23, 2021. *Id.* at 405.

6. The '374 Application (Ex. 1008)

105. The applicant filed the '374 application on November 18, 2021 as a continuation of the '873 application. Ex. 1008 at 1-37. The set of 6 initial claims included a sole independent claim 1, which read in total “1. A method of delivering one or more media clips to a user, comprising: collecting user information including demographic information, one or more interests of the user, and a prior history of interacting with prior media clips; and delivering the one or more media clips to the user based on the collected user information, wherein the one or more media clips are concatenated together.” *Id.* at 16-17.

106. The first office action rejected all pending claims on a variety of double patenting grounds. *Id.* at 45-50. Additionally, claims 1 and 6 were rejected as anticipated by Marcus156. *Id.* at 51-53:

Claims 1 and 6 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Marcus (Patent No US 6,032,156). Hereinafter, referenced as Marcus.

Regarding claim 1, Marcus discloses a method of delivering one or more media clips to a user (Figure 6), comprising:

collecting user information including demographic information (Col. 3 line 66 – col 4 line 19; demographics of the target audience), one or more interests of the user (Col. 10 lines 10-21; viewer information may include preference information for the viewer), and a prior history of interacting with prior media clips (Col. 10 lines 10-21; viewer information may include the history of items viewed by the viewer);

and delivering the one or more media clips to the user based on the collected user information, wherein the one or more media clips are concatenated together (Col. 13 lines 23-45, co. 14 lines 13-34 figures 5

and 6; assembling 615 a concatenated stream of clips for a user based on the viewer information characteristics).

Regarding claim 6, Marcus discloses the method of claim 1; moreover, Marcus discloses that each of the one or more media clips is associated with a header (e.g. header 110) including both content information (Col. 3 lines 7-8 figure 2; content tags 120 contain information identifying content of the clips) and aesthetic information for the each of the one or more media clips (Col. 3 lines 34-46, col. 13 lines 45-53, col. 14 lines 25-32; video clip information may include parameters such as aesthetic features).

Id. at 52-53.

107. Claims 2 and 3 were rejected as obvious over the combination of Marcus¹⁵⁶ and Mackintosh:

Claims 2 and 3 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Marcus in view of Mackintosh et al. (Patent No US 6,317,784). Hereinafter, referenced as Mackintosh.

Regarding claim 2, Marcus discloses the method of claim 1; moreover, Marcus discloses the media clips (Col. 13 lines 23-45, co. 14 lines 13-34 figures 5 and 6; assembling 615 a concatenated stream of clips for a user based on the viewer information characteristics).

However, it is noted that Marcus is silent to explicitly disclose selecting a media clip from the one or more media clips, wherein the selected media clip is related to a product; and delivering an advisement clip detailing the product to the user.

Nevertheless, in a similar field of endeavor Mackintosh discloses selecting a media clip from the one or more media clips, wherein the selected media clip is related to a product (e.g. advertisements); and delivering an advisement clip detailing the product to the user (Col. 16 lines 9-37 figure 7; delivering advertisement to users, e.g. McDonalds advertisement, Sony advertisement, etc.).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Marcus by specifically

providing the elements mentioned above, as taught by Mackintosh, for the predictable result of generating profit from delivery of advertisement content to multiple viewers.

Regarding claim 3, Marcus discloses the method of claim 1; moreover, Marcus discloses the media clips (Col. 13 lines 23-45, co. 14 lines 13-34 figures 5 and 6; assembling 615 a concatenated stream of clips for a user based on the viewer information characteristics).

However, it is noted that Marcus is silent to explicitly disclose selecting a media clip from the one or more media clips, wherein the selected media clip is a part of a second media clip; and delivering the second media clip to the user (Col. 16 lines 9-37 figure 7; delivering advertisement to users, e.g. McDonalds advertisement, Sony advertisement, etc.).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Marcus by specifically providing the elements mentioned above, as taught by Mackintosh, for the predictable result of generating profit from delivery of advertisement content to multiple viewers[.]

Id. at 53-55.

108. Additionally, pending claims 4 and 5 pertaining to simultaneous playback of multiple video clips were rejected as obvious over the combination of Marcus, Mackintosh, and U.S. Patent Publication No. 2002/0166122 (“Kikinis”).

Id. at 55-57.

109. On May 17, 2023, the applicant filed a response to non-final office action, including amendments and remarks. *Id.* at 74-82. The pending claims were amended as follows:

1. (Currently Amended) A method of delivering one or more media clips to a user, comprising:
~~collecting user information including demographic information, one or more interests of the user, and a prior history of interacting with prior media clips; and~~
delivering the one or more media clips to the user ~~based on the collected user information~~, wherein the one or more media clips are concatenated together;
selecting a media clip from the one or more media clips, wherein the selected media clip is related to a product; and
delivering an advertisement clip detailing the product to the user.

2. (Cancelled).

3. (Currently Amended) ~~The method of claim 1, further comprising:~~ A method of delivering one or more media clips to a user, comprising:
delivering the one or more media clips to the user, wherein the one or more media clips are concatenated together;
selecting a media clip from the one or more media clips, wherein the selected media clip is a part of a second media clip; and
delivering the second media clip to the user.

4. (Currently Amended) The method of claim ~~[[2]]~~ 1, further comprising:
providing an information display and partitioning the information display to comprise:
a ~~first~~ first viewing area; and
a plurality of mini-display regions;
playing, in the first viewing area of the information display, the selected media clip; ~~and~~
playing, in the first viewing area of the information display, the advertisement clip;
and
playing, in the plurality of mini-display regions of the information display, the selected media clip ~~when the first viewing area is simultaneously playing the advisement clip.~~

5. (Currently Amended) The method of claim 3, further comprising:

providing an information display and partitioning the information display to comprise:

a ~~first~~ first viewing area; and

a plurality of mini-display regions;

playing, in the first viewing area of the information display, the selected media clip; ~~and~~

playing, in the first viewing area of the information display, the advertisement clip;

and

playing, in the plurality of mini-display regions of the information display, the selected media clip ~~when the first viewing area is simultaneously playing the second media clip.~~

6. (Original) The method of claim 1, wherein each of the one or more media clips is associated with a header including both content information and aesthetic information for the each of the one or more media clips.

7. (New) A method of delivering one or more media clips to a user, comprising:

delivering the one or more media clips to the user, wherein the one or more media clips are concatenated together;

providing an information display and partitioning the information display to comprise:

a first viewing area; and

a plurality of mini-display regions;

playing, in the first viewing area of the information display, a first media clip from the one or more media clips;

delivering a second media clip to the user;

playing, in the first viewing area of the information display, the second media clip; and

playing, in one of the plurality of mini-display regions of the information display, the selected first media clip.

Id. at 76-77.

110. The applicant argued that the newly-amended claims were not anticipated or rendered obvious by the previously-cited prior art:

Without acquiescing to the propriety of the Office's previous rejections, and solely in the interest of expediting prosecution of the present application, the subject matter of claim 2 has been incorporated into claim 1, thus rendering the alleged anticipation rejection by Marcus moot.

Withdrawal of the rejections under 35 U.S.C. 102 are respectfully requested. As acknowledged by the Office on p. 11 of the Office Action, Marcus is silent regarding "selecting a media clip from the one or more media clips, wherein the selected media clip is related to a product; and delivering an advisement clip detailing the product to the user" as recited in original claim 2 and now in amended claim I and "selecting a media clip from the one or more media clips, wherein the selected media clip is a part of a second media clip; and delivering the second media clip to the user" as recited in original and amended claim 3. Applicant respectfully submits that Mackintosh, as well as the rest of the cited art, fails to resolve these deficiencies in Marcus.

Mackintosh discloses that clicking on a window in a history bar "bring[s] up the information for that selection. For example, in the embodiment illustrated in FIG. 7, history window 306 shows the five most recent tracks played If, for example, the listener wanted to obtain additional information about the Beatles selection that was being played, the user may simply click on the Beatles window and that information can be displayed on data window 302." See Mackintosh, Col. 16, lines 10-21.

Regarding amended claim 1, there is no disclosure or suggestion that the additional or supplemental information in Mackintosh is an advertisement clip, as recited. Rather, if anything, the advertisements cited by the Office, "e.g. McDonalds advertisement, Sony advertisement, etc." on p. 11 of the Office Action, correspond to "a media clip" recited in claim 1. Applicant respectfully notes that "a media clip" and "an advertisement clip" refer to differently recited clips. Mackintosh provides various examples of supplemental information when broadcast material is an advertising spot, e.g., "electronic or virtual coupons" (Col. 11, lines 20-21), "a coupon that is printed" (Col. 11, lines 23-24) "a special alphanumeric code" (Col. 11, line 25), or "contest materials" (Col. 11, line 29), none of which suggest the need for or benefit of a second advertisement clip. Thus, even if Mackintosh suggested that the media clip could be an

advertisement clip, Mackintosh fails to disclose or suggest a second advertisement clip.

Thus, Marcus in view of Mackintosh fails to disclose or suggest “selecting a media clip from the one or more media clips, wherein the selected media clip is related to a product; and delivering an advertisement clip detailing the product to the user” as recited in amended claim 1. As Kikinis, whether alone or in combination with the other cited art, does nothing to resolve these deficiencies, claim 1 is patentable over the prior art of record.

Regarding amended claim 3, it is patentable for similar reasons as discussed above with respect to amended claim 1. Claim 3 recites that “the selected media clip is a part of a second media clip”. Even if one of the advertisements of Mackintosh discussed in Col. 16, lines 14-17, could be considered the “selected media clip” or the “second media clip”, there is no disclosure or suggestion that the advertisement is “part of” another media clip, as recited. Again, as discussed above with respect to amended claim 1, Mackintosh provides no indication or suggestion that the additional or supplemental information be another media clip, of which the selected media clip is a part of.

Accordingly, Marcus in view of Mackintosh fails to disclose or suggest “selecting a media clip from the one or more media clips, wherein the selected media clip is a part of a second media clip; and delivering the second media clip to the user” as recited in amended claim 3. As Kikinis, whether alone or in combination with the other cited art, does nothing to resolve these deficiencies, claim 3 is patentable over the prior art of record.

Dependent claims are patentable at least by virtue of their respective dependencies on an allowable claim and/or for further features recited therein. Accordingly, Applicant submits that claims 1 and 3-6 are allowable.

Specifically regarding claims 4-5, Applicant respectfully submits that there would be no reason to modify Marcus in view of Mackintosh in the manner described in the Office Action because the windows in the history bar Mackintosh are specifically used “to bring up information for that selection”, not to play a different media clip in the same “first viewing area”, as recited, and additionally playing the selected media

clip “in the plurality of mini-display regions of the information display”, as also recited. Thus, absent the benefit of Applicants’ own disclosure, there is no suggestion to modify Marcus in view of Mackintosh to arrive at what is claimed.

Further, even if there were a reason to modify Marcus in view of Mackintosh (which Applicant disagrees with, as discussed above), Applicant respectfully submits that Kikinis is nonanalogous art to the presently claimed disclosure. Kikinis is concerned solely with the use of computer screen space for the static presentation and discovery of media assets. Usually, the assets themselves were not displayed, but text and still images of the referenced media. This electronic programming guide (EPG) usage runs contrary to the spirit of the present disclosure.

At the time of original filing of the present disclosure, there was a strong interest in interactive program guides to aid overburdened consumers of media to discover and filter potential offerings of interest. The present disclosure ran contrary to that interest in passive index-style discovery. The present disclosure integrates discovery into the process of active consumption of audio and video streams of content.

The present disclosure enables streams of potential content to be filtered in real- or near-real-time to allow the presentation of potentially relevant content to a user based on their stored profile, viewing/listening history, and/or stated desires as expressed by an active query. This last sub-category of media filtration most closely hews to the intent of Kikinis but is still distinctly different in that search results, in accordance with the present disclosure, are presented at all times as part of an ongoing media stream.

Accordingly, Applicant submits that a person of ordinary skill in the art would not look to EPGs, which are static, in the context of the present disclosure, which improves the search process in dynamic audio and/or video streams for consumer consumption. For these additional reasons, Kikinis fails to resolve the deficiencies of Marcus in view of Mackintosh with respect to claims 4 and 5, and claims 4-5 are allowable.

Id. at 78-81.

111. The examiner issued an office action on September 13, 2023, rejecting all pending claims. *Id.* at 88-106. Regarding the alleged difference between a “media clip” and an “advertisement clip,” the examiner wrote:

Additionally, the applicant argues that a media clip and an advertisement clip refer to different recited clips (remarks page 6). However, it is noted that independent claim 3 does not even use the argued language, claim 3 actually claims that the media clip is part of the second clip, e.g. advertisement. Moreover, the examiner notes that the claims are broad in scope and fail to explicitly and positively recite that these claimed clips are distinct from each other, as argued by the applicant.

In fact, the disclosure of the instant applications teaches otherwise, and clearly defines that the menu options area 220 comprises mini-display regions 230 of the recently displayed clip in viewing area 210; paragraphs [0023] [0024] figure 2. Wherein the user selects information for display in the current viewing area 210 by selecting a clip from one or the mini-display regions 230; paragraph [0026]; similar to the teachings of Mackintosh.

Id. at 92-93.

112. Regarding the analogous nature of Kikinis and the obviousness of simultaneous display of videos and/or thumbnail images, the examiner wrote:

In this case, the examiner notes that Kikinis is related to the analogous art of content distribution, e.g. A/V television content, and how to present the available content to the user (Figure 2). Additionally, the examiner notes that Kikinis was merely introduced to show that it was well-known in the art at the time of the invention to present “playing” mini video windows, such as 205A-X, which not only contain thumbnail pictures but thumbnail video streams of real-time programming for the selected channel; paragraph [0030]. All for the predictable result of reviewing other selectable content available to the user for selection by previewing the video image (Kikinis - paragraph [0030]).

Additionally, nothing in the disclosure of Kikinis would explicitly or implicitly discourage one of ordinary skill in the art from implementing the disclosed teachings of Kikinis along with the disclosure of the Marcus and Mackintosh references. Therefore, Kikinis is analogous to the art and directly relevant to the claimed features, as noted on the instant office action.

Id. at 94.

113. The examiner also issued a variety of double patenting rejections. *Id.* at 94-99.

114. Regarding the disclosures of Marcus¹⁵⁶, the examiner stated:

Regarding claim 1, Marcus discloses a method of delivering one or more media clips to a user (Figure 6), comprising:

delivering the one or more media clips to the user, wherein the one or more media clips are concatenated together (Col. 13 lines 23-45, co. 14 lines 13-34 figures 5 and 6; assembling 615 a concatenated stream of clips for a user based on the viewer information characteristics).

However, it is noted that Marcus is silent to explicitly disclose selecting a media clip from the one or more media clips, wherein the selected media clip is related to a product; and delivering an advisement clip detailing the product to the user.

...

Regarding claim 3, Marcus discloses a method of delivering one or more media clips to a user (Figure 6), comprising:

delivering the one or more media clips to the user, wherein the one or more media clips are concatenated together (Col. 13 lines 23-45, co. 14 lines 13-34 figures 5 and 6; assembling 615 a concatenated stream of clips for a user based on the viewer information characteristics).

However, it is noted that Marcus is silent to explicitly disclose selecting a media clip from the one or more media clips, wherein the selected media clip is a part of a second media clip; and delivering the second media clip to the user (Col. 16 lines 9-37 figure 7; delivering

advertisement to users, e.g. McDonald's advertisement, Sony advertisement, etc.).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Marcus by specifically providing the elements mentioned above, as taught by Mackintosh, for the predictable result of generating profit from delivery of advertisement content to multiple viewers[.]

Regarding claim 6, Marcus and Mackintosh disclose the method of claim 1; moreover, Marcus discloses that each of the one or more media clips is associated with a header (e.g. header 110) including both content information (Col. 3 lines 7-8 figure 2; content tags 120 contain information identifying content of the clips) and aesthetic information for the each of the one or more media clips (Col. 3 lines 34-46, col. 13 lines 45-53, col. 14 lines 25-32; video clip information may include parameters such as aesthetic features).

115. *Id.* at 100-02.

116. Regarding modification of Marcus to display multiple clips simultaneously, the examiner wrote:

Regarding claim 4, Marcus and Mackintosh disclose the method of claim 1; however, it is noted that Marcus is silent to explicitly disclose providing an information display and partitioning the information display to comprise: a first viewing area; and a plurality of mini-display regions; playing, in the first viewing area of the information display, the selected media clip; playing, in the first viewing area of the information display, the advertisement clip; and displaying, in the plurality of mini-display regions of the information display, the selected media clip.

Nevertheless, in a similar field of endeavor Mackintosh discloses providing an information display and partitioning the information display to comprise: a first viewing area (Col. 12 lines 55-63, col. 13 lines 4-15 figure 7 and 12; window 302/526 display content currently being played in a larger screen area);

and a plurality of mini-display regions (Col. 15 line 15- col. 16 line 23 figures and 12; sliding history window 306/525 comprising sub-windows 525a for the five most recently displayed clips in chronological order);

playing, in the first viewing area of the information display, the selected media clip (Col. 12 lines 55-63, col. 13 lines 4-15 figure 7 and 12; display content currently being played; e.g. Fleetwood Mac in window 302/526);

playing, in the first viewing area of the information display, the advertisement clip (Col. 12 lines 55-63, col. 13 lines 4-15 figure 7 and 12; display content currently being played; e.g. McDonalds);

and displaying, in the plurality of mini-display regions of the information display (e.g. sub-windows 525a), the selected media clip (Col. 12 lines 55-63, col. 13 lines 4-15 figure 7 and 12; window 302/526 display content currently being played in a larger screen area, which is later moved to the sliding history window 306/525 comprising sub-windows 525a for the five most recently displayed clips in chronological order).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Marcus by specifically providing the elements mentioned above, as taught by Mackintosh, for the predictable result of presenting multiple clip information to the user in order to allow the user to navigate amongst them.

However, it is noted that Marcus and Mackintosh are silent to explicitly disclose that the displayed plurality of mini-display regions (Col. 15 line 15- col. 16 line 23 figures 7 and 12; sliding history window 306/525) comprises playing, in the plurality of mini-display regions, the second media clip.

Nevertheless, in a similar field of endeavor Kikinis discloses that the displayed plurality of mini-display regions (e.g. image area 205A-X) comprises playing, in the plurality of mini-display regions, the selected media clip (Paragraphs [0027]-[0029] figure 2; displaying window 101 displaying the currently tuned channel while simultaneously displaying individual image area 205A-X. Wherein the individual image areas

205A-X contain thumbnail video streams of real-time programming for the selected channel; paragraph [0030]).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Marcus and Mackintosh by specifically providing the elements mentioned above, as taught by Kikinis, for the predictable result of reviewing other selectable content available to the user for selection (Kikinis - paragraph [0030]).

Id. at 103-05.

117. On February 13, 2024, the applicant submitted a request for continued examination accompanied by claim amendments and remarks. *Id.* at 117-28.

Pending independent claim 1 was amended as follows:

1. (Currently Amended) A method of delivering one or more media clips to a user, comprising:
delivering the one or more media clips to the user, wherein the one or more media clips are concatenated together;
selecting a media clip from the one or more media clips, wherein the selected media clip is **a video or audio clip** related to a product; and
delivering an advertisement clip, **longer than the video or audio clip**, detailing the product to the user.

Id. at 121.

118. Pending independent claim 3 was amended as follows:

3. (Currently Amended) A method of delivering one or more media clips to a user, comprising:
delivering the one or more media clips to the user, wherein the one or more media clips are concatenated together;
selecting a media clip from the one or more media clips, wherein the selected media clip is a **part of a second media clip portion of an entire media program**; and
delivering the **entire media program** ~~second media clip~~ to the user.

119. New claims 9-17 were also added. *Id.* at 123-24.

120. The applicant argued that the newly-added limitations distinguished the combination of Marcus156 and Mackintosh:

Without acquiescing to the propriety of the Office's previous rejections, and solely in the interest of expediting prosecution of the present application, claims 1, 3 and 7 have been amended. Specifically, Applicant respectfully submits that the prior art of record, whether alone or in combination, fails to disclose suggest at least:

selecting a media clip from the one or more media clips, wherein the selected media clip is a video or audio clip related to a product; and

delivering an advertisement clip, longer than the video or audio clip, detailing the product to the user[.]

as recited in amended claim 1 (emphasis added),

selecting a media clip from the one or more media clips, wherein the selected media clip is a portion of an entire media program; and

delivering the entire media program to the user[.]

as recited in amended claim 3 (emphasis added), and

playing, in one of the plurality of mini-display regions of the information display, the selected first media clip in a loop[.]

as recited in amended claim 7 (emphasis added).

Regarding claim 1, Marcus as modified by Mackintosh discloses a sliding history window 306/525 and supplemental information in a main window 302/526, as discussed in Mackintosh at Col. 16, lines 1-37. Applicant respectfully notes that the history window merely displays a “history of advertising or other tracks included in the broadcast material”. See Mackintosh at Col. 16, lines 5-6. In contrast, amended claim 1 recites “wherein the selected media dip is a video or audio clip related to a product”. Thus, Marcus as modified by

Mackintosh is deficient at least because the combination fails to disclose or suggest (1) a “selected media clip” that is “a video or audio clip related to a product” and (2) “delivering an advertisement clip, longer than the video or audio clip”.

Therefore, even if Marcus could be modified with Mackintosh (which Applicant is not conceding), the resulting combination would still fail to disclose or suggest at least: “wherein the selected media clip is a video or audio clip related to a product” and “delivering an advertisement clip, longer than the video or audio clip”, as recited in claim 1, because neither Marcus nor Mackintosh discloses or suggests providing an audio or video clip that is selected which results in an advertisement (which is longer than the selected clip) being delivered. That is, the proposed combination of references based on Marcus fails to show, describe or suggest all claimed features.

Regarding claim 3, Marcus as modified by Mackintosh discloses a sliding history window 306/525 and supplemental information in a main window 302/526, as discussed in Mackintosh at Col. 16, lines 1-37. Applicant respectfully notes that the history window merely displays a “history of advertising or other tracks included in the broadcast material”. See Mackintosh at Col. 16, lines 5-6. In contrast, amended claim 3 recites “wherein the selected media clip is a part of an entire media program”. Thus, Marcus as modified by Mackintosh is deficient at least because the combination fails to disclose or suggest (1) a “selected media clip” that is “a portion of an entire media program” and (2) “delivering the entire media program” to a user.

Therefore, even if Marcus could be modified with Mackintosh (which Applicant is not conceding), the resulting combination would still fail to disclose or suggest at least: “wherein the selected media clip is a portion of an entire media program” and “delivering the entire media program to the user”, as recited in claim 3, because neither Marcus nor Mackintosh discloses or suggests providing a portion of a media program that is selected, which results in an entire media program being delivered to a user. That is, the proposed combination of references based on Marcus fails to show, describe or suggest all claimed features.

Regarding claim 7, Marcus as modified by Mackintosh and Kikinis discloses image areas 205A-X that contain thumbnail video streams, as discussed in Kikinis at Para. [0030]. Applicant respectfully notes that

the thumbnail video streams are of “real-time programming of selected channels”. In contrast, amended claim 7 recites “playing, in one of the plurality of mini-display regions of the information display, the selected first media clip in a loop”. Thus, Marcus as modified by Mackintosh is deficient at least because the combination fails to disclose or suggest playing the selected first media clip in a loop.

Therefore, even if Marcus could be modified with Mackintosh and Kikinis (which Applicant is not conceding), the resulting combination would still fail to disclose or suggest at least: “playing, in one of the plurality of mini-display regions of the information display, the selected first media clip in a loop”, as recited in claim 7, because neither Marcus nor Mackintosh nor Kikinis discloses or suggests playing the selected first media clip in a loop in a mini-display region. That is, the proposed combination of references based on Marcus fails to show, describe or suggest all claimed features.

For at least these reasons, Applicant respectfully submits that independent claims 1, 3 and 7 are patentable. Dependent claims are patentable at least by virtue of their respective dependencies on an allowable claim and/or for further features recited therein. Accordingly, Applicant submits that claims 1 and 3-16 are allowable.

Id. at 126-28. As of July 14, 2024, this was the last filing in the prosecution history of the '374 application.

C. Level of Ordinary Skill in the Art

121. I understand that a person of ordinary skill in the art (“POSITA”) is a hypothetical person who is presumed to be aware of all pertinent art, thinks along conventional wisdom in the art, and is a person of ordinary creativity—not an automaton. In deciding the level of ordinary skill, I understand that the following factors may be considered:

- The levels of education and experience of persons working in the field;

- The types of problems encountered in the field; and
- The sophistication of the technology.

122. I understand that the asserted claims of the asserted patents must be evaluated from the perspective of a POSITA. In my opinion, the relevant art for this patent relates to delivery of video content over a computer network. I understand that the relevant point in time for determining the qualifications of a person of ordinary skill in the state-of-the-art is the time of the alleged invention, which I assume to be the earliest filing date for the patent. Here, I understand that the earliest invention date contended by Patent Owner is June 24, 2003.

123. In my opinion, a POSITA at the time of the alleged invention would have had at least a Bachelor of Science in electrical engineering, computer engineering, or similar fields and at least two years of practical experience in the field of media processing and/or transmission.

124. This level of skill is approximate, and more experience would compensate for less formal education, and vice versa. I meet these criteria now and met them at the time of the alleged invention. I have applied this level of skill in my analysis. My opinions would not change if a slightly higher or lower level of ordinary skill applied.

D. Claim Construction

125. I have been instructed by Petitioner to perform my technical analysis of the disclosures of the prior art by applying the plain and ordinary meaning of all claim terms. I have not been asked to opine on the correctness of any claim constructions.

126. I reserve the right to provide additional opinions concerning claim construction or the application of certain claim constructions to the prior art, as appropriate, and to respond to any particular claim construction-related argument advanced by PO and/or its expert.

127. As described in more detail in the remainder of this declaration, the prior art discloses or renders obvious the challenged claims under any reasonable potential claim interpretation.

E. Challenged Claims

128. I understand that claims 1-20 are at issue in this petition for *inter partes* review. They are reproduced below for reference.

Claim	Limitation
1[pre]	1. A method of providing video programming to a user, comprising the steps of:
1[a]	searching, by one or more computing devices, one or more databases of video clips in response to a user request wherein each video clip comprises a plurality of frames,
1[b]	selecting, by one or more of said computing devices, a plurality of first video clips responsive to the user request, wherein each

Claim	Limitation
	first video clip is associated with media content comprising one or more second video clips;
1[c]	selecting, by one or more of said computing devices, a plurality of first transitional clips for association with at least some of the selected first video clips;
1[d]	determining a first temporal sequence for concatenating the selected first video clips and first transitional clips;
1[e]	wherein said selecting, determining and concatenating is performed according to one or more first templates, to define first video programming, wherein said one or more first templates comprise a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and
1[f]	presenting said first video programming in a first area of a display associated with one of said computing devices;
1[g]	concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display;
1[h]	receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;
1[i]	selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection;
1[j]	selecting, by one or more of said computing devices, a plurality of video clips including the second video clip, and determining a second temporal sequence for concatenating the selected plurality of second video clips, wherein said selecting, determining, and concatenating is performed in accordance with one or more first or second templates, to define second video programming;

Claim	Limitation
1[k]	wherein said one or more first or second templates comprises a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and
1[l]	presenting said second video programming on a display associated with one or more of said computing devices.
2	A method according to claim 1, wherein at least one of said video clips of said first or second video programming is a transitional clip.
3	A method according to claim 2, wherein at least one of said video clips of each of said first and second video programming is a transitional clip.
4	A method according to claim 3, wherein said transitional clip is a hosting clip consistent with at least one template variable.
5	A method according to claim 2, wherein at least one transitional clip is a hosting clip consistent with at least one template variable.
6	A method according to claim 2, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.
7	A method according to claim 6, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.
8	A method according to claim 6, wherein one or more video clips comprise advertising.
9	A method according to claim 6, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.

Claim	Limitation
10	A method according to claim 6, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.
11	A method according to claim 6, wherein said stored associations between said video clips comprise a plurality of media tags.
12	A method according to claim 6, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.
13	A method according to claim 1, wherein said first and second video programming are each presented as a stream of information to the user.
14	A method according to claim 1, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.
15	A method according to claim 14, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.
16	A method according to claim 1, wherein one or more video clips comprise advertising.
17	A method according to claim 1, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.
18	A method according to claim 1, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.

Claim	Limitation
19	A method according to claim 1, wherein said stored associations between said video clips comprise a plurality of media tags.
20	A method according to claim 1, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.

IV. The Prior Art

129. All of the underlying technologies of the '185 patent claims were well-known to a POSITA. I briefly describe here some of the relevant technologies.

A. State of the Art

130. Since the earliest days of radio and television broadcasting, media providers have sought to understand their audience and cater to their preferences. Nielsen Media Research—the precursor to today's Nielsen TV Ratings—started publishing its Nielsen Radio Index in 1942. In the 1950s, Nielsen moved to television and began installing “Adimeters” in sample households to monitor channels being watched and introduced the Nielsen Station Index service, which included reports on the audience of any given show, network, and programming hour. Media providers then competed with each other to develop programming that captured the largest audience share, allowing providers to charge more for commercials airing with programming that had a larger audience share. In turn,

advertisers began selecting specific programming for their commercials based on market share of their target audience.

131. In the 1970s, television programming became further customizable via the advent of local programming insertion. Cable providers would dedicate specific slots within their national programming for local affiliates to insert their own, custom news segments, messages, or commercials, as content was transmitted from a centralized, nationwide “headend,” to a more local market headend or cable hub (where localized programming insertion occurs), to individual subscribers. Thus, media programming became customizable on a geographical basis, catering to the preferences of the local audience.

132. Even with local programming insertion, the broadcast nature of television continued to be a limiting factor to the granularity with which television programming could be customized. Traditionally, analog television was broadcast over the air or via cable to a group of households (e.g., all households served by a particular market center headend or cable hub), and at a minimum, that group of households would all receive the same programming. Thus, there historically was a minimum addressable audience size—media providers were generally unable to target programming to a single, specific household.

133. This began to change in the 1990s with the introduction of electronic program guides (EPGs) capable of receiving data on available programming and

controlling channel tuning. A user was able to browse a menu of available television programming and use a remote control to select a particular program and tune to the same. In addition to such manual control, EPGs opened the door to automatically customizing the user experience based on locally-collected user data. For example, U.S. Patent No. 6,177,931 (filed in 1998) describes an EPG that monitors user interactions with the EPG to generate viewer profile information, and then uses that information to customize the presentation of advertising to the user, e.g., by automatically tuning to a particular advertising channel during a commercial break. As another example, U.S. Patent No. 8,813,122 (filed in 1999) describes generating a “virtual channel” of programming specifically selected based on the user’s profile and then automatically tuning to that programming according to a defined schedule.

134. The next step in the evolution of custom media programming was brought by advances in computer technology. By the early- to mid-1990s, most components of the modern Internet were in place—including web servers that support HTTP over TCP/IP traffic—and codecs for transmitting video over computer networks (such as MPEG) had also been standardized. Indeed, the first movie was broadcast over the Internet in May 1993, the first live performance in June 1993, and the first episodic television show in September 1997. Importantly, unlike prior broadcast technology, TCP/IP permitted a web server to receive a

request from a particular user and respond to that particular user's request with media programming—including media programming customized for that particular user.

135. One limiting factor was that most households in the 1990s did not have sufficient Internet speeds to support video streaming. Standard definition video requires about 0.5 – 1.5 Mbps. In the 1990s, only 56 kbps dial-up was available to the vast majority of U.S. households with Internet access—far below the speed necessary to stream television-quality video. This began to change with the introduction of broadband Internet connections. By 2000, about 7 million U.S. households had broadband Internet subscriptions. This rendered transmitting video programming over the Internet commercially desirable, leading to a rapid rise in Internet-based media companies from 2000 onwards.

B. Marcus156 (Ex. 1009)

136. U.S. Patent No. 6,032,156 (“Marcus156”) was filed on April 1, 1998 and issued on February 29, 2000. I therefore understand that Marcus156 is prior art to the '185 patent under at least pre-AIA 35 U.S.C. § 102(b).

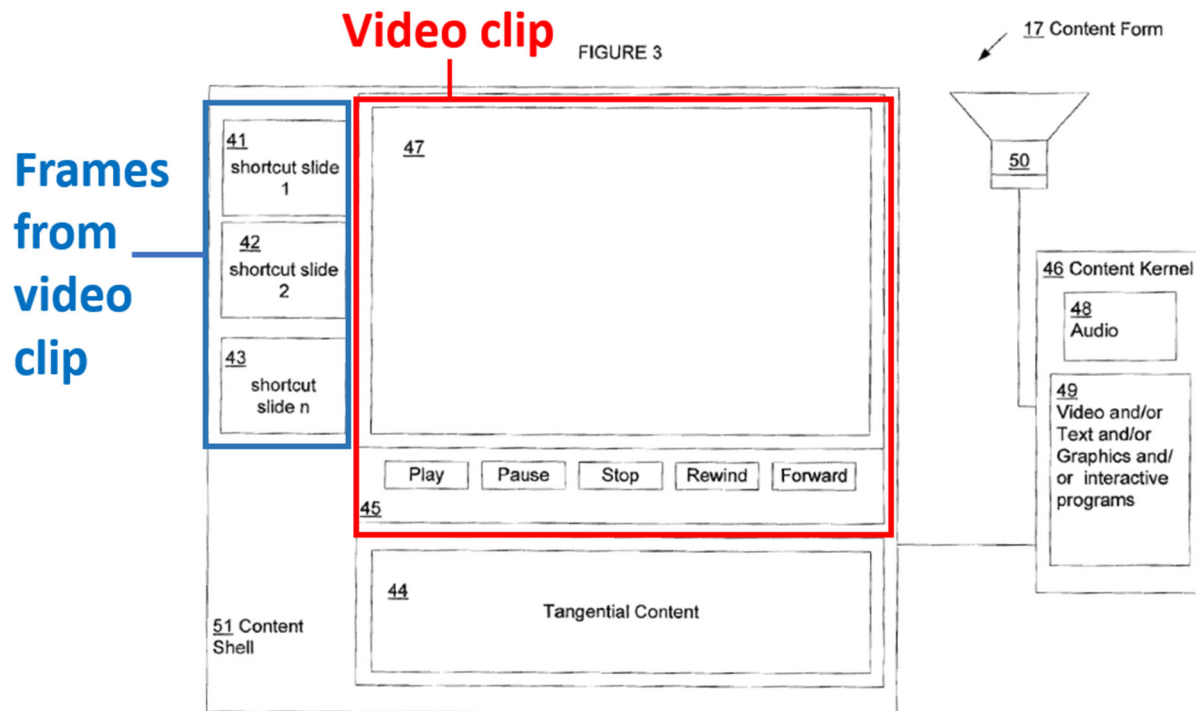
137. Marcus156 names the same inventor, Dwight Marcus, as the '185 patent, and relates to the same technology described and claimed in the '185 patent. Marcus156 describes a video system where templates are used to select video clips, find associated transition clips, and concatenate them together into a stream of video programming. *See, e.g.*, Marcus156 at Abstract (“Media elements are selected in

response to a request for media programming, and arranged in a temporal organization. . . . Transitions between audiovisual clips are determined by the system based on information stored in the database and predetermined preferences as to types of transitions. . . . Upon completion of the selection of the media elements, the sequence, and the transitions, the media elements are assembled into a media program”).

C. Rosen (Ex. 1011)

138. Rosen is a patent application publication that was filed on November 13, 2001 and published on October 14, 2004. I therefore understand that Rosen is prior art to the '185 patent under at least pre-AIA 35 U.S.C. § 102(e).

139. Rosen is directed to an enhanced presentation system for displaying video alongside slides and “tangential content.” Rosen at Abstract; [0001] (“The invention relates generally to multimedia presentations. Particularly, this invention relates to a method of generating and playing a multimedia content presentation in which a video is shown and tangential content is presented at preset points in the video.”). As depicted in Figure 3, a video is “playing in a main window 47”:



Id. at [0023]; Fig. 3 (annotated). Shortcut slides 41-43 are displayed next to the playing video. *Id.*; *see also id.* at [0025] (“It is to be understood that there may be any number of shortcut slides, and they may be structured in various graphical ways in the content shell 51.”). The shortcut slides contain frames of the playing video. *Id.* at [0027]. The user may select a shortcut slide to initiate an action. *Id.* at [0023], [0027].

D. Sull (Ex. 1010)

140. Sull is a patent application publication that was filed on July 23, 2001 and was published on June 6, 2002. Therefore, I understand that Sull is prior art to the '185 patent under at least pre-AIA 35 U.S.C. § 102(b).

141. Sull is directed to a system for “tagging, indexing, searching, retrieving, manipulating, and editing video images on a wide area network such as the Internet.” Sull at Abstract. In particular, Sull describes “multimedia bookmarks” that point to a particular location in a video and are displayed as a frame. Sull at Fig. 9; [0205]; [0220]. A user may select one of the multimedia bookmark frames to trigger a search for compatible video clips. *Id.* at [0225]; [0221]; [0024]; [0060].

E. Hosken (Ex. 1015)

142. Hosken is a PCT patent application publication that was filed on July 14, 2000 and was published on January 25, 2001. I understand that Hosken is therefore prior art to the '185 patent under at least pre-AIA 35 U.S.C. § 102(b).

143. Hosken is directed to recommendation algorithms for “automated generation of source content alternatives based on content relationships and user behavioral patterns to support the recommendation of alternative content source.” Hosken at 1:10-14. Hosken notes that “a user’s appreciation of particular content items can suggest the user’s likely appreciation of other content titles.” *Id.* at 2:9-11. Therefore, items from a first search result set can be selected by the user and used to perform successive, more refined searches. *Id.* at 9:4-8 (“Based on the review and consideration of the recommendation set, including as applicable any presented content samples, the user may request a further search and production of a new recommendation set, typically identifying a prior recommendation media

content item as part of the request, or request the purchase and delivery of a media content item.”).

F. Byers (Ex. 1018)

144. Byers is an issued United States Patent that was filed on April 26, 2000 and issued on February 19, 2008. I therefore understand that Byers is prior art to the ’185 patent under at least pre-AIA 35 U.S.C. § 102(e).

145. Byers is directed to “dynamically altering a portion of a digital video image based upon a user profile.” Byers at Abstract. In particular, targeted advertising can be dynamically inserted into a video stream. *Id.* at 2:48-52 (“[T]argeted product placement refers to placing products into scenes from a play, television show, or movie. Such placement can lead to a boon in sales and publicity for the products so advertised.”).

G. Ferman (Ex. 1012)

146. Ferman is a U.S. Patent Publication published on May 16, 2002. I therefore understand that Ferman is prior art to the ’185 patent under at least pre-AIA 35 U.S.C. § 102(b).

147. Ferman is directed to a system for “managing audiovisual information,” and in particular for presenting that information in a way that is suited to the particular viewing preferences of the users. Ferman at [0001], [0089]. The system described in Ferman “records and presents to the user audio and video information

based upon the user's prior viewing and listening habits, preferences, and personal characteristics.” *Id.* at [0091]. Ferman uses a robust system of metadata templates to match user preferences with a database of video clips to create dynamic, interactive video programming. *Id.* at [0093] (“By providing a separate description scheme describing the program(s) 10, the user 14, and the system 12, the three portions (program, user, and system) may be combined together to provide an interactivity not previously achievable.”).

V. Ground 1: Marcus156 Renders Claims 1-20 Obvious

148. In my opinion, Marcus156 alone renders challenged claims 1-20 obvious. Unless otherwise noted below, it is my opinion Marcus156 discloses each claim and claim element.

A. Claim 1

1. Element 1[pre]: “1. A method of providing video programming to a user, comprising the steps of:”

149. In my opinion, Marcus156 discloses this element.

150. *A method of providing video programming to the user:* Marcus156 teaches that its techniques are used for “delivering information via customized video programming.” Marcus156 at 10:23-27. Marcus156’s abstract also discusses “creating audiovisual programming [with] media elements, such as audiovisual clips.” *Id.* at Abstract; *see also id.* at 2:10-11 (“The media elements are then assembled into media programming.”). The programming is provided to a user as

“interactive generation of programming.” *Id.* at 14:46-58. The additional steps are described below.

2. **Element 1[a]: “searching, by one or more computing devices, one or more databases of video clips in response to a user request wherein each video clip comprises a plurality of frames,”**

151. In my opinion, Marcus156 discloses this element.

152. *Searching, by one or more computing devices, one or more databases of video clips in response to a user request:* Marcus156 discloses a database of video clips. Marcus156 at 1:63-65 (“A database is provided which contains selected information about each of a large number of media elements.”). The database includes various types of data about video clips, as shown in Figure 2:

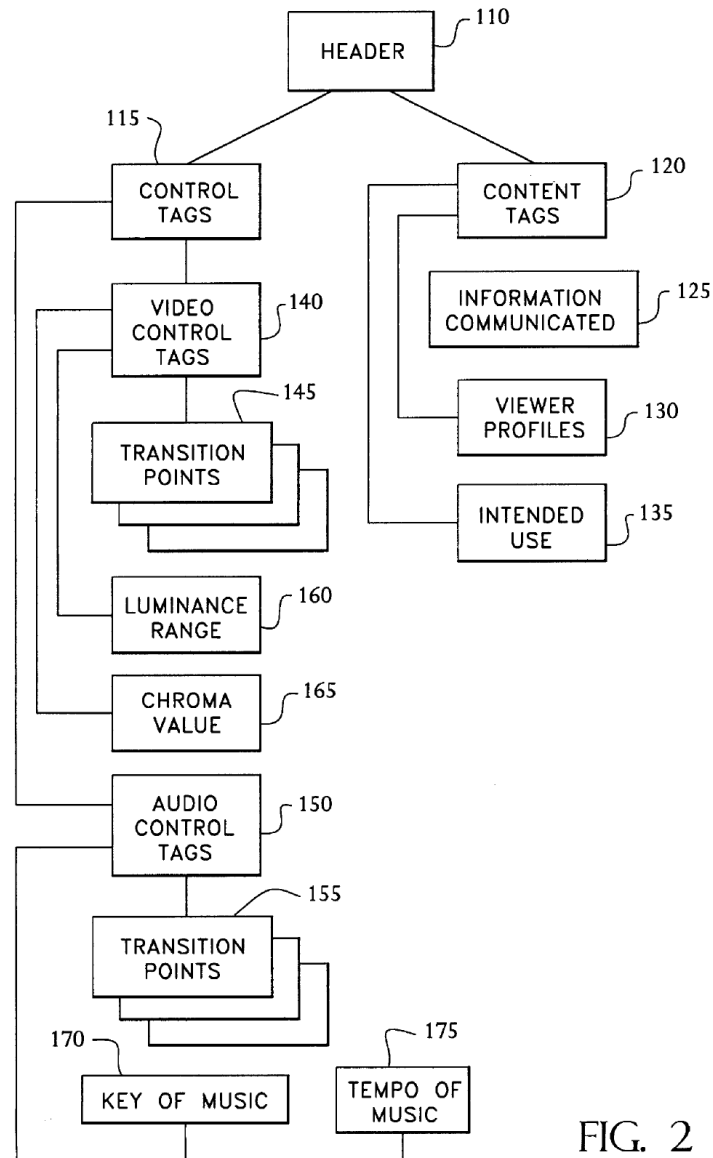


FIG. 2

Id. at Fig 2; 2:35-37 (“FIG. 2 is a schematic diagram indicating information contained in a database according to the invention.”); 2:67-3:2 (“Database 100 contains unique identifying information for each clip and has associated therewith additional information often arranged in a hierarchical manner.”). Among other things, the database contains media tags attached to video clips. *Id.* at 5:40-43 (“The database is created by identifying each clip or other asset and defining values of the

control tags and content tags for each.”); cl. 1 (“maintaining a database containing selected information about each of a plurality of media elements”).

153. Thus, it is my opinion that Marcus156’s database 100 is a claimed “database of video clips.” *See also* Ex. 1003 at 81 (database need not store binary data for video files: “What is most significant is that we propose that accessible to our command-interpreter program, we maintain a database of links to the requested underlying media of that those links are inherent in the tagging of the viewed media and are thus passed to the server.”), 28 (“the source of the full-length media (which might be remote to the database (or not) and may in fact be an entirely different broadcast network or channel or URL address probably fitted with VOD capabilities”).

154. Marcus156 also describes a library 70 which stores the large underlying video data for the video clips. *Id.* at 2:56-58 (“Computer 20 communicates in a suitable manner with library 70 of audiovisual clips 80.”); 2:61-64 (“Digitized audiovisual clips 80 are stored in library 70. Each audiovisual clip is suitably identified, such as by the name of one or more files which make up the clip.”). Additionally, Marcus156 describes a “database of viewer information” which “defines the identity of each viewer, including name and address information, and may also include social, economic and medical data regarding the viewer, the history of items viewed by the viewer, and a file of preference information for the viewer.”

Marcus156 at 10:10-16. The system described in Marcus156 uses this “database of viewer information” along with the clip database and the library to select and assemble the video programming. Therefore, it is my opinion that database 100, along with at least library 70 and the “database of viewer information,” constitute the claimed “one or more databases of video clips.”

155. However, to the extent that the scope of the claimed “database” is ultimately limited to Marcus156’s database 100, it is my opinion that Marcus156 would still disclose all relevant database-related claim elements. As detailed below, database 100 is searched in response to a user request as required by limitation 1[a], and the claim term “database” only occurs in limitation 1[a], not in any other limitation of claim 1 or in any other claim of the ’185 patent.

156. Additionally, whether the library 70 is a “database” is at least an implementation detail or design choice, where a POSITA would have been motivated to make obvious changes to the data architecture of the system with a reasonable expectation of success. As one example, a POSITA would have been motivated to integrate the raw video data in library 70 into database 100. In my opinion, this would have improved the system by eliminating the need to separately access library 70 after selecting video clips based on database 100. As another example, a POSITA would have been motivated to implement library 70 as a database. Doing this would have speeded up the lookup and retrieval of video data

from library 70. I note that Marcus156 discloses storage of the video clips in the library. A POSITA reading Marcus 156 would understand that there would need to be a way of implementing the storage library. There would be a finite set of choices a POSITA would consider on how to implement the storage functionality. One of most likely options would be to use a database. The advantages and disadvantages of using a database would have been known to a POSITA at the time of the invention. Therefore, a POSITA would have found that using a database for the storage library would have been an obvious design alternative to try, and could have implemented that solution with a reasonable likelihood of success. Further supporting the reasonable expectation of success, Marcus156 already describes the use of databases to store various types of information and provides significantly more detail than the '185 patent itself as to their use.

157. *Searching . . . in response to a user request:* The databases are searched in response to a user request. *Id.* at 1:67-2:2 (“The method provides for selecting some of those media elements *in response to a request for media programming . . .*”) (emphasis added); 2:25-28 (“In a first step, from a database containing information concerning the audiovisual media elements, certain audiovisual media elements are selected.”); 4:60-61 (“[T]he database is mined for suitable assets.”); 14:33-34 (“[T]he assets may be selected based on demographic search criteria.”). Marcus156 also refers to the user request as a “query.” *Id.* at 9:9-

11 (“[T]he user creates queries depending on the qualities desired for the organization levels to be identified.”); 13:51-52 (“The user can select assets through a suitable query without attention to temporal organization.”). The user may be a video editor using Marcus156’s “List Editor/Previewer” (7:10-15) or an end-user using Marcus156’s “interactive system” (14:35-67).

158. For reference, the appendix to the ’597 provisional also refers to user requests as “queries.” *E.g.*, Ex. 1003 at 25 (“One model’s technical focus is the creation of a mediating layer that lies between the database of media assets and the viewer. This mediating layer insulates the consumer from the tedium of sorting through media elements that match his/her query and simply presents the results of a user’s query in a sophisticated, entertaining way.”), 60 (“Again it’s important to note that, although this process was developed to act as a front-end for the delivery of short video clips in response to a query, it will also function well for delivering text and image results to a user in a traditional manner.”).

159. ***Searching . . . by one or more computing devices:*** The searching of the database is performed “by one or more computing devices,” namely computer 20. Marcus156 states that “Computer 20 is also suitably associated with a database 100.” Marcus156 at 2:66-67. One way that the search may be performed is by an editing program, which may include an automatic expert system, running on computer 20. Marcus156 at 4:52-5:2; 9:11-12 (“The editor program then identifies

suitable templates, modules or sequences based on the queries.”); 8:13-29 (“[A]n expert system included within the editor program selects and orders one or more suitable templates [U]sing the requirements in the selected modules and the inherent structural characteristics of sequences, the expert system or truth table selects and orders suitable sequences”). Furthermore, Marcus156’s claim 1 specifies that the searching is done automatically—i.e., by a computing device according to data and algorithms. *Id.* at cl. 1 (“automatically selecting a plurality of said media elements in response to a request for media programming”).

160. ***Wherein each video clip comprises a plurality of frames:*** The video clips described in Marcus156 are, like any video, comprised of a plurality of frames. *See, e.g.*, Marcus156 at 9:48-51 (“[C]ompression formats, of which MPEG is an example, which achieve performance using predictive compression which includes both forward and backward looking frames.”); 2:61-62 (“Digitized audiovisual clips 80 are stored in library 70.”). For example, Marcus156 discusses a type of transition that is measured in a particular amount of frames. *Id.* at 3:39-41 (“24-frame dissolve”). As another example, when it discusses adding transitional clips to a video sequence, Marcus156 describes working with the transitions on a frame-by-frame basis. *Id.* at 7:54-57 (“The first frame of each segment is decoded into separate buffers and then the transition function is applied to each matching pixel of both images to produce the third rendered frame buffer.”).

3. **Element 1[b]: “selecting, by one or more of said computing devices, a plurality of first video clips responsive to the user request, wherein each first video clip is associated with media content comprising one or more second video clips;”**

161. In my opinion, Marcus156 discloses this element.

162. *Selecting, by one or more of said computing devices, a plurality of first video clips responsive to the user request:* The system described in Marcus156 (*i.e.*, computer 20) selects a set of video clips that satisfy the user request. *See, e.g.*, Marcus156 at 2:25-28 (“In a first step, from a database containing information concerning the audiovisual media elements, certain audiovisual media elements are selected.”); 2:6-7 (“The system selects and orders the media elements according to the data in the request”); 1:67-2:2 (“The method provides for selecting some of those media elements in response to a request for media programming”); 4:60-61 (“[T]he database is mined for suitable assets.”); 9:9-11 (“[T]he user creates queries depending on the qualities desired for the organization levels to be identified.”); 13:51-52 (“The user can select assets through a suitable query without attention to temporal organization.”); 14:33-34 (“[T]he assets may be selected based on demographic search criteria.”); 8:29-32 (“[U]sing the requirements in the selected sequences and the inherent structural characteristics of the clips, the expert system or truth table selects and orders suitable clips”); cl. 1 (“automatically selecting a plurality of said media elements in response to a request for media programming”).

163. *Wherein each first video clip is associated with media content comprising one or more second video clips:* Marcus describes using metadata, including tags and templates, to associate each video clip in the database with other video clips based on numerous potential criteria. Therefore, it is my opinion that any selected video clip (“first video clip”) is associated with other video clips (“media content comprising one or more second video clips”) because they, for example, share the same or similar tags, or are otherwise identified as being similar on a particular metric according to templates. *Id.* at 3:25-30 (“One or more fields may be provided within the content tag to define content. The system by which the content information is defined is preferably organized to provide for relationships among type of information. Such relationships may be levels or generality, and other subject matter relationships.”); 3:13-14 (“The content tags contain information defined by a suitable system of classification of information.”). Thus, not only are the clips associated by similar tags and template, Marcus¹⁵⁶’s system treats that interrelationship as an association, and uses it to groups clips in its internal data structures.

164. For example, the associations may be hierarchical, and cause the video clips to be sorted into various high-level and low-level categories (and thus associated). *Id.* at 2:67-3:2 (“Database 100 contains unique identifying information for each clip and has associated therewith additional information often arranged in a

hierarchical manner.”); 3:31-33 (“The organizational structure of the database may be hierarchical, with each layer of hierarchy defining a specific set of organizational principles.”); 3:60-65 (“Here are defined the type of assets and the core content descriptions of those assets. By way of example, the types of assets may be defined as training, informational, and entertainment assets. Examples of core subject matter would be ‘medical’, at the highest level, ‘health management’, at a lower level, and ‘diabetes mellitus’, at a still lower level.”). A POSITA would have recognized that, in the diabetes example, all of the video clips in the database which are tagged “diabetes mellitus” are associated with each because they share the common tag. They are also associated (at a different level of abstraction or hierarchy) with other video clips that are tagged as “health management.”

165. In another relevant section, Marcus¹⁵⁶ describes a video editor configuring the system to prioritize certain types of associations between video clips, which will cause certain (but not all) commonly-associated clips to become part of the same video stream. *Id.* at 4:9-14 (“[T]he instructional designer might define that the block of programming content called EFFECTS ON THE HEART is presented across three fundamental age groups, two levels of detail (summary/cursory and in-depth), both gender specific groups and four distinct ethnicity components.”); 8:3-5 (“The [identity] and order of the clips that comprise a program may be defined by a

variety of methods varying in the degree of control exercised by the individual user.”).

4. Element 1[c]: “selecting, by one or more of said computing devices, a plurality of first transitional clips for association with at least some of the selected first video clips;”

166. In my opinion, Marcus156 discloses this element.

167. *Selecting, by one or more of said computing devices, a plurality of first transitional clips for association with at least some of the selected first video clips:* A POSITA would have considered modifying an existing clip to make a transition to be a selection of a transitional clip. The appendix to the ’597 provisional presents modification of existing video clips (as disclosed in Marcus156) as an example of selecting. *E.g.*, Ex. 1003 at 34 (“types of transitions (hard cut, fade-to-black, white-burst, and so-on)”). Similarly, Marcus156 also considers modification of clips to be the selections of a transitional clip. *E.g.*, Marcus156 at 8:33-35 (“The expert system or truth table also makes *transition selections* at the beginning and end of the audio and video portions of each clip”) (emphasis added). In this scenario, each modified portion of a video clip becomes one of the claimed “transitional clips.” Thus, it is my opinion that modifying existing clips qualifies as “selecting . . . a plurality of first transitional clips.”

168. Once a set of video clips has been selected, Marcus156’s system selects an accompanying set of transitional clips. Marcus156 at 2:29-31 (“Transitions

between the media elements are automatically selected.”); 14:13-17 (“The system is also capable of arranging assets in an order, and creating transitions between assets, to create a concatenated stream of audiovisual programming that has transitions that appear to the viewer to have been edited by a human editor.”); 14:43-45 (“The system will then generate suitable transitions and add those assets to the programming.”). The transitions can comprise free-standing video clips, or they can be created by modifying the start or end of other video clips. *E.g., id.* at 8:33-35 (“The expert system or truth table also makes transition selections at the beginning and end of the audio and video portions of each clip, as shown by block 525 of FIG. 5.”).

169. Marcus¹⁵⁶ also discloses transitions created by inserting a distinct clip (a “bridge” or “floater”) between two other clips (for example, inserting a percussion clip between the two clips lacking percussion) rather than by modifying the surrounding clips. *Id.* at 8:51-56 (“[T]he system may identify a suitable type of transition for both clips, add a bridge between the clips, or determine that one of the clips must be discarded. For example, if the keys of the music in two adjacent clips are incompatible, the system may add a bridge consisting of a burst of percussion in the audio portion of the program.”); 4:28-29 (“elements which are free to float temporally in order to accomplish certain transitions”); 4:34-38 (“The free floating elements may have various linking geometries or parameters at the opening and

closing thereof. Such elements can be used to bridge elements that cannot themselves be linked because a direct link is either disallowed or would involve use of disfavored transitions.”).

170. Marcus156’s transitions may be selected either before or after the initial clips have been selected:

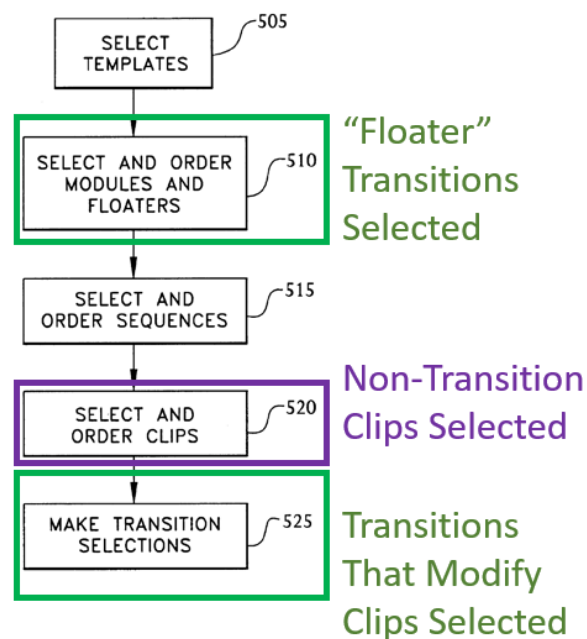


FIG. 5

Id. at Fig. 5 (annotated); 2:42-43 (“FIG. 5 is a flow chart showing steps in the assembly of an edit decision list in the method according to the invention.”); 7:3-7 (“At first, a sample order of clips, without transitions, is analyzed. The order and identities of the clips is then revised. When the order and identities of the clips has been finalized, the transitions are computed.”).

171. *Selecting . . . by one or more of said computing devices:* The transitions are selected “by one or more of said computing devices,” namely computer 20. Marcus¹⁵⁶ at 14:18-21 (“[T]he system contains default transition choices, to permit assembly of programming with fluid transitions without the intervention of a human editor.”); 2:29-31 (“Transitions between the media elements are automatically selected.”); 9:40-41 (“The assembly program may also dictate particular transitions.”). As with the selection of video clips in response to a user query, editing software running on a computing device, and potentially also using an expert system, selects transitions. *Id.* at 8:36-38 (“[T]he expert system or truth table may discard an element and replace it with a substitute element, or employ suitable transition methods.”); 8:48-51 (“The editor program may apply expert system software or truth tables to determine whether a direct transition between any two clips meets aesthetic requirements.”).

172. *Selecting . . . for association with at least some of the selected first video clips:* In my opinion, the transitions are associated with at least some of the selected first video clips. In particular, each transition is associated in the system with the other clips that it transitions between. *Id.* at 7:47-49 (“The transition process will require the identification of both an exit and lead-in point, and a transition function specification.”). Transitions may be selected and associated with other clips based on particular aesthetic requirements derived from the content and format

of the clips that are being transitioned between. *Id.* at 8:48-51 (“The editor program may apply expert system software or truth tables to determine whether a direct transition between any two clips meets aesthetic requirements.”); 3:36-39 (“A given stylistic approach 300 can be defined through the use of default values which will be employed, for example, in the treatment of transitions between adjacent clips.”). In my opinion, transitions like these that share the stylistic or aesthetic approach of the surrounding clips (and this commonality is also tracked by the system as a whole) are “associated with” those clips, as claimed.

173. In my opinion, the transitions are also associated with selected clips because they meet the system’s transition criteria or share tags. Marcus156 at 4:26-38; 4:48-51. Moreover, a POSITA would have recognized that the transitions are also associated with at least some of the selected first video clips because they meet the transition criteria that the system determines for each clip. For example, Marcus156 discusses how video clips may have control tags for “preferred or required transitions, including length and type of transition, a dominant chroma value for each of the opening and closing of the clip, start and end points for dialog and action.” *Id.* at 5:27-33. In my opinion, transitions that fulfill the criteria for a preferred or required transition of another video clip are “associated with” that clip, as claimed.

5. Element 1[d]: “determining a first temporal sequence for concatenating the selected first video clips and first transitional clips;”

174. In my opinion, Marcus156 discloses this element.

175. *Determining a first temporal sequence for concatenating the selected first video clips and first transitional clips:* Marcus156 discloses that its system determines a temporal order for the various clips and the accompanying transitions between them. *Id.* at 1:67-2:3 (“The method provides for selecting some of those media elements in response to a request for media programming, and selecting a temporal organization for the media elements.”); 2:28-31 (“A temporal sequence for the selected elements is designated. Transitions between the media elements are automatically selected.”); cl. 1 (“automatically selecting a temporal organization for said selected media elements”).

176. Marcus156’s system includes a MODULE LAYER or SEQUENCE LAYER that considers numerous factors in deciding how to temporally arrange the video programming. *Id.* at 4:22-26 (“Below the instructional design layer are the smaller organizational elements which allow for elasticity in the specifics of the implementation of the temporal design. This is called the MODULE LAYER 320 and in special instances the SEQUENCE LAYER.”). This layer considers factors that affect how clips and transitions are ordered relative to each other. *Id.* at 4:26-34 (“Fundamental to this layer are weighting factors which control likelihood of

asset use, and allow for the deployment of elements which are free to float temporally in order to accomplish certain transitions and effective deployment of those elements which are slave to the temporality functions.”). The control tags associated with the clips may dictate a particular temporal structure to the video programming. *Id.* at cl. 8 (“selecting a temporal order for said two elements, and determining based on information in said control tags whether said two elements may be assembled in the selected temporal order”).

177. ***For concatenating:*** Marcus156 specifically describes the assembling of clips and transitions into a sequenced final video as “concatenating.” Marcus156 at 14:13-17 (“The system is also capable of arranging assets in an order, and creating transitions between assets, to create a concatenated stream of audiovisual programming that has transitions that appear to the viewer to have been edited by a human editor.”); 3:44-46 (“Numerous other aesthetic or programmatic choice that affect the delivery and placement of media elements in the concatenated stream can be defined.”). The final video stream is concatenated from component parts. *E.g.*, *id.* at 7:5-9 (“When the order and identities of the clips has been finalized, the transitions are computed. The transitions are then executed and inserted into the stream by use of a system of removal of portions of the tag extremities used in temporal transitions.”); 9:23-32 (“Once the set of clips has been defined, the

assembly of the clips from the library takes place. . . . The assembly program uses the transition information in assembling the clips into the program.”).

178. In my opinion, claim 1 does not require selecting appropriate transitions before determining a temporal sequence, and hence the order of operations in Marcus156 discloses the claim language.

179. Notably, a POSITA would have recognized that nothing in the nature of these method steps requires a particular order. Workable video recommendation systems can use a variety of approaches and times to sequencing. The selection of an appropriate transition can be facilitated by knowing the order of the other clips first. Moreover, a POSITA would discern nothing in the intrinsic evidence of the ’185 patent that requires the selection of transitional clips before the sequence order for the video programming is determined. I also note that claim 1 of the ’185 does contain an express timing requirement elsewhere, where it specifies “concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips.” This other claim language would have indicated to a POSITA that no order is required by these steps, because the applicant knew how to require a strict order when it wanted.

180. Moreover, Marcus156 discloses the claimed steps even if they must be performed in order. Floater/bridge transitions are selected before a sequence is determined. Marcus156 at Fig. 5. And in Marcus156, additional sequencing is

performed after transitional clips are selected. For example, even after the sequence of video clips is determined, if a transitional clip is selected to bridge two other clips, then the subsequent step of inserting that transition clip into the video programming further sequences that transition clip relative to the other video clips.

6. **Element 1[e]: “wherein said selecting, determining and concatenating is performed according to one or more first templates, to define first video programming, wherein said one or more first templates comprise a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and”**

181. In my opinion, Marcus¹⁵⁶ discloses this element.

182. *Wherein said selecting, determining and concatenating is performed . . . to define first video programming:* An assembly program creates a stream of video from the selected video clips and transitional clips, which in my opinion “defin[es] first video programming.” *Id.* at 9:22-39 (“Assembly Program. Once the set of clips has been defined, the assembly of the clips from the library takes place. This is accomplished by an assembly program. Referring to FIG. 6, the assembly program reads a file of clip identifiers and transition information created by the editor program as shown by block 605, and calls to the library for each identified clip as shown by block 610. The identified clips are then transmitted from the library to a suitable destination. The assembly program uses the transition information in assembling the clips into the program.”); 14:13-15 (“The system is

also capable of arranging assets in an order, and creating transitions between assets, to create a concatenated stream of audiovisual programming”); 14:7-11 (“The system also provides for the automatic creation of new programming, uniquely configured to the requirements of the viewer, but which may be viewed passively.”).

183. *Wherein said selecting, determining and concatenating is performed according to one or more first templates:* Marcus156’s system includes a “template layer” that is involved in and consulted during the selecting, determining and concatenating steps. Thus, in my opinion, these steps are “performed according to” the templates, as claimed. Figure 3 depicts the template layer:

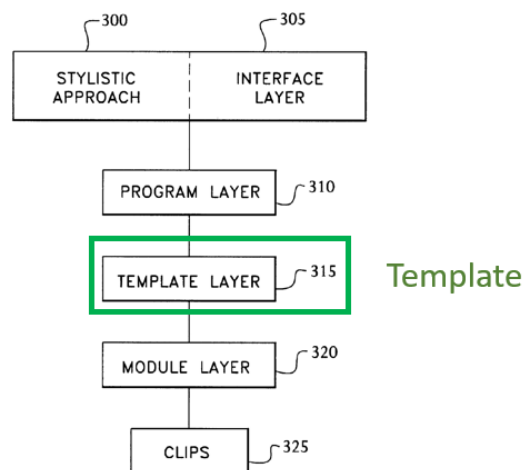


FIG. 3

Id. at Fig. 3 (annotated); 8:13-15 (“[A]n expert system included within the editor program selects and orders one or more suitable templates”); 9:9-13 (“[T]he

user creates queries depending on the qualities desired for the organization levels to be identified. The editor program then identifies suitable templates, modules or sequences based on the queries.”). For example, Marcus156 describes the selection of transitional clips as being performed according to one or more templates. Marcus156 at Abstract (“Transitions between audiovisual clips are determined by the system based on information stored in the database and predetermined preferences as to types of transitions.”). Marcus156 also describes selecting clips according to “weights” present in a template. *Id.* at 9:11-21 (“The editor program then identifies suitable templates, modules or sequences based on the queries. The relative importance of different items in the queries may be weighted. The weighting may be supplied by the user in response to suitable prompts, or may be generated automatically by the system based on existing programming. Using suitable relationships among data items, the user may be presented by the editor program with one or more templates, modules or sequences that represent a fundamental response, another set the represent a secondary response, a third set that represent a tertiary response, and so forth.”).

184. Moreover, in Marcus156’s system, one or more templates are selected before selecting, ordering, and concatenating video clips and transitional clips, as depicted in Figure 5:

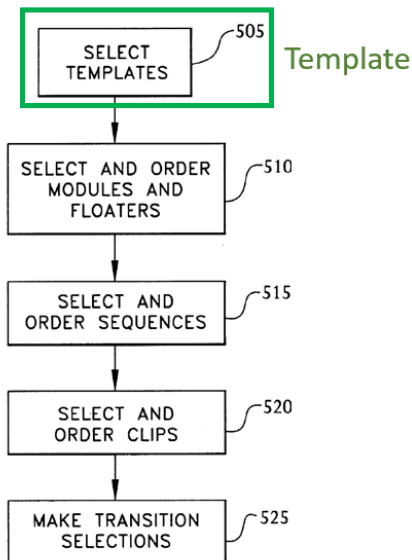


FIG. 5

Id. at Fig. 5 (annotated). Thus, it is my opinion that the subsequent (and claimed) steps all are done in accordance with (“performed according to”) the requirements and preferences of the selected template. *E.g., id.* at 8:19-25 (“Using the requirements in the selected templates and the demographic characteristics and other requirements specific to the program, as well as the inherent structural characteristics of the modules and sequences, the expert system or truth table then selects and orders one or more suitable modules and floater sequences, as shown by block 510 of FIG. 5.”); *see also id.* at 3:67-5:9, 8:19-25; 8:39-43; 9:2-6; 14:22-24. It is further my opinion that the templates comprise a plurality of template variables, namely the data

contained in the templates. *E.g.*, Marcus156 at 3:67-4:3 (“a family of defining values”).

185. *Wherein said one or more first templates comprise a plurality of variables corresponding to (a) information about the user:* Marcus156’s templates explicitly include demographic and psychographic parameters (which a POSITA would have considered as “a plurality of variables”) that are used to customize the video stream to the characteristics of the viewer. Marcus156 at 3:66-4:5 (“Next in the hierarchy is the instructional design layer, or TEMPLATE LAYER 315. This layer is characterized by a family of defining values which describe the range of the target audience in specific demographic and psychographic terms. Additionally, the overall outline of the subject matter is contained in this layer and is associated with demography where appropriate.”); 6:33-40 (“The coding interface 440 is the GUI designed to perform the entry of the medical and socio-demographic selection codes which apply to the clip. These include the ICDs for which this clip is applicable, the socio-demographic and medical attributes of those patients assumed to be the most likely potential viewers of the clip, and any special fields which when present in the client/patient profile will cause this clip to be selected.”); 13:47-51 (“[T]he audio and video assets in the database are characterized by such parameters as content, aesthetic features, and suitability to viewers with certain demographic or psychographic characteristics.”).

186. In my opinion, the demographic information is “information about the user” because it restricts the intended audience for the video programming based on certain user characteristics. *E.g., id.* at 4:9-14 (“For example, the instructional designer might define that the block of programming content called EFFECTS ON THE HEART is presented across three fundamental age groups, two levels of detail (summary/cursory and in-depth), both gender specific groups and four distinct ethnicity components.”); 14:30-34 (“The individual assets may have demographic characteristics. For example, an audiovisual clip may feature actors of a particular ethnic group, geographic origin, age, or other demographic characteristic. As a result, the assets may be selected based on demographic search criteria.”). In my opinion, the plain and ordinary meaning of “information about the user” would encompass this type of information. I examined the specification of the ’185 patent, the prosecution history of the family of the ’185 patent, and the appendix to the ’597 provisional, and did not discern any statement from the patentee restricting the scope of the claim term “information about the user.” I also note that a collection of general characteristics ultimately maps to specific people in the viewing audience.

187. Additionally, Marcus156’s system tracks information about individual users in its “database of viewer information” and in its “patient profile.” *E.g., id.* at 10:12-20 (“This database defines the identity of each viewer, including name and address information, and may also include social, economic and medical data

regarding the viewer, the history of items viewed by the viewer, and a file of preference information for the viewer.”); *see also id.* at 6:49-52 (“The Patient profile entry interface 415 is the GUI designed to perform the entry of the general, medical and socio-demographic codes for the specific patient for which the product is to be produced.”). This information is used to select video clips and transitions, and to sequence video programming. *Id.* at 10:16-20 (“In an on-line environment, the viewer may be prompted to provide viewer identification information which may be entered into a database or may be retained temporarily as a viewer data record for use in creation of customized programming.”); 4:56-59 (“The inhibition mechanism may be responsive to psychographic characteristics of the user, such as age, level of education, or even reason for the query.”). In my opinion, this information is also “information about the user,” as claimed.

188. In my opinion, this information is a “template,” as claimed. The plain and ordinary meaning of “template” does not require any particular data structure. Consistent with this, the intrinsic evidence does not define “template.” In particular, a POSITA would not restrict the scope of the “templates” claimed in the ’185 patent to the specific data structures that are part of the TEMPLATE LAYER 31 in Marcus156.

189. I additionally note that the appendix from the ’597 provisional says that a template can be code, in addition to data, and can employ complex programming

logic. Ex. 1005 at 34 n.6 (“Make note that programming line-numbers, looping, counting, and branching commands are a part of the template structures in one embodiment. Templates may be written in a programming language beneath a template guide, or may be simple data tables interpreted by the template parser.”).

190. I also note that the appendix contemplates multiple different templates addressed to different types of metadata. *Id.* at 35 (“In one embodiment, the master template controls, or limits, the choice of clips – and these will include transitional and hosting clips. The look-and-feel template is actually likely to be simply a convenient GUI-oriented place to change subjective factors like pacing, transitional feel, frequency of advertising, numbers and type of station-ID’s and the like by means of placing new search/tag attributes on the master template itself.”).

191. Finally, I understand based on its complaint filed in Eastern District of Texas, PO has accused TikTok’s video recommendation algorithms as a whole of being a claimed “template.” In my opinion, a POSITA would consider the “database of viewer information” or “patient profile” in Marcus156 to also be a “template,” applying the same apparent interpretation of “template” that PO used for infringement purposes.

192. To the extent that some particular data format is ultimately required in order to be a claimed “template,” it would have been obvious to include Marcus156’s “database of viewer information” or “patient profile” in its template layer. A

POSITA would have had a motivation to convert the data into the template format in order to better use the information to select video clips to form video programming. Synchronizing the data format would have simplified and streamlined the system. Templates provide a consistent structure for data, making it easier to organize and retrieve information. Standardized formats improve data quality and reduce errors in data entry and interpretation. Templates can streamline data processing and analysis by providing a uniform structure. As the database grows, templates make it easier to manage and update large amounts of user information. Template formats often enhance compatibility and interoperability between different systems or software applications. This can improve data sharing and integration across various platforms or departments. Templates can facilitate easier data input and updates for both users and administrators. Additionally, by simplifying data management and reducing the need for manual data handling, templates can lead to cost savings in the long run.

193. A POSITA would have had a reasonable expectation of success in any such modification because it would entail a routine conversion of information in one data format to another. Converting data into structured formats is a common practice in information technology and data management. Database management systems and template creation tools were well-established technologies as of the time of invention of the '185 patent. There were numerous extant software solutions and

frameworks designed for such conversions as of the time of invention of the '185 patent, and documentation for those tools would have been available to a POSITA. In particular, I note that Marcus156's system already explicitly uses data in template format, and also already includes software systems to import data pertaining to new video clips and users (*id.* at 6:15-7:9), both of which would reduce the amount of time and effort to make the modification because much of the software scaffolding is already present in Marcus156's system. Databases typically store information in structured formats (e.g., tables, fields) that lend themselves well to template conversion. The structured nature of database information facilitates mapping to template fields. Moreover, a person of ordinary skill in the art would routinely work with both databases and document templates, providing a pool of relevant expertise. Additionally, the conversion process could have been thoroughly tested and refined using sample data before full implementation. The conversion could also be implemented incrementally, allowing for adjustments and improvements along the way.

194. *Wherein said one or more first templates comprise a plurality of variables corresponding to . . . (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other:* Marcus156's template layer is described as containing "the temporal organizational templates embodied in this layer." *Id.* at 4:5-7. The "temporal organization" includes viewing scenarios of two

or more selected clips in a temporal relationship to each other. *E.g., id.* at 4:7-9 (“Here the instructional designer, or interactive author, defines the preferred temporal modes of presentation of the universe of assets.”); 4:14-21 (“Within this multi-dimensional array of program assets, the instructional designer might also define that the material be presented in the preferred sequence of—
INTRODUCTION TO THE HEART, IMPACT OF DIABETES ON THE
CARDIOVASCULAR-VASCULAR STRUCTURES, EFFECTS OF DIET,
EFFECTS OF EXERCISE, Q&A SESSION, SUMMARY.”); 8:39-43 (“For example, if it is not possible to assemble acceptable sequences in the manner dictated by the modules under the template, the expert system or truth table may discard one or more modules and select alternatives that meet the requirements of the template.”); 14:22-25 (“[A]ssets may be ordered by using a *template* that imposes sequential requirements on the assets”) (emphasis added).

195. In my opinion, these “sequential requirements” and “temporal modes of presentation” are “viewing scenarios of two or more selected video clips in a temporal relationship to each other,” as claimed. For example, they can describe an expected audience and subject matter for a sequence of temporally-ordered clips. *Id.* at 4:9-14 (“[T]he instructional designed might define that the block of programming content called EFFECTS ON THE HEART is presented across three fundamental age groups, two levels of detail (summary/cursory and in-depth), both gender

specific groups and four distinct ethnicity components.”). This is a “scenario” for how the video programming is ultimately viewed.

196. I note that the ’185 patent specification does not contain the term “viewing scenario.” Additionally, “viewing scenario” only shows up once in the appendix to the ’597 provisional application, where its meaning is not restricted. Ex. 1005 at 25 (“Unlike earlier database architectures, this model has found a way to separate templates that describe every possible potential viewing scenario from the actual media that might be viewed.”). Therefore, in my opinion, the intrinsic evidence does not support a narrow interpretation of “viewing scenario” that would somehow exclude the “sequential requirements” or “temporal modes of presentation” described in Marcus156. The plain and ordinary meaning of “viewing scenario” to a POSITA at the time of the invention would encompass the scenarios described in Marcus156.

197. *Wherein said one or more first templates comprise a plurality of variables corresponding to . . . (c) style requirements:* The templates described in Marcus156 additionally contain constraints on the style of the video programming. *E.g.*, Marcus156 at 9:2-6 (“Templates might be provided to content providers or even to end-users, which would allow a specific *style of editing*, audio cutting or mixing, or program formation; perhaps a template may be provided *associated with an individual editor*. Such decision-capture systems already exist for other uses

which could be adapted to the assembly of audio and video.”) (emphasis added); 14:22-29 (“[A]ssets may be ordered by using a *template* that imposes sequential requirements on the assets The attributes of the assets impose certain limitations on the ordering of the assets. For example, it may not be possible for *aesthetic reasons* to place a certain clip immediately before or after a certain other clip.”) (emphasis added); 3:41-46 (“In this example, the master look-and-feel level of organization might define hard-cut as the aesthetically optimal, and therefore the default, choice. Numerous other aesthetic or programmatic choices that affect the delivery and placement of media elements in the concatenated stream can be defined.”).

7. Element 1[f]: “presenting said first video programming in a first area of a display associated with one of said computing devices;”

198. In my opinion, Marcus156 discloses this element.

199. *Presenting said first video programming in a first area of a display associated with one of said computing devices:* Marcus156 discloses “OUTPUT AND DISPLAY DEVICES” as components of computer 20. *Id.* at Fig. 1; 2:51-54 (“Referring to FIG. 1, there is depicted a system 10 according to the invention. System 10 includes a computer 20 including processor 30, input devices 40, output and display devices 50 and memory device 60.”). Marcus156’s computer 20 is the claimed “said computing device.” Marcus156’s computer 20’s display device 50 is

the claimed “display associated with one or more of said computing devices.” As described further below, the same “said computing device” that performs the searching and selecting steps can also present the video programming. When the first video programming has been assembled, it is displayed on a display device 50 of computing device 20. Marcus156 describes several ways that this may happen, any one of which (in my opinion) independently discloses or renders obvious this element.

200. As a first example, Marcus156 describes the first video programming being presented as part of the user interface for a “List Editor/Previewer” program, which allows a user to interact with video produced by the editor program. *See, e.g., id.* at 8:2 (“Programs are created from clips by an editor program.”); 6:3-7 (“4. A decision/transition list editor 430 with built-in transition previewer for both audio and video components of clips. 5. A production player/recorder which embodies an engine for both audio and video transitions.”); 6:9-12 (“Each component is started by the user selecting an icon from an appropriate folder. The user interface is preferably be graphical in nature.”); 7:11-15 (“The List Editor/Previewer 430 is the GUI designed to provide the production editor with the ability to change, add to and preview the EDL that has been produced by the editor program 425 in response to information entered in the Patient Profile Entry Interface 415.”). Thus, using the “List Editor/Previewer” program and its graphical user interface, a video editor is

presented with the assembled first video programming in a first area of a display associated with computer 20. For instance, Marcus156 describes the “preliminary EDL” as “being played.” Marcus156 at 7:51-52.

201. As a second example, Marcus156 describes the first video programming being assembled by a “production player” or “assembly program” component and then output as a finalized digitized video stream to be displayed. *See, e.g., id.* at 7:36-41 (“[T]he player output thread will be started taking its input instructions from the player FIFO. The player output thread will operate asynchronously, transferring each data component specified in the player FIFO to the appropriate hardware component, such as video and audio decoder modules”); 7:59-60 (“The video and audio is then re-compressed and written to the output transition file.”); 9:30-32 (“The assembly program uses the transition information in assembling the clips into the program.”).

202. As a third example, Marcus156 discusses users “receiving the programming online through a modem or Internet link” and then “playing of the video sequence.” *Id.* at 11:4-10; *see also id.* at 14:35-40 (describing an “on-line environment” where “the system may provide delivery of a predetermined audiovisual program so long as the viewer does not seek to interact with the system, e.g., by positioning a mouse pointer on the screen and clicking”). Thus, in this on-line environment, the programming is played on the screen of a computer.

- 8. Element 1[g]: “concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display;”**

203. In my opinion, Marcus156 renders this element obvious.

204. *Concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display:* As an initial matter, Marcus156 describes creating media streams that include both video elements and still elements such as individual pictures, which are similar to frames. *Id.* at 13:63-67 (“[A] database can be constructed for use with still photographs to be arranged into a program with an audio track. Interactive assets, such as two-dimensional or three-dimensional model and graphical user interfaces can also be catalogued in a database of this type.”); cl. 3 (“The method of claim 1, wherein said media elements are still photographs, and said media programming comprises a series of said still photographs.”).

205. Furthermore, Marcus156 describes the system tracking different areas of a display screen that is playing the assembled first video programming. *Id.* at 14:46-52 (“When the user moves a pointer or mouse to a particular portion of the screen, and clicks on the screen, the system may select, for example, assets incorporating more detailed information on certain portions of the heart

corresponding to material on the portion of the screen where the mouse was located when clicked.”).

206. Additionally, Marcus156 discloses varying the contents of the displayed media in different portions of the screen. *Id.* at 14:61-67 (“The system may, for example, provide more interesting visual data in areas of the screen that lead to the selection of material that the viewer should see, or provide smooth transitions from audiovisual assets incorporated in the program as a result of viewer input to audiovisual assets incorporated in the program to achieve a predetermined purpose, e.g., to convey certain information to the viewer.”).

207. Marcus156 also discusses displaying a companion window alongside a playing video, where the companion window includes different user interface controls. *Id.* at 11:4-10 (“For viewers who are receiving the programming online through a modem or Internet link, this might be a window which remains open and active during the playing of the video sequence, or which becomes open and active upon the presentation of the appropriate superimposed strings, which allows the viewer the opportunity to record that sequence.”).

208. Furthermore, Marcus156 describes the “List Editor/Previewer” program which “is the GUI designed to provide the production editor with the ability to change, add to and preview the EDL that has been produced by the editor program 425 in response to information entered in the Patient Profile Entry Interface.” *Id.* at

7:11-15; 6:3-5 (“4. A decision/transition list editor 430 with built-in transition previewer for both audio and video components of clips.”); 6:11-14 (“The user interface is preferably be graphical in nature. Actual keyed data entry is preferably kept to a minimum where most entries are made by selecting an option from a set or an entry from a list.”).

209. Given Marcus156’s guidance in the sections cited above, it is my opinion that a POSITA, to the extent necessary, would have found it obvious to slightly modify Marcus156’s user interface to add a concurrent display of a frame.

210. There are several aspects of Marcus156’s system that could be enhanced to accomplish this. As a first example, the “more detailed information on certain portions of the heart corresponding to material on the portion of the screen where the mouse was located when clicked” would most naturally include a still frame of the portion of the video. As a second example, the content in the companion window “which remains open and active during the playing of the video sequence” would logically include a still frame of the playing video that included the superimposed strings. As a third example, the interface for the “List Editor/Previewer” application would display thumbnails of clips in the preliminary edit decision list (EDL) as well as a way to play the video programming. In my opinion, making any one of these individual modifications would be approximately the same amount of work. Implementing all of them would not be significantly more

work than implementing one, due to commonalities in the changes. Implementing any one of them would be sufficient to render this limitation obvious.

211. In my opinion, a POSITA would have been motivated to add this type of interface to Marcus156's system. Adding the concurrent display would have improved the user interface by allowing the user to view more relevant information at one time. Displaying a still frame alongside the running video enables enhanced precision in editing because it allows editors to examine specific moments in detail while maintaining context within the overall video. Having both the running video and a still frame visible simultaneously allows for easy comparison between frames or sections of the video. This would have been particularly useful for color correction, continuity checking, and assessing the impact of edits. Editors can more easily spot issues like motion blur, focus problems, or artifacts in individual frames while still seeing how these frames fit into the moving sequence. Providing a thumbnail view of one or more video frames along the playing video would allow the user to navigate through the playing video stream in a more efficient manner. For example, displaying the frames would give a user more time to click on the screen to get additional relevant information on what is being displayed. As another example, it would increase workflow efficiency by allowing editors to make decisions about specific frames without repeatedly pausing and resuming the video playback. Adding a frame display would have allowed better integration with other

video editing or viewing tools that operate on a frame-by-frame basis. Adding such an additional graphical element would comport with Marcus156's instructions that "[t]he user interface is preferably . . . graphical in nature." *Id.* at 6:11-12.

212. In my opinion, a POSITA would have had a reasonable expectation of success in making such a software modification because Marcus156 explicitly discloses software techniques to extract a frame image from a video in the context of preparing transitional clips, and the same technique could be used to generate the frame images to be concurrently displayed. *Id.* at 7:54-57 ("The first frame of each segment is decoded into separate buffers and then the transition function is applied to each matching pixel of both images to produce the third rendered frame buffer."). Thus, the core functionality for both video playback and frame extraction already existed within Marcus156's system. More generally, techniques for extracting and displaying individual frames from a video stream were well-established in the field of video processing as of the time of the invention. Video codecs and container formats were already highly standardized, making it straightforward to access and manipulate video data consistently. For example, Marcus156 discusses and employs the standard MPEG video file format, for which multiple tools existed as of the time of the invention to extract frames, such as FFmpeg. Moreover, this type of dual display setup was already in wide use in video editing systems such as Avid at the time of the invention of the '185 patent. It was also in common usage in other

systems such as picture-in-picture TV displays and in sports replay systems. Display technologies and graphics libraries available at the time of the invention readily supported multiple viewport rendering, making it feasible to show different content in separate areas of the screen, without much programming work. Similarly, most GUI frameworks used in video editing software at the time of the invention of the '185 patent supported dynamic layout changes and multiple widget areas, facilitating the addition of a new display area. These libraries and frameworks were supported by extensive documentation, open-source projects, and developer communities which provided resources and potential solutions for common challenges in video software development. Additionally, the computing hardware typically in use at the time of the invention (*i.e.*, Marcus156's computer 20) had more than sufficient processing power and memory to handle simultaneous video playback and still frame display. Well-established optimization techniques for video processing and display could have been employed to ensure smooth performance of the concurrent display feature. The feature could be implemented incrementally, starting with basic functionality and refining it over time, reducing development risks.

213. As I noted earlier, making any one of the three proposed individual modifications would be approximately the same amount of work, and each would have a reasonable expectation of success. Implementing all of them would not be significantly more work than implementing one, due to commonalities in the

changes, and would also have a reasonable expectation of success. Implementing any one of them would be sufficient to render this limitation obvious.

9. Element 1[h]: “receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;”

214. In my opinion, Marcus156 renders this element obvious.

215. *Receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display:* For the reasons I gave above in the context of element 1[g], Marcus156 renders obvious the concurrent display of a playing video clip and a frame from that video clip in different areas of the display.

216. Marcus156 also discloses that a user can select displayed elements via, for example, a mouse click. Marcus156 at 14:35-43 (“The system may permit, in an on-line environment, switching between a predetermined program and interactive video. For example, the system may provide delivery of a predetermined audiovisual program so long as the viewer does not seek to interact with the system, e.g., by positioning a mouse pointer on the screen and clicking. At this point, the system may select assets that are appropriate to the particular information that the viewer appears to be seeking based on the information on the screen at the time.”); 14:46-52 (“When the user moves a pointer or mouse to a particular portion of the screen, and clicks on the screen, the system may select, for example, assets

incorporating more detailed information on certain portions of the screen where the mouse was located when clicked.”).

217. Additionally, because the “List Editor/Previewer” provides “the ability to change, add to and preview the EDL,” it entails selecting relevant user interface elements. *Id.* at 7:11-15 (“The List Editor/Previewer 430 is the GUI designed to provide the production editor with the ability to change, add to and preview the EDL that has been produced by the editor program 425 in response to information entered in the Patient Profile Entry Interface 415.”).

218. More generally, Marcus156’s system supports “viewer input” to perform a search for video content. *Id.* at 14:59-67 (“The system may also be employed in an interactive system that leads the viewer to select a desired sequence. The system may . . . provide smooth transitions from audiovisual assets incorporated in the program as a result of viewer input to audiovisual assets incorporated in the program to achieve a predetermined purpose . . .”).

219. In light of these disclosures, it is my opinion that it would have also been obvious for Marcus156’s system, already enhanced as described in the previous claim element to concurrently display a video and a frame from that video, to also track a user selection of a frame and for computer 20 to receive data about that selection.

220. In my opinion, a POSITA would have been motivated to make such a modification. Once a decision has been made to concurrently display a playing video and a frame from that video (as detailed in the previous element), then adding interactive feedback to user selections of the frame (and thus tracking the user selection of the frame) would have improved the functionality and responsiveness of the system. For example, Marcus156 already discloses allowing a user to click on a frame in a video clip to run a second search for other video clips. The same motivation would apply for allowing the user to click on the concurrently displayed frame, too, and indeed doing so would further the benefit of affording the user additional time to click on that particular frame (which may have already passed in the main video but is still visible in the concurrently-displayed frame). Allowing users to directly select frames with a mouse click or tap provides an intuitive and accurate way to pinpoint exact moments in the video. Mouse clicks or taps are familiar and natural interactions for most users. This method of frame selection is more user-friendly and efficient than alternatives like manually entering timecodes or scrubbing through a timeline. Moreover, tracking the user selection would improve the accessibility of the system. For users with certain disabilities, selecting frames by a mouse click may be easier than using keyboard shortcuts or timeline scrubbing. The ability to select frames complements the concurrent display of

running video and a static frame. Users can keep context (via the running video) while focusing on specific moments (via frame selection).

221. In my opinion, a POSITA would have had a reasonable expectation of success in making such a modification. Determining what displayed user interface element a user click corresponds to is a routine software function that was present in the first graphical user interfaces such as the Xerox Alto from 1973. The existing graphical interface code used by Marcus156's "List Editor/Previewer" would include such code, or it could be added with minimal programming effort, as could any necessary code to transmit the selection. Programming frameworks and development environments in common usage as of the time of the invention of the '185 patent already provided robust mechanisms for handling user inputs like mouse clicks. Integrating these events into Marcus156's system (to the extent they are not already handled) would use well-understood and widely implemented techniques. Frame selection functionality existed in many video players and video editing tools as of the time of the invention, providing proven concepts that could be adapted and enhanced for Marcus156's system. The feature could be implemented incrementally, starting with basic functionality and then refining and optimizing it, reducing the risk of major integration issues.

- 10. Element 1[i]: "selecting, by one or more of said computing devices, a second video clip based on at least one stored**

association between the second video clip and at least one first video clip indicated by said user selection;”

222. In my opinion, Marcus156 discloses this element, or at least renders it obvious.

223. *Selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection:* Marcus156 describes how based on a user interaction with the display of an existing stream of video programming that was already generated by the system, the can perform interactive additional searches and selections of video clips, including the generation of a new stream of video programming. *See, e.g., id.* at 14:40-45 (“At this point, the system may select assets that are appropriate to the particular information that the viewer appears to be seeking based on the information on the screen at the time. The system will then generate suitable transitions and add those assets to the programming.”); 14:46-55 (“When the user moves a pointer or mouse to a particular portion of the screen, and clicks on the screen, the system may select, for example, assets incorporating more detailed information on certain portions of the screen where the mouse was located when clicked. The system applies suitable transitions to the assets and incorporates these assets into the programming interactively in response to mouse clicks or other suitable input from the viewer.”); 7:11-15 (“The List Editor/Previewer 430 is the GUI designed to provide the production editor with the

ability to change, add to and preview the EDL that has been produced by the editor program”).

224. While Marcus156’s example in column 14 addresses “a predetermined program,” (*id.* at 14:35-37: “The system may permit, in an on-line environment, switching between a predetermined program and interactive video”), Marcus156’s disclosure is not so limited. In particular, Marcus156’s description of allowing the user to interact with video programming and generate new programming by selecting a portion of the screen is agnostic as to whether that video programming is “predetermined” or itself selected based on a first search. At a minimum, it would have been obvious to apply the same teachings to both “predetermined” and dynamically generated video programming. In my opinion, a POSITA would have been motivated to do so because the same benefits of allowing the user to interact with a “predetermined” program to, for example, search for “particular information” related to that program equally apply to allow the user to search for additional information related to a targeted program. In my opinion, a POSITA would also have had a reasonable likelihood of success in making any such modification, since the same selection mechanics can be applied to either type of video.

225. The selection of a second video clip is based on the stored associations between the prior, selected clip and the new, second clip, including because both videos may have the same or similar tags. *Id.* at 9:16-21 (“Using suitable

relationships among data items, the user may be presented by the editor program with one or more templates, modules or sequences that represent a fundamental response, another set that represent a secondary response, a third set that represent a tertiary response, and so forth.”); 3:26-30 (“The system by which the content information is defined is preferably organized to provide for relationships among types of information. Such relationships may be levels of generality, and other subject matter relationships.”); 14:40-45 (“At this point, the system may select assets that are appropriate to the particular information that the viewer appears to be seeking based on the information on the screen at the time. The system will then generate suitable transitions and add those assets to the programming.”); 14:46-55 (“When the user moves a pointer or mouse to a particular portion of the screen, and clicks on the screen, the system may select, for example, assets incorporating more detailed information on certain portions of the screen where the mouse was located when clicked. The system applies suitable transitions to the assets and incorporates these assets into the programming interactively in response to mouse clicks or other suitable input from the viewer.”). Thus, where a user clicks to obtain “more detailed information on certain portions of the heart” (*id.* at 16:46-52), both sets of clips could be associated with the tags “IMPACT OF DIABETES ON THE CARDIOVASCULAR-VASCULAR STRUCTURES” and “DIABETIC HEART,”

while the second set of clips are further associated with the more specific tag “LEFT VENTRICLE.” *Id.* at 4:14-21, 4:40-45.

11. **Element 1[j]: “selecting, by one or more of said computing devices, a plurality of video clips including the second video clip, and determining a second temporal sequence for concatenating the selected plurality of second video clips, wherein said selecting, determining, and concatenating is performed in accordance with one or more first or second templates, to define second video programming;”**

226. In my opinion, Marcus156 discloses this element.

227. *Selecting, by one or more of said computing devices, a plurality of video clips including the second video clip, and determining a second temporal sequence for concatenating the selected plurality of second video clips, wherein said selecting, determining, and concatenating is performed in accordance with one or more first or second templates, to define second video programming:* The same systems, data, templates, and algorithms that Marcus156’s system used to select a plurality of first video clips, determine a temporal sequence for them, and concatenate them are used in the second search. Therefore, in my opinion, all of the evidence and analysis reviewed above in the context of elements 1[b], 1[d], 1[e], and 1[i] applies to this element as well.

12. **Element 1[k]: “wherein said one or more first or second templates comprises a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of**

two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and”

228. In my opinion, Marcus156 discloses this element.

229. *Wherein said one or more first or second templates comprises a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements:* The templates used in follow-up searches in Marcus156’s system can be the same or different as the initial search (discussed in the context of claim element 1[e] above). Marcus156 discloses that when a search is being conducted, the expert system “selects and orders one or more suitable templates” as a beginning step. *Id.* at 8:13-15. The interactive searches described in Marcus156 at 14:35-67 use templates, just as any initial or non-interactive search does. And for the reasons discussed above in the context of element 1[e], those templates contain “a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements.”

13. Element 1[l]: “presenting said second video programming on a display associated with one or more of said computing devices.”

230. In my opinion, Marcus156 discloses this element.

231. *Presenting said second video programming on a display associated with one or more of said computing devices:* After performing a follow-up search

and defining second video programming, the programming is displayed on the computing device's display, just like the first video programming is. *See, e.g.,* Marcus156 at 14:52-55 (“The system applies suitable transitions to the assets and incorporates these assets into the programming interactively in response to mouse clicks or other suitable input from the viewer.”); 14:61-67 (“The system may, for example, provide more interesting visual data in areas of the screen that lead to the selection of material that the viewer should see, or provide smooth transitions from audiovisual assets incorporated in the program as a result of viewer input to audiovisual assets incorporated in the program to achieve a predetermined purpose, e.g., to convey certain information to the viewer.”). I provided additional detail in the context of claim element 1[f] above.

B. Claim 2: “A method according to claim 1, wherein at least one of said video clips of said first or second video programming is a transitional clip.”

232. In my opinion, Marcus156 discloses this claim.

233. *Wherein at least one of said video clips of said first or second video programming is a transitional clip:* As I described above in the context of claim element 1[c] above, Marcus156's system includes transitional clips (either “bridge”/“floater” clips or clips that are modified versions of other clips to add transition features) in the first video programming. And as described in the context of element 1[j], the same types of transitional clips are also included in the second

video programming. I provided additional details above in the context of claim elements 1[b], 1[d], and 1[e].

C. Claim 3: “A method according to claim 2, wherein at least one of said video clips of each of said first and second video programming is a transitional clip.”

234. In my opinion, Marcus156 discloses this claim.

235. *Wherein at least one of said video clips of each of said first and second video programming is a transitional clip:* Transitional clips can be present in the initial assembled video stream (“first video programming”) and also in the interactive video stream that is assembled in response to further user input (“second video programming”). Therefore, for the same reasons I detailed above in the context of claim 2, claim 3 is also disclosed by Marcus156.

D. Claim 4: “A method according to claim 3, wherein said transitional clip is a hosting clip consistent with at least one template variable.”

236. In my opinion, Marcus156 discloses this claim, or at least renders it obvious.

237. *Wherein said transitional clip is a hosting clip consistent with at least one template variable:* I note that the term “hosting clip” is described broadly in the appendix. *E.g.*, Ex. 1005 at 34 (“The clips are generally interspersed with ‘hosting’ segments. These are clips that create the overall mood and personality of the system experience.”), 71 (“By ‘hosting’ we mean the addition of commentary, chat,

monologue, and the like – usually through the agency of a human host. While the hosting of a program assembled through the use of templates isn't required, some form of user-specific customization is highly desirable, and this is best done through the addition of interstitial segments that cohere the flow of cybernetically-assembled clips.”), 74 (“Any form of entertainment inserts/content or hosting segments can be the subject of this model.”).

238. Marcus156 discloses transitional clips that are hosting clips. In particular, Marcus156 teaches that video clips in the database 100 can be tagged with fields that are particular to actors that appear in the clips. Marcus156 at 3:21-25 (“Fields may be provided to indicate technical complexity, specific items of information conveyed, demographic characteristics of actors in clips, style of script, and other suitable information.”); 14:30-34 (“The individual assets may have demographic characteristics. For example, an audiovisual clip may feature actors of a particular ethnic group, geographic origin, age, or other demographic characteristic. As a result, the assets may be selected based on demographic search criteria.”). Marcus156 gives an example of a media tag “TALKING HEAD” which may be applied to clips, and therefore used to select the clip for inclusion in the video stream as a transitional clip. *Id.* at 4:39-48. A talking head clip features a person (*i.e.*, a host) talking. The '185 patent's appendix describes such talking head clips from Marcus156 as hosting clips. Ex. 1005 at 102 (“Take the case of a human host

who wants to comment upon a given clip or clip set. These clips are defined as Hosting clips – that is, they are perhaps ‘talking heads’ which set tone and keep continuity throughout the presentation of the series of disparate clips.”). Like other clips, these types of actor-specific or talking head clips can be “bridge” or “floater” transitional clips that are used to link other clips. Marcus156 at 4:22-38; *see also* claim 2 above. If such a clip is selected as a linking clip, it is my opinion that a POSITA would have considered it to be a “hosting clip” that is also a “transitional clip.”

239. Additionally, Marcus156 discloses that the hosting clips are “consistent with at least one template variable.” Like any other transitional clip, the hosting clips must comport with the contents of the selected templates in order to be included in the final video programming; if the content of a hosting clip is inconsistent with the template then it will not be included. Marcus156 at 8:19-25 (“Using the requirements in the selected templates and the demographic characteristics and other requirements specific to the program, as well as the inherent structural characteristics of the modules and sequences, the expert system or truth table then selects and orders one or more suitable modules and floater sequences, as shown by block 510 of FIG. 5.”); *see also, generally*, element 1[e] above.

240. Moreover, Marcus156 teaches that transitional clips can be selected based on overall “look-and-feel,” just as the appendix describes hosting clips as

creating “the overall mood and personality of the system experience.” Marcus156 at 3:41-46 (“[T]he master look-and-feel level of organization might define hard-cut as the aesthetically optimal, and therefore the default, choice. Numerous other aesthetic or programmatic choices that affect the delivery and placement of media elements in the concatenated stream can be defined.”). This is another indication that the hosting clips are consistent with at least one template variable.

241. In my opinion, a POSITA would have been motivated to make any needed modification to Marcus156 to use hosting clips as transitional clips. First, including hosting clips would have allowed the concatenated video stream to take the form of traditional television shows that include a host, such as nature documentaries, late-night talk shows, or sequences of music videos hosted by a video jockey. Using hosting clips as transitions has been ubiquitous since the early days of television. Using such a format in the context of Marcus156 would thus breed familiarity for users (the clips would be presented in a format the users are already familiar with) and also expand the potential repertoire of the video programming (by for example allowing for introductions to video programming users may otherwise be unfamiliar with). Second, including hosting clips would allow for additional types of transitions used for comedic or dramatic effect (“And now for something completely different...” / “A few hours later...”). This would improve the quality and potential breadth of the delivered video programming.

242. In my opinion, a POSITA would have had a reasonable expectation of success in making any needed modification to Marcus156 to use hosting clips as transitional clips. Given Marcus156's existing disclosure of tracking talking head clips, a POSITA would have been able to exploit the existing data tracking these types of clips. Additionally, a POSITA would have been able to modify the existing templates in Marcus156's system to introduce requirements to use the hosting clips as transitional clips, given the extensive sequencing and clip association capabilities detailed in Marcus156.

E. Claim 5: "A method according to claim 2, wherein at least one transitional clip is a hosting clip consistent with at least one template variable."

243. In my opinion, Marcus156 discloses this claim, or at least renders it obvious.

244. *Wherein at least one transitional clip is a hosting clip consistent with at least one template variable:* In my opinion, Marcus156 discloses or renders obvious this claim for the same reasons given above in the context of claim 4.

F. Claim 6: "A method according to claim 2, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming."

245. In my opinion, Marcus156 discloses this claim.

246. *Wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips*

at specific locations in the first or second video programming: The temporal organization templates used by Marcus156 include restrictions on placement of clips at specific locations in the programming, as described earlier in the context of claim 1[e]. For example, the variables define a “preferred sequence” of clips ordered by content type. *Id.* at 4:14-21. And more explicitly, “assets may be ordered by using a template that imposes sequential requirements on the assets”). *Id.* at 14:22-25. Because these “sequential requirements” place clips in temporal relationships to each other, they are “sequencing limitations for incorporating video clips at specific locations” in the programming.

G. Claim 7: “A method according to claim 6, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.”

247. In my opinion, Marcus156 discloses this claim.

248. *Wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame:* In my opinion, Marcus156’s templated sequencing limitations disclose each of these three categories. The limitations may be length requirements or requirements for presentation within a particular time frame. *Id.* at 4:65-5:2 (“The system may set, or the user may select, a length of time

for the program, and the inhibition mechanism will operate to limit the total number of clips selected to the selected time, as well as choosing clips according to suitability for the viewer and the viewer's purpose."); 5:27-33 ("Numerous optional control tags may be used, such as, with reference to video information, . . . preferred or required transitions, including length and type of transition . . ."). Marcus¹⁵⁶ also discloses limitations on the order of clips. *Id.* at 4:14-21 ("Within this multi-dimensional array of program assets, the instructional designer might also define that the material be presented in the preferred sequence of—INTRODUCTION TO THE HEART, IMPACT OF DIABETES ON THE CARDIOVASCULAR-VASCULAR STRUCTURES, EFFECTS OF DIET, EFFECTS OF EXERCISE, Q&A SESSION, SUMMARY.").

H. Claim 8: "A method according to claim 6, wherein one or more video clips comprise advertising."

249. In my opinion, Marcus¹⁵⁶ discloses this claim.

250. *Wherein one or more video clips comprise advertising:* Video clips in the database ("assets") may be tagged as advertising or promotion. *Id.* at 3:52-55 ("At this level of organization, the behaviors of ancillary assets such as promotional segments, information identifying the system, advertisements and news-flashes are defined."). In addition to explicitly describing "advertisements" and "promotional segments," Marcus¹⁵⁶ also discloses that clips may be classified as "market-driven," which a POSITA would understand to include advertising. *Id.* at 3:55-57

(“These assets embody aesthetic, program or instructional design, as well as market-driven, or viewer defined behaviors.”). Consistent with this, Marcus156 describes in the Background of the Invention section that that customized advertising programming is valuable. Marcus156 at 1:19-26 (“Customization of audiovisual programing is useful in many applications. For example, in advertising certain products, and in particular automobiles, one promotional technique is to prepare promotional videotapes which are sent to potential customers on their request. The desirability of customizing such videotapes to demographic or other characteristics of individual consumers are of course substantial.”). One of the objects of Marcus156’s invention is to facilitate improved customization of video programming, such as the customized advertising example, by automating the process and eliminating the need for human video editors. *Id.* at 1:37-56. Thus, a POSITA would have recognized that one or more clips in Marcus156’s system contain advertising.

I. Claim 9: “A method according to claim 6, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.”

251. In my opinion, Marcus156 discloses this claim.

252. *Wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history:* As described above in the context of claim element 1[e], the templates in Marcus156’s

system are explicitly described as containing demographic and psychographic information. Additionally, the viewing history of individual users is tracked in the “database of viewer information.” *Id.* at 10:12-16 (“This database [of viewer information] defines the identity of each viewer, including name and address information, and may also include social, economic and medical data regarding the viewer, *the history of items viewed by the viewer*, and a file of preference information for the viewer.”) (emphasis added).

J. Claim 10: “A method according to claim 6, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.”

253. In my opinion, Marcus156 discloses this claim.

254. *Wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates:* Marcus156’s system uses templates to select additional clips associated with a given clip, for example by limiting the universe of clips which may be associated with each other based on an additional restriction in the template. *Id.* at 3:67-4:5 (“This [template] layer is characterized by a family of defining values which describe the range of the target audience in specific demographic and psychographic terms. Additionally, the overall outline of the subject matter is contained in this layer and is associated with demography where appropriate.”).

255. The templates are used to select video clips based on associations with other video clips, including by using a tag or attribute of one clip as a necessary requirement for another clip. *E.g., id.* at 4:56-64 (“The inhibition mechanism may be responsive to psychographic characteristics of the user, such as age, level of education, or even *reason for the query*. The tags are responsive to this type of information. The inhibition mechanism *may be modified dynamically as the database is mined for suitable assets*. The inhibition mechanism may be viewed as in a multi-dimensional set of psychographic requirements. Clips can be evaluated for their responsiveness in the various dimensions.”) (emphasis added); 4:66-5:2 (“[T]he inhibition mechanism will operate to limit the total number of clips selected to the selected time, as well as choosing clips according to suitability for the viewer and *the viewer’s purpose*.”) (emphasis added).

K. Claim 11: “A method according to claim 6, wherein said stored associations between said video clips comprise a plurality of media tags.”

256. In my opinion, Marcus156 discloses this claim.

257. *Wherein said stored associations between said video clips comprise a plurality of media tags*: Marcus156 discloses a variety of media tags (which it also refers to as “header information”) that are attached to particular video clips. *Id.* at 3:4-8 (“Individual items of information within the header will be referred to herein as tags. The tags are divided into two general categories, namely control tags 115

and content tags 120. Content tags 120 contain information identifying the content of the clip.”); 5:3-4 (“[C]ontrol tags define audio and video components of clips.”); 5:27-33 (“Numerous optional control tags can be used, such as, with reference to video information, a luminance range 160 for the opening and closing of the clip, preferred or required transitions, including length and type of transition, a dominant chroma value for each of the opening and closing of the clip, start and end points for dialog and action.”).

- L. Claim 12: “A method according to claim 6, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.”**

258. In my opinion, Marcus¹⁵⁶ discloses this claim.

259. *Wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history:* The media tags include at least context, subject matter, and demographics. *Id.* at 3:21-26 (“Fields may be provided to indicate technical complexity, specific items of information conveyed, demographic characteristics of actors in clips, style of script, and other suitable information. One or more fields may be provided within the content tag to define content.”); 4:41-48 (“These elements carry tags which define their specific content, such as: DIABETIC HEART, LEFT VENTRICLE, DAMAGE TO,

HYPERGLYCEMIA, MALE, AGE 50, TALKING HEAD. The first three content tags will be noted as being in hierarchical order from most general to most specific. The next two are examples of demographic tags, and the final tag is a simple example of a tag denoting style.”).

260. The demographic information in the tags in turn includes information about source or region. *Id.* at 14:30-33 (“The individual assets may have demographic characteristics. For example, an audiovisual clip may feature actors of a particular ethnic group, geographic origin, age, or other demographic characteristic.”).

261. The tags include psychographic information. *Id.* at 13:47-51 (“[T]he audio and video assets in the database are characterized by such parameters as content, aesthetic features, and suitability to viewers with certain demographic or psychographic characteristics.”); 4:61-63 (“The inhibition mechanism may be viewed as in a multi-dimensional set of psychographic requirements.”).

M. Claim 13: “A method according to claim 1, wherein said first and second video programming are each presented as a stream of information to the user.”

262. In my opinion, Marcus156 discloses this claim.

263. *Wherein said first and second video programming are each presented as a stream of information to the user:* In Marcus156’s system, the finalized first and second video programming is assembled and presented to the user as a unified

video stream. That is, the user sees a sequence of video programming where the clips and transitions proceed sequentially, one after the other. *See, e.g., id.* at 3:44-46 (“Numerous other aesthetic or programmatic choices that affect the delivery and placement of media elements in the *concatenated stream* can be defined.”) (emphasis added); 5:35-37 (“This type of information can be derived by suitable software reviewing the data *stream* before or after the clips are concatenated”) (emphasis added); 7:7-10 (“The transitions are then executed and inserted into the *stream* by use of a system of removal of portions of the tag extremities used in temporal transitions.”) (emphasis added); 14:13-17 (“The system is also capable of arranging assets in an order, and creating transitions between assets, to create a concatenated stream of audiovisual programming that has transitions that appear to the viewer to have been edited by a human editor.”); 9:30-32 (“The assembly program uses the transition information in assembling the clips into the program.”); cl. 1 (“assembling said media elements into media programming”).

264. I understand that during prosecution of the '185 patent, the examiner stated that “Marcus [is] silent to explicitly disclose that said first and second video programming are each presented as a stream of information to the user.” Ex. 1005 at 222. In my opinion, this statement is not accurate. As detailed above, there should be no dispute that Marcus¹⁵⁶ discloses presenting the video programming as a stream of information to the user.

- N. Claim 14: “A method according to claim 1, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.”**

265. In my opinion, Marcus156 discloses this claim for the same reasons given above in the context of claim 6, which uses similar language.

- O. Claim 15: “A method according to claim 14, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.”**

266. In my opinion, Marcus156 discloses this claim for the same reasons given above in the context of claim 7, which uses similar language.

- P. Claim 16: “A method according to claim 1, wherein one or more video clips comprise advertising.”**

267. In my opinion, Marcus156 discloses this claim for the same reasons given above in the context of claim 8, which uses similar language.

- Q. Claim 17: “A method according to claim 1, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.”**

268. In my opinion, Marcus156 discloses this claim for the same reasons given above in the context of claim 9, which uses similar language.

- R. Claim 18: “A method according to claim 1, wherein the step of selecting a second video clip based on said association with at least**

one first video clip is performed according to at least one of the first or second templates.”

269. In my opinion, Marcus156 discloses this claim for the same reasons given above in the context of claim 10, which uses similar language.

S. Claim 19: “A method according to claim 1, wherein said stored associations between said video clips comprise a plurality of media tags.”

270. In my opinion, Marcus156 discloses this claim for the same reasons given above in the context of claim 11, which uses similar language.

T. Claim 20: “A method according to claim 1, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.”

271. In my opinion, Marcus156 discloses this claim for the same reasons given above in the context of claim 12, which uses similar language.

VI. Ground 2: Marcus156 and Rosen Render Claims 1-20 Obvious

272. In my opinion, Marcus156 as a base reference, with modifications according to the teachings of Rosen, renders challenged claims 1-20 obvious.

A. Specific Proposed Modification to Marcus156’s System

273. For either the “interactive system” described in Marcus156 at 14:35-67 or the “List Editor/Previewer” at 5:58-6:14 and 7:10-67, Marcus156’s system would be enhanced, according to the teachings of Rosen, to display one or more frames of the currently-playing video in a separate area of the display from the video. The

system would be further enhanced to allow the user to select the frame in the concurrently displayed frame, for example by tapping or clicking, and to use that selection, in combination with all of the existing templates, tags, and other data in the database, to perform a second search for relevant video clips, which would then be displayed in the same manner that Marcus156 displays the results of an initial user query. The resulting system would practice all limitations of claims 1-20.

B. Motivation to Combine

274. In my opinion, a POSITA would have been motivated to modify Marcus156's system according to the teachings of Rosen for at least the following reasons.

275. *Showing frame content provides more information to the user:* Providing more concurrent information to the user would be commercially valuable given the typical user preference for increased information density. Adding the concurrent display would have improved the user interface by allowing the user to view more relevant information at one time. Displaying a still frame alongside the running video enables enhanced precision in editing because it allows editors to examine specific moments in detail while maintaining context within the overall video.

276. *Showing frame content time-shifts the display of information to prevent a user from missing critical information:* Displaying lagging frames in a

separate, concurrent window would give the user an expanded chance to click the right frame (or area within that frame), even if they missed it when the frame initially was displayed as part of the video. It would increase workflow efficiency by allowing editors to make decisions about specific frames without repeatedly pausing and resuming the video playback.

277. *Showing a frame, allowing a user selection, and performing a second search allows the video programming to be progressively refined:* The interaction pattern in the combined system would have resulted in superior video programming delivered to the user, because it would have allowed the video programming to be progressively refined with each user selection of a frame. Marcus156 describes this as an advantage of its techniques, and it would also be an advantage of the combined system, which would motivate a POSITA to make the combination. Marcus156 at 14:55-58 (“This interactive generation of programming for passive viewing contrasts with existing systems that include a predetermined loop if the viewer does not provide inputs.”). Among other advantages, progressive refinement of a result set through user selections of items from the previous result set would increase the accuracy and pertinence of the video programming.

278. *Similar field of art:* Both Marcus156 and Rosen pertain to video creation and editing systems. Both systems describe video playback. Given the

commonalities of the references, the teachings of Rosen would have been reasonably pertinent to a POSITA seeking to improve Marcus156's system.

279. ***Design choice:*** There are many different ways to organize a user interface for a video editing or consumption systems. Modifying Marcus156's user interface to present a frame and allow user selection of it, causing a subsequent search, would have been within a realm of minor design decisions that a POSITA would naturally make during the coding of such a system. Adding concurrent display of a video and a frame would comport with Marcus156's instructions that "[t]he user interface is preferably . . . graphical in nature." *Id.* at 6:11-12.

C. Reasonable Expectation of Success

280. In my opinion, a POSITA would have had a reasonable expectation of success in modifying Marcus156's system according to the teachings of Rosen, in the manner discussed above, for at least the following reason:

281. ***Software modification would be predictable:*** The modification can be implemented entirely in software and does not require any hardware modifications. Given the generally-acknowledged predictability of the software arts, a POSITA would have had a reasonable expectation of success in making the proposed modification. For example, Rosen employs well-understood XML markup. *E.g.*, Rosen at [0035]-[0161]. More generally, techniques for extracting and displaying individual frames from a video stream were well-established in the field of video

processing as of the time of the invention. Video codecs and container formats were already highly standardized, making it straightforward to access and manipulate video data consistently. For example, Marcus156 discusses and employs the standard MPEG video file format, for which multiple tools existed as of the time of the invention to extract frames, such as FFmpeg. Moreover, this type of dual display setup was already in wide use in video editing systems such as Avid at the time of the invention of the '185 patent. It was also in common usage in other systems such as picture-in-picture TV displays and in sports replay systems. Display technologies and graphics libraries available at the time of the invention readily supported multiple viewport rendering, making it feasible to show different content in separate areas of the screen, without much programming work. Similarly, most GUI frameworks used in video editing software at the time of the invention of the '185 patent supported dynamic layout changes and multiple widget areas, facilitating the addition of a new display area. These libraries and frameworks were supported by extensive documentation, open-source projects, and developer communities which provided resources and potential solutions for common challenges in video software development.

282. *Marcus156 already teaches extraction of a frame from a video, display of still images, and receiving a user selection of an item on a screen:* The system described in Marcus156 already encompasses techniques for extracting a

frame from a video. Marcus156 at 7:54-57 (“The first frame of each segment is decoded into separate buffers and then the transition function is applied to each matching pixel of both images to produce the third rendered frame buffer.”). And it teaches that its techniques can be used with still images (like frames). *Id.* at 13:62-65 (“For example, a database can be constructed for use with still photographs to be arranged into a program with an audio track. Interactive assets, such as two-dimensional or three-dimensional models and graphical user interfaces can also be catalogued in a database of this type.”); cl. 13 (“wherein said media elements are still photographs, and said media programming comprises a series of said still photographs”). Marcus156’s system also already processes user selections of displayed elements on a screen. *Id.* at 14:37-43 (“For example, the system may provide delivery of a predetermined audiovisual program so long as the viewer does not seek to interact with the system, e.g., by positioning a mouse pointer on the screen and clicking. At this point, the system may select assets that are appropriate to the particular information that the viewer appears to be seeking based on the information on the screen at the time.”). Because these capabilities are already disclosed by Marcus, a POSITA would have been able to implement a concurrent display of a video and a frame from that video, where the user can select the frame, with a reasonable expectation of success. Additionally, the computing hardware typically in use at the time of the invention (*i.e.*, computer 20) had more than

sufficient processing power and memory to handle simultaneous video playback and still frame display. Well-established optimization techniques for video processing and display could have been employed to ensure smooth performance of the concurrent display feature. The feature could be implemented incrementally, starting with basic functionality and refining it over time, reducing development risks.

D. All Claims and Claim Elements Other Than 1[g], 1[h], and 1[i]

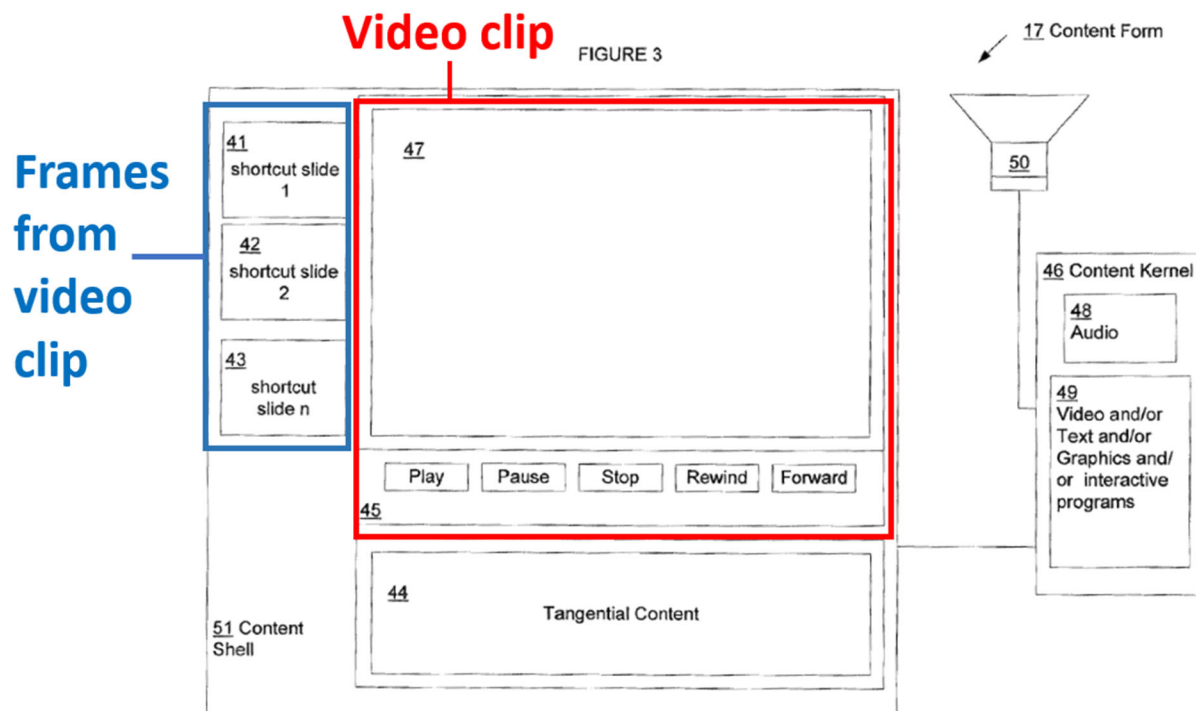
283. In my opinion, all claims 1-20 and claim elements are disclosed or rendered obvious by Marcus156 for the reasons given above in ground 1. No other reference is required in combination with Marcus156 in order to disclose or render obvious claims 1-20.

284. However, for the purposes of this ground I have been asked to consider as an additional theory a ground where I assume that Marcus156's disclosures are insufficient to teach elements 1[g] (concurrent display of a frame from the displayed first video clips in a second area of a display), 1[h] (receiving a user selection corresponding to a frame), and 1[i] (selection of a second video clip based on associations between the second clip and the video containing the selected frame). That I make this assumption in the context of ground 2 should not suggest that I have accepted that those elements are necessarily missing from Marcus156.

E. Rosen Discloses Element 1[g]: “concurrently with said presenting said first video programming, presenting one or more frames

from one or more displayed first video clips from the first video programming in a second area of said display;”

285. *Concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display:* Rosen’s user interface displays a playing video in main window 47 and shortcut slides 41-43 next to it:



Rosen at Fig 3 (annotated), [0023], [0025], Fig. 3.

286. The shortcut slides contain frames of the playing video. *Id.* at [0027] (“At any desired point in the video, the author may choose and extract a playback point P0 (or *frame*) from the video 64. The playback of the video during authoring

is the paused 65 and a shortcut slide in the content shell 51 is associated with the playback point P0. A still image of the video at the playback point is captured 66 and the shortcut slide in the content shell 51 is filled with the captured image.”) (emphasis added); [0028]. Therefore, in my opinion, Rosen discloses the concurrent display of a playing video and a frame from that video, as required by the claim language.

F. Rosen Discloses Element 1[h]: “receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;”

287. *Receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display:* In Rosen’s system, the user may select a shortcut slide to initiate an action. *Id.* at [0023] (“The predetermined images in the shortcut boxes 41-43 are selectable by a user or otherwise ‘activated automatically’ (if so scripted) and may initiate an event. For example, selecting an image may cause a ‘jump’ to a particular scene in the video or may activate other multimedia content in the main window.”); [0027] (“The author may also associate tangential content (such as text or a clipped video segment) with the added shortcut slide. As slide event is then chosen 68 for activation of the shortcut slide. For example, a shortcut slide may be activated during execution if a user clicks on it with a mouse or uses some other

input method.”). Thus, in my opinion, Rosen discloses the computing device receiving a user selection corresponding to a frame of the playing video.

G. Rosen Discloses Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection;”

288. *Selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection:* As described immediately above, Rosen discloses that its system can process a user selection of a shortcut slide, and that can be used to trigger a wide variety of different actions. *Id.* at [0031] (“A specific event is then chosen 114 for the shortcut box. The author may choose from a variety of event paths that will execute at the point P1 during video playback in the content shell. Exemplary event paths may include, but are not limited to, the appearance of the still image of the video 119 taken at P1, the appearance of a predetermined image 118, an interactive text box 117, another video 116, or an audio program 115 standing alone or in combination with any other event path or a web browser.”). In the context of Marcus¹⁵⁶’s system, the teachings of Rosen would allow an additional supplemental search to be performed by the system when a frame is selected by the user.

VII. Ground 3: Marcus156 and Sull Render Claims 1-20 Obvious

289. In my opinion, Marcus156 as a base reference, with modifications according to the teachings of Sull, renders challenged claims 1-20 obvious.

A. Specific Proposed Modification to Marcus156's System

290. For either the “interactive system” described in Marcus156 at 14:35-67 or the “List Editor/Previewer” at 5:58-6:14 and 7:10-67, Marcus156's system would be modified, according to the teachings of Sull, to display one or more frames of the currently-playing video in a separate area of the display from the video. The system would be further modified to allow the user to select the frame, for example by tapping or clicking, and to use that selection, in combination with all of the existing templates, tags, and other data in the database, to perform a second search for relevant video clips, which would then be displayed in the same manner that Marcus156 displays the results of an initial user query. The resulting system would practice all limitations of claims 1-20.

B. Motivation to Combine

291. A POSITA would have been motivated to modify Marcus156's system according to the teachings of Sull for at least the following reasons.

292. *Showing frame content provides more information to the user:* Described above in ground 2 and fully applicable here because the motivations are not unique to the combination of Marcus156 and Rosen.

293. *Showing frame content time-shifts the display of information to prevent a user from missing critical information:* Described above in ground 2 and fully applicable here because the motivations are not unique to the combination of Marcus156 and Rosen.

294. *Showing a frame, allowing a user selection, and performing a second search allows the video programming to be progressively refined:* Described above in ground 2 and fully applicable here because the motivations are not unique to the combination of Marcus156 and Rosen.

295. *Similar field of art:* Both Marcus156 and Sull pertain to video creation and editing systems. Both systems describe video playback. Given the commonalities of the references, the teachings of Sull would have been reasonably pertinent to a POSITA seeking to improve Marcus156's system.

296. *Design choice:* Described above in ground 2 and fully applicable here because the motivations are not unique to the combination of Marcus156 and Rosen.

C. Reasonable Expectation of Success

297. A POSITA would have had a reasonable expectation of success in modifying Marcus156's system according to the teachings of Sull, in the manner discussed above, for at least the following reason:

298. *Software modification would be predictable*: Described above in ground 2 and fully applicable here because the considerations are not unique to the combination of Marcus156 and Rosen.

299. *Marcus156 already teaches extraction of a frame from a video, display of still images, and receiving a user selection of an item on a screen*: Described above in ground 2 and fully applicable here because the considerations are not unique to the combination of Marcus156 and Rosen.

D. All Claims and Claim Elements Other Than 1[g], 1[h], and 1[i]

300. In my opinion, all claims 1-20 and claim elements are disclosed or rendered obvious by Marcus156 for the reasons given above in ground 1. No other reference is required in combination with Marcus156 in order to disclose or render obvious claims 1-20.

301. However, I have been asked to consider as an additional theory a ground here where I assume that Marcus156's disclosures are insufficient to teach elements 1[g] (concurrent display of a frame from the displayed first video clips in a second area of a display), 1[h] (receiving a user selection corresponding to a frame), and 1[i] (selection of a second video clip based on associations between the second clip and the video containing the selected frame). That I make this assumption in the context of ground 3 should not suggest that I have accepted that those elements are necessarily missing from Marcus156.

- E. Sull Discloses Element 1[g]: “concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display;”

302. *Concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display:* Sull discloses that a playing video is displayed concurrently, next to frames from that video that represent multimedia bookmarks. This is depicted in Figure 9:

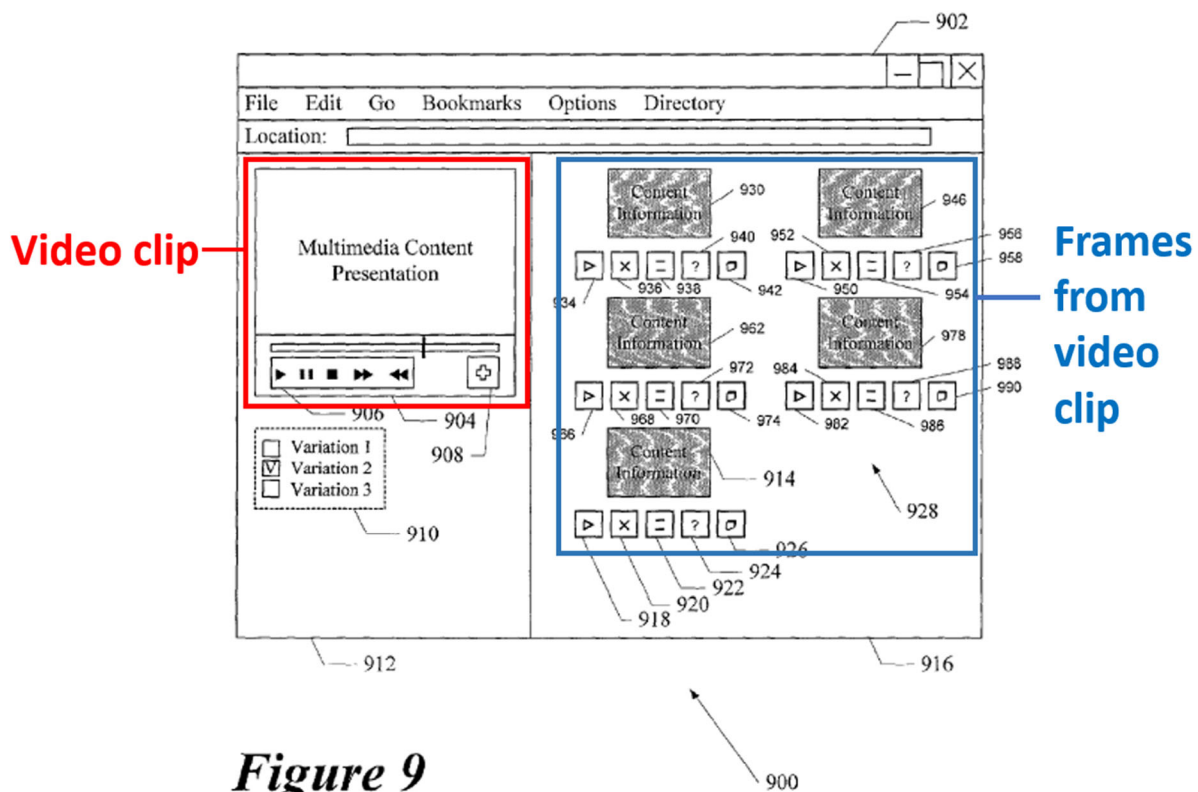


Figure 9

Sull at Fig. 9 (annotated); [0205] (“Fig. 9 shows an example of a user interface incorporating the multimedia bookmark of the present invention. The user interface

900 is composed of a playback area and a bookmark list 916.”); *id.* (“[T]he playback area 912 is also composed of a multimedia player 904 [which] provides various buttons 906 for normal VCR (Video Cassette Recorder) controls such as play, pause, stop, fast forward and rewind.”); *id.* (“[I]n the bookmark list 916, the saved bookmark is visually displayed with its content information. For example, a spatially reduced thumbnail image corresponding to the temporal location of interest saved by a user in the case of a multimedia bookmark is presented to help the user to easily recognized the previously bookmarked content of the video.”); [0220] (“Examples of visual features for the multimedia bookmark include, but are not limited to, captured frames in JPEG image compression format or color histograms of the frame.”).

F. Sull Discloses Element 1[h]: “receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;”

303. *Receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display:* Sull discloses that a user may select one of the displayed multimedia bookmarks (frames) and that the selection may be received at a computing device. *Id.* at [0194] (“The user selects one of the saved multimedia bookmarks from the user’s bookmark list.”); [0225] (“When a user at a client side

selects a bookmarked image, the client sends the image data to the server as a query frame.”).

- G. Sull Discloses Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection;”**

304. *Selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection:* Sull discloses that the user selection of the multimedia bookmark (frame) is communicated to a computing device and that a search for additional video clips can be performed based on the information. *Id.* at [0225] (“When a user at a client side selects a bookmarked image, the client sends the image data to the server as a query frame.”); [0221] (“[A] content-based multimedia search is performed by using the audio-visual features of the bookmark as query input in step 1518. The result of step 1518 is that the resulting multimedia segments have audiovisual features similar to the query.”); [0024] (“For a query image/video clip given by a user, these methods search the databases for the images/videos that are most similar to the query. In other words, the goal of the image/video search is to find the best matches to the query image/video from the database.”); [0060] (“The present invention also provides a method for searching for relevant multimedia content based on at least one feature saved in a multimedia bookmark. The method preferably includes transmitting at least one feature saved

in a multimedia bookmark from a client system to a server system in response to a user's selection of the multimedia bookmark. The server may then generate a query for each feature received and, subsequently, use each query generated to search one or more storage devices. The search results may be presented to the user upon completion.”); [0214] (“The content information of a particular segment, stored in a multimedia bookmark, may be used to find other relevant information about the particular segment. For example, a frame-based video search may be employed to find other video segments similar to the particular video segments.”). Therefore, a POSITA would have understood Sull to disclose this element.

VIII. Ground 4: Marcus156, Rosen, and Hosken Render Claims 1-20 Obvious

305. In my opinion, Marcus156 as a base reference, with modifications according to the teachings of Rosen and Hosken, renders challenged claims 1-20 obvious.

A. Specific Proposed Modification to the Combined System of Marcus156 and Rosen

306. The combined system of Marcus156 and Rosen would be further modified to perform additional searches for video clips (using the existing capabilities of Marcus156) based on the user selection of a displayed frame.

B. Motivation to Combine

307. In my opinion, a POSITA would have been motivated to modify the combined system of Marcus156 and Rosen according to the teachings of Hosken for at least the following reasons.

308. *Improved video clip recommendations:* A POSITA would have been motivated to add Hosken’s techniques to the combined system of Marcus156 and Rosen because progressive refinement of a result set through user selections of items from the previous result set would increase the accuracy and pertinence of the video programming. Hosken at 4:17-20 (“An advantage of the present invention, therefore, is that the provided combination of content and collaborative recommendation systems enables the delivery of recommendations that are particularly tailored to the personalized interests of a user.”). The use of stored associations between video clips as disclosed in both Marcus156 and Hosken would additionally improve the relevance of the clips selected for the second video programming.

309. *Similar field of art:* The ’185 patent, Marcus156, and Hosken all pertain to automated video recommendation systems. *E.g.*, Hosken at 4:9-11 (“Thus, the operation of the present system reflects the consideration that media content items, such as music, video, and other forms of content, can be interrelated based on multiple characterizing attributes.”); 28:25-27 (“[T]he present invention can also be

used to recommend other items, such as videos, digital music (e.g. MP3 files), television shows, books and other consumer entertainment media content.”).

C. Reasonable Expectation of Success

310. A POSITA would have had a reasonable expectation of success in modifying the combined system of Marcus156 and Rosen, further according to the teachings of Hosken, in the manner discussed above, for at least the following reason:

311. *Software modification would be predictable:* Described above in ground 2.

312. *Marcus156 already teaches some degree of successive searching:* As described above in ground 1, Marcus156 already discloses and renders obvious this claim element. To the extent it does not, its close similarity with the claim language would have motivated a POSITA add progressive search refinement to Marcus156 because all of the building blocks are already present. *E.g.*, Marcus156 at 14:40-45 (“At this point, the system may select assets that are appropriate to the particular information that the viewer appears to be seeking based on the information on the screen at the time. The system will then generate suitable transitions and add those assets to the programming.”).

D. All Claims and Claim Elements Other Than 1[i]

313. In my opinion, all claims 1-20 and claim elements other than element 1[i] (selection of a second video clip based on associations between the second clip and the video containing the selected frame) are disclosed or rendered obvious by Marcus156 for the reasons given above in ground 1, and by Marcus156 and Rosen for the reasons given above in ground 2. For this ground, I have been asked to consider a hypothetical scenario where 1[i] is not rendered obvious by Marcus156 and Rosen. I note that this hypothetical should not be read to assume that I believe that the combination of Marcus156 and Rosen does not render obvious 1[i].

E. Hosken Discloses Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection;”

314. *Selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection:* Hosken discloses performing a series of searches, where a user selection of one item in the previous search results is used in combination with “associations” between content items, to produce another search set. *E.g.*, Hosken at 9:4-8 (“Based on the review and consideration of the recommendation set, including as applicable any presented content samples, the user may request a further search and production of a new recommendation set, typically identifying a prior recommendation media content

item as part of the request”); 10:13-16 (“Input requests from a user 38, such as requests to find media content in some way similar to an identified media content item, are submitted for processing through system processes 42 to produce a responsive recommendation set back to the [user] 38.”); 11:9-21 (“In the first case, a user 38 selects a media content item known and of interest to the user from a master list of media content items. The selection is submitted to the system processes 42 for autonomous consideration against those identifiable through the content relations database 26 *that are linked by some association*, such as particularly or cumulatively considered characterizing attributes, to the media content item selected by the user 38. . . . As the product of this autonomous consideration, the system processes 42 produce a set of media content items that, as considered, have the strongest relationship connections to the user selected media content item or items based on this specific evaluation approach.”) (emphasis added). The associations between content items are stored in Hosken’s system, just as they are in Marcus156’s system. Hosken at 7:20-22 (“Explicit profiling data provided by the user and implicitly derived from referral system 20 processes are preferably processed 22 and stored 24 by the server system 18.”); 9:21-23 (“The implicit information gathered from user actions is preferably processed 22 and stored as an addition and refinement of the profile data 24 previously stored.”).

IX. Ground 5: Marcus156, Sull, and Hosken Render Claims 1-20 Obvious

315. In my opinion, for the same reasons given above in the context of grounds 3 and 4, the combined system of Marcus156, Sull, and Hosken would render obvious claims 1-20. This ground is similar to ground 4 except that instead of Rosen, this ground uses Sull. In my opinion, nothing in Sull's techniques would preclude also using Hosken's techniques. The motivations to combine and the reasonable expectations of success that I detailed above in grounds 3 and 4 are also applicable here.

X. Ground 6: Marcus156, Rosen, Hosken, and Byers Render Claims 8 and 16 Obvious

316. In my opinion, Marcus156 as a base reference, with modifications according to the teachings of Rosen, Hosken, and Byers renders challenged claims 8 and 16 obvious.

A. Specific Proposed Modification to the Combined System of Marcus156, Rosen, and Hosken Using Byers

317. The combined system of Marcus156, Rosen, and Hosken described in ground 4 would be further modified according to the teachings of Byers to include advertising clips in the video programming, including by dynamically editing other clips. Part of this modification would entail adding advertising material to the clip database and tracking additional advertising-related parameters in Marcus156's

templates. The result would be that “one or more video clips comprise advertising,” as required by claims 8 and 16 of the ’185 patent.

B. Motivation to Combine

318. *Commercial benefit to advertising*: Because insertion of advertising into video programming would enable a new revenue stream for the video creator, a POSITA would have had an economic motivation to do it. *E.g.*, Byers at 2:51-52 (“Such placement can lead to a boon in sales and publicity for the products so advertised.”), Marcus156 at 3:52-57 (“At this level of organization, the behaviors of ancillary assets such as promotional segments, information identifying the system, advertisements and news-flashes are defined. These assets embody aesthetic, program or instructional design, as well as market-driven, or viewer defined behaviors.”).

319. *Commercial benefit to user-targeted advertising*: The user-targeted recommendation techniques disclosed in Marcus156, Hosken, and Byers would all allow improved advertising that is customized to the demographics, interests, or other facets of the user, thus increasing the pertinence and, ultimately, the click-through or conversion rate of the targeted advertising. Byers at 2:57-67 (“One problem with product placement is that the placement is not tailored for an individual user. . . . Utilizing the present invention alleviates this problem by facilitating the placement of replacement images that are related to a user profile . . .”).

320. ***Same field of art:*** Like Marcus156 and the '185 patent, Byers is directed to creation and presentation of dynamic video programming. *E.g.*, Byers at 1:40-43 (“Therefore, a need exists for a method and apparatus for dynamically altering digital video images to allow customization of a digital video stream based upon characteristics associated with a user.”).

C. Reasonable Expectation of Success

321. ***Software modification would be predictable:*** As with the other reference combinations, this change would be accomplished in software and would entail predictable changes to Marcus156’s system.

322. ***Byers provides explicit instructions on how to modify clips to insert advertising:*** Byers discloses underlying mechanisms to insert the advertising content. *E.g.*, Byers at 3:67-4:5 (“[S]ervice node 103 receives the digital video stream, determines a replacement image to insert in the digital video stream, performs target geometry transforming and lighting model, digitally alters the digital video stream by inserting the replacement image, and transmits the altered video image to broadband access network 105.”); 5:60-66 (“DSP 209 preferably accomplishes the replacement of the original element by calculating the geometry of the replacement image, determining the lighting present in the original image, determining the transparency of the replacement area in the original image, determining the object order and motion, and overwriting selected portions of the

frame buffer that include the original video image.”). The hardware used in this process is similar to Marcus156’s disclosed hardware. Byers at 5:45-67; Marcus156 at 2:51-56. Marcus156’s disclosed hardware could perform Byers’ video editing algorithms.

323. *Byers’ instructions on adding advertising parallel Marcus156’s instructions on how to modify clips to insert transitions:* Marcus156 provides instructions to the user on extracting frames from a video clip and modifying them dynamically according to Marcus156’s templates to add transitions. Marcus156 at 7:47-63. The same technique could be used to add advertising information.

D. Byers Discloses and Renders Obvious “One or More Video Clips Comprise Advertising”

324. Although Marcus156 either discloses or renders obvious this limitation, as described above in the context of ground 1 and claims 8 and 16, to the extent it does not, it would have been obvious to further modify the combined system of Marcus156, Rosen, and Hosken to use Byers’ advertising techniques.

325. Byers discloses an “advertiser database” that contains advertising-related information. Byers at 4:19-25 (“Advertiser database 113 includes advertising information that relates to advertisers. Such information is preferably organized to allow the retrieval of digital images from image database 117 that can be inserted into the digital video stream. For example, advertiser database can include information relating to thresholds that advertisers are willing to utilize.”).

326. The information in the database is used to dynamically insert ads into a video stream. *Id.* at 2:48-51; 3:1-4 (“A replacement image for a product, such as a product that an actor is using in a movie or television show, can be altered for each desired user based upon factors associated with the intended recipient of the video stream.”).

XI. Ground 7: Marcus156, Sull, Hosken, and Byers Render Claims 8 and 16 Obvious

327. For the same reasons given above in the context of grounds 5 and 6, it is my opinion that the combined system of Marcus156, Sull, Hosken, and Byers would render obvious claims 8 and 16, which specify that “one or more video clips comprise advertising.” This ground is similar to ground 6 except that instead of Rosen, this ground uses Sull. In my opinion, nothing in Sull’s techniques or Hosken’s techniques would preclude also using Byers’ techniques. I do not rely on any disclosure in Rosen or Sull to establish the obviousness of claims 8 or 16, to support a motivation to combine, or to show a reasonable expectation of success.

XII. Ground 8: Ferman Renders Claims 1-7, 9-15, and 17-20 Obvious

328. In my opinion, Ferman alone renders challenged claims 1-7, 9-15, and 17-20 obvious. Unless otherwise noted below, it is my opinion Ferman discloses each claim and claim element.

A. Claim 1

1. Element 1[pre]: “1. A method of providing video programming to a user, comprising the steps of:”

329. In my opinion, Ferman discloses this element.

330. *A method of providing video programming to the user:* Ferman’s system provides customized interactive video programming to a user. Ferman at [0092] (“First, referring to FIG. 1 the video, image, and/or audio information 10 is provided or otherwise made available to a user and/or a (device) system. Second, the video, image, and/or audio information is presented to the user from the system 12 (device), such as a television set or a radio. Third, the user interacts both with the system (device) 12 to view the information 10 in a desirable manner and has preferences to define which audio, image, and/or video information is obtained in accordance with the user information 14.”); [0106] (“The intelligent management and consumption of audiovisual information using the multi-part description stream device provides a next-generation device suitable for the modern era of information overload. The device responds to changing lifestyles of individuals and families, and allows everyone to obtain the information they desire anytime and anywhere they want.”).

2. Element 1[a]: “searching, by one or more computing devices, one or more databases of video clips in response to a user request wherein each video clip comprises a plurality of frames,”

331. In my opinion, Ferman discloses this element.

332. ***One or more databases of video clips:*** Ferman's system includes a program description scheme, which is a database of video clips. *Id.* at [0114] ("In addition, the program description scheme and system description scheme in an advanced VCR or other system will enable the user to browse, search, and filter audiovisual programs."); [0117] ("The search engine uses the information contained in the program description scheme to retrieve programs on the basis of their content."); [0264] ("Referring to FIG. 24, the device may use the program descriptions provided by any suitable source describing the current and/or future audio and/or video content available from which a filtering agent 520 selects the appropriate content for the particular user(s). The content is selected based upon the usage preference description for a particular user identification(s) to determine a list of preferred audio and/or video programs."); [0266] ("[P]referred programs can be determined as a result of filtering program descriptions according to user's filtering preferences."); [0315] ("Referring to FIG. 31, the filter agent 600 receives the user preference description 604 and interprets the information contained within the user preference description 604 at block 610 using information from a mapping table 614. The filter agent 600 also builds a model, such as a combinatorial model, of the user criteria indicating the desired user criteria at block 612 using information from the mapping table 614. The resulting model or otherwise set of criteria, is then applied

against the available program descriptions 602 at block 616 to select the desired programs.”).

333. *Searching...in response to a user request:* Ferman’s system searches the program description scheme in response to a user request. *Id.* at [0114] (“In addition, the program description scheme and system description scheme in an advanced VCR or other system will enable the user to browse, search, and filter audiovisual programs.”); [0264] (“Referring to FIG. 24, the device may use the program descriptions provided by any suitable source describing the current and/or future audio and/or video content available from which a filtering agent 520 selects the appropriate content for the particular user(s). The content is selected based upon the usage preference description for a particular user identification(s) to determine a list of preferred audio and/or video programs.”); [0266] (“[P]referred programs can be determined as a result of filtering program descriptions according to user’s filtering preferences.”).

334. The user request may be triggered by an interaction or may be a user preference description or a set of filtering preferences. *Id.* [0315] (“Referring to FIG. 31, the filter agent 600 receives the user preference description 604 and interprets the information contained within the user preference description 604 at block 610 using information from a mapping table 614. The filter agent 600 also builds a model, such as a combinatorial model, of the user criteria indicating the

desired user criteria at block 612 using information from the mapping table 614. The resulting model or otherwise set of criteria, is then applied against the available program descriptions 602 at block 616 to select the desired programs.”); [0266] (“[P]referred programs can be determined as a result of filtering program descriptions according to user’s filtering preferences.”).

335. *Searching . . . by one or more computing devices:* The search is performed by a computing device running a software agent. *Id.* at [0094] (“The user description scheme provides information to a software agent that in turn performs a search and filtering on behalf of the user by possibly using the system description scheme and the program description scheme information.”); [0092] (“First, referring to FIG. 1 the video, image, and/or audio information 10 is provided or otherwise made available to a user and/or a (device) system. Second, the video, image, and/or audio information is presented to the user from the system 12 (device), such as a television set or a radio. Third, the user interacts both with the system (device) 12 to view the information 10 in a desirable manner and has preferences to define which audio, image, and/or video information is obtained in accordance with the user information.”); Fig. 1. Components of the computing device are shown in Figure 2:

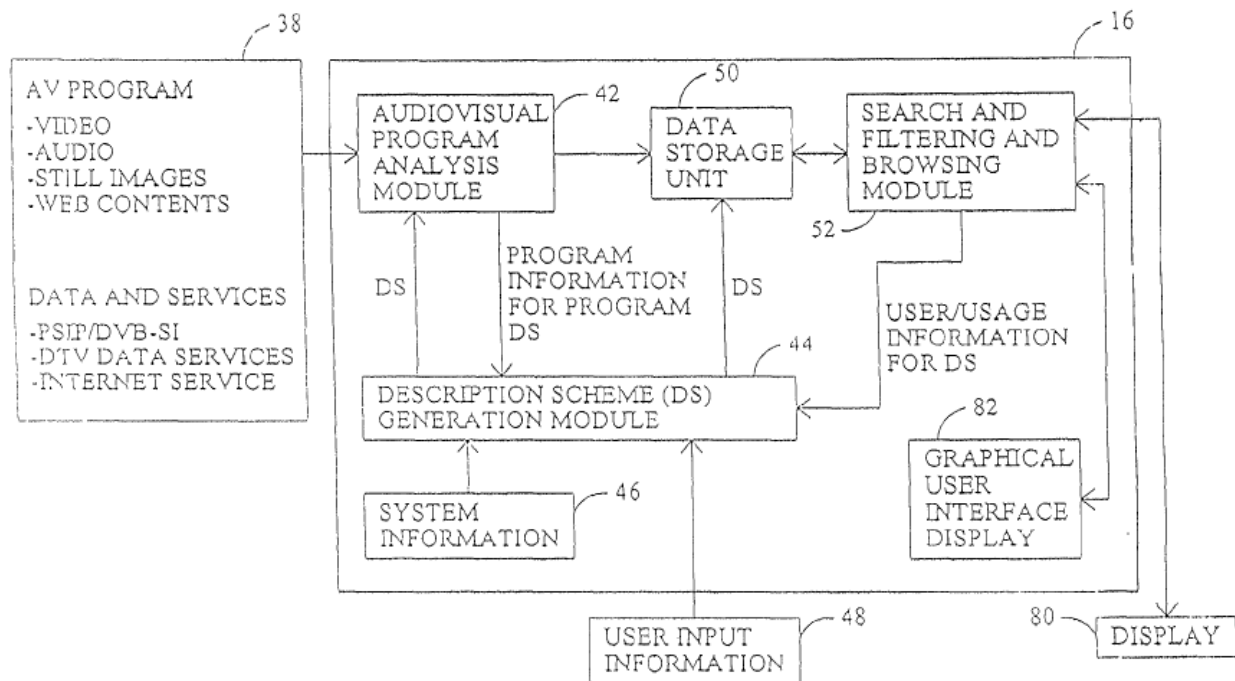


FIG. 2

Id. at Fig. 2; [0007] (“FIG. 2 is an exemplary embodiment of the audio-visual system, including an analysis module, of FIG. 1.”); [0102] (“Referring to FIG. 2, the audio, visual, or audio-visual program 38, is received by the system 16. . . . The system 16 may include any device(s) suitable to receive any one or more of such programs 38[.]”).

336. ***Wherein each video clip comprises a plurality of frames:*** The clips in the database are comprised of frames, like any other video. Ferman at [0094] (“The program views define logical structures of the frames of a video that define how the video frames are potentially to be viewed suitable for efficient browsing.”); [0107] (“The highlights may be generated by audio or video analysis, or the program

description scheme includes data indicating the frames that are presented for a 5 minute highlight.”); [0252] (“The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video.”).

3. Element 1[b]: “selecting, by one or more of said computing devices, a plurality of first video clips responsive to the user request, wherein each first video clip is associated with media content comprising one or more second video clips;”

337. In my opinion, Ferman discloses this element.

338. *Selecting, by one or more of said computing devices, a plurality of first video clips responsive to the user request:* Ferman’s system selects video clips that fulfill the user request. *Id.* at [0117] (“The search engine uses the information contained in the program description scheme to retrieve programs on the basis of their content.”); [0264] (“Referring to FIG. 24, the device may use the program descriptions provided by any suitable source describing the current and/or future audio and/or video content available from which a filtering agent 520 selects the appropriate content for the particular user(s). The content is selected based upon the usage preference description for a particular user identification(s) to determine a list of preferred audio and/or video programs.”); [0315] (“Referring to FIG. 31, the filter agent 600 receives the user preference description 604 and interprets the information contained within the user preference description 604 at block 610 using information from a mapping table 614. The filter agent 600 also builds a model, such as a

combinatorial model, of the user criteria indicating the desired user criteria at block 612 using information from the mapping table 614. The resulting model or otherwise set of criteria, is then applied against the available program descriptions 602 at block 616 to select the desired programs.”); Fig. 28 (“FILTERING AND SEARCH” component leads to “PREFERRED PROGRAMS”); [407].

339. From this initial set, based on the user preferences, Ferman’s system will select a subset of video clips corresponding to specific events or highlights. *Id.* at [0103] (“[I]f a user prefers to watch a 5 minute video highlight of a particular program, such as a basketball game, the analysis module 42 may invoke a knowledge based system 90 (FIG. 3) to determine the highlights that form the best 5 minute summary.”); [094] (“a key frame...providing multiple levels of summary ranging from coarse to fine.”); [0104] (“highlight generation functionality for a particular program”); [0107] (describing a user viewing session where some programs (“the Microsoft trial”) are presented un-edited and other programs (“sporting events”) are automatically summarized and culled into sub-segments). I note that Ferman refers to this functionality of breaking down a long-form program into subsequences as “filtering.” *Id.* at [0119].

340. The result of these operations is a set of recommended video clips that are presented to the user. *Id.* at [105] (“The SFB module 52 may perform filtering, searching, and browsing of the programs 38, on the basis of the information

contained in the description schemes.”). In my opinion, this constitutes the claimed “plurality of first video clips” in Ferman.

341. *Wherein each first video clip is associated with media content comprising one or more second video clips:* Each first video clip in is associated with a set of second video clips, including the full program, other clips that are part of the program but not included in a filtered set, or other, different filtered subsequences of clips from the same program. *Id.* at [0103] (“[I]f a user prefers to watch a 5 minute video highlight of a particular program, such as a basketball game, the analysis module 42 may invoke a knowledge based system 90 (FIG. 3) to determine the highlights that form the best 5 minute summary.”); [0104] (“[T]he user description scheme may be suitable for triggering the highlight generation functionality for a particular program and thus generating the preferred views and storing associated data in the program description scheme.”); [0107] (“The user is interested in basketball games and indicates a desire to view games. A set of title frames is presented on the display that captures an important moment of each game. The user selects the Chicago Bulls game and indicates a desire to view a 5 minute highlight of the game. The system automatically generates highlights. The highlights may be generated by audio or video analysis, or the program description scheme includes data indicating the frames that are presented for a 5 minute highlight. The system may have also recorded web-based textual information

regarding the particular Chicago-Bulls game which may be selected by the user for viewing. If desired, the summarized information may be recorded onto a storage device, such as a DVD with a label.”). In my opinion, these are the claimed “media content comprising one or more second video clips” in Ferman.

342. In my opinion, the associations are contained within at least the program view descriptions, program description scheme, and segment description scheme. [0104] (cited above, program description scheme); [0094] (program view descriptions); [0252] (segment description scheme).

4. Element 1[c]: “selecting, by one or more of said computing devices, a plurality of first transitional clips for association with at least some of the selected first video clips;”

343. *Selecting, by one or more of said computing devices, a plurality of first transitional clips for association with at least some of the selected first video clips:* Ferman’s system will filter and segment a video program “into story units, scenes, and shots.” *Id.* at [0252] (“For example, a video program may be segmented into story units, scenes, and shots, from which the segment description scheme 450 may contain such information as a table of contents.”). The set of segments includes multiple major events in the video clip, including events which transition between other clips. *Id.* at [0094] (“A program description scheme 18 related to the video, still image, and/or audio information 10 preferably includes two sets of information, namely, program views and program profiles. The program views define logical

structures of the frames of a video that define how the video frames are potentially to be viewed suitable for efficient browsing. For example the program views may contain a set of fields that contain data for the identification of key frames, segment definitions between shots, highlight definitions, video summary definitions, different lengths of highlights, thumbnail set of frames, individual shots or scenes, representative frame of the video, grouping of different events, and a close-up view.”); [0252] (“The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video. . . . Such a segment tree may be used to define[] a temporal structure of a video program.”); [0255] (“Referring again to FIG. 13, the semantic structure description scheme 404 is used to specify semantic features of a video program (or audio), e.g. semantic events. In a similar manner to the syntactic structure description scheme, the semantic structure description scheme 404 preferably includes three modules, namely an event description scheme 480, an object description scheme 482, and an event/object relation graph description scheme 484. The event description scheme 480 may be used to form relationships between different events of the video normally consisting of multiple frames of the video. An event description scheme 480 may contain another event description scheme 480 to form a segment tree. Such an event segment tree may be used to define a semantic index table for a video program. Multiple event trees may be created and thereby

creating multiple index tables. For example, a video program may include multipole events, such as a basketball dunk, a fast break, and a free throw, and the event description scheme may contain such information as an index table. The event description scheme may also contain references which link the event to the corresponding segments and/or regions specified in the syntactic structure description scheme. On[e] example of an event description scheme is shown in FIG. 19.”). In my opinion, a POSITA would have considered video segments that are between two other defined events or segments to be a “transitional clip.” In particular, a POSITA would have recognized that the set of sub-segments for given preferred program are related to each other as a sequence of events that pertain to the same overall program. *Id.* at [0094] (“grouping of different events”). Based on the user’s viewing preferences, a set of initial preferred programs can be supplemented by sub-segments from those programs that are automatically selected. In my opinion, a POSITA would consider these video segments, which are thematically and structurally related to other surrounding video clips, to be “transitional clips.”

344. Additionally, it is my opinion that, to the extent PO contends that these clips do not qualify as “transitional clips,” it would have been obvious for a POSITA to modify Ferman’s system to more explicitly select clips for inclusion in the final video programming based on their ability to transition between other events. In my

opinion, a POSITA would have been motivated to do so by at least Ferman's expressed desire to "include content suited to [the user's] particular viewing preferences." *Id.* at [0089]. Moreover, doing so would provide more compelling and cohesive video programming. In my opinion, a POSITA would have had a reasonable expectation of success in making any necessary changes to Ferman's system to more explicitly track transitions, including because Ferman already categorizes sub-sections of a video program by their contextual relationship to other sections. *Id.* at [0094].

345. ***Selecting . . . by one or more of said computing devices:*** For the same reasons that I gave above in the context of element 1[a], the selecting is done by the same computing devices that performs the searching step.

346. ***Selecting . . . for association with at least some of the selected first video clips:*** The selected transitional clips are associated with the first video clips via at least the segment description scheme 450. Ferman at [0252] ("The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video."). Additionally, in my opinion they are associated at least via the program description scheme. *Id.* at [094]; [0104] ("[T]he user description scheme may be suitable for triggering the highlight generation functionality for a particular program and thus generating the preferred views and storing associated data in the program description scheme.").

5. Element 1[d]: “determining a first temporal sequence for concatenating the selected first video clips and first transitional clips;”

347. In my opinion, Ferman discloses this element.

348. *Determining a first temporal sequence for concatenating the selected first video clips and first transitional clips:* The selection of video clips and transitional clips in turn determines a temporal sequence for the concatenated video programming. Ferman at [0252] (“The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video. . . . Such a segment tree may be used to define[] a temporal structure of a video program.”). Moreover, the video programming is sequenced based on information in the program description scheme and/or event description scheme. *Id.* at [0095] (sequencing program views using the program description scheme: “In the close-up view, a certain image object, e.g., a famous basketball player such as Michael Jordan, can be viewed up close by playing back a close-up sequence that is separate from the original program.”); [0255] (“The event description scheme 480 may be used to form relationships between different events of the video normally consisting of multiple frames of the video.”).

349. *For concatenating:* The responsive video clip and transitions are concatenated and presented as a stream of video programming to the user. *Id.* at [0607] (“PlayStream Play content from input stream”); [0107] (“The user selects the

Chicago Bulls game and indicates a desire to view a 5 minute highlight of the game.

The system automatically generates highlights. The highlights may be generated by audio or video analysis, or the program description scheme includes data indicating the frames that are presented for a 5 minute highlight.”); *id.* (“The user . . . invokes the summarization functionality of the system to scan the tape and prepare a summary. The user can then view the summary the next morning to quickly discover the baby’s looks, and playback segments between the keyframes of the summary to catch a glimpse of the crying baby.”); [0111] (“[O]ne can also personalize television viewing, for example, by plugging the smart card into a remote control that in turn will autonomously command the television receiving system to present the user information about current and future programs that fits the user’s preferences.”); [0114] (“The user then may choose frame view, shot view, key frame view, or highlight view, depending on their availability and user’s preference. These views can be readily invoked using the relevant information in the program description scheme, especially in program views. The user at any time can start viewing the program either in parts, or in its entirety.”). The system can generate a highlight video of a given length that consists of multiple subsidiary highlight or event clips concatenated and presented together. *Id.* at [0267] (15-minute highlight); Fig. 10; [0107] (5-minute highlight). A POSITA would have recognized that the act of

generating a single video highlight from multiple previously separate clips qualifies as “concatenating” under the plain and ordinary meaning.

350. It is also my opinion that, to the extent that PO may argue that concatenation of the video clips is not explicitly described in Ferman, it would have been obvious for a POSITA to modify Ferman’s system to record and allow viewing of archival copies of the sequence of programming viewed by a user, which would be a concatenated video stream. In my opinion, a POSITA would have been motivated to make the modification because it would allow its system to achieve better VCR-like functionality. It is additionally my opinion that a POSITA would have been motivated to make the modification also because concatenating clips would result in a more coherent viewing experience. Furthermore, it is my opinion that a POSITA would have had a reasonable expectation of success in making the modification, at least because Ferman’s system already has video recording capability. Ferman at [0107].

6. **Element 1[e]: “wherein said selecting, determining and concatenating is performed according to one or more first templates, to define first video programming, wherein said one or more first templates comprise a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and”**

351. In my opinion, Ferman discloses this element.

352. *Wherein said selecting, determining and concatenating is performed . . . to define first video programming:* As mentioned in the elements above, Ferman's system selects, determines, and concatenates clips to form a final video stream ("first video programming"). *Id.* at [0102]; [0105] (filtered video programming output by SFB module); [0107] (summarized highlight video programming output); [0111]; [0114].

353. To the extent PO contends that Ferman's system itself fails to define video programming consisting of video data, it is my opinion that it would have been obvious for a POSITA to modify Ferman's system to do so. A POSITA would have been motivated to do so to reduce interdependency among different software components and minimize the chance for service interruptions due to bandwidth congestion between a viewing device and a video source. A POSITA would have a reasonable expectation of success in doing so because techniques for combining and manipulating underlying video data (e.g., via transcoding) were well-known at the time of the invention.

354. *Wherein said selecting, determining and concatenating is performed according to one or more first templates:* These steps are performed according to at least a template referred to in Ferman as the "user description scheme." Ferman at [0097] ("The user description scheme 20 preferably includes the user's personal preferences, and information regarding the user's viewing history such as for

example browsing history, filtering history, searching history, and device setting history. The user's personal preferences includes information regarding particular programs and categorizations of programs that the user prefers to view. The user description scheme may also include personal information about the particular user, such as demographic and geographic information, e.g. zip code and age. The explicit definition of the particular programs or attributes related thereto permits the system 16 to select those programs from the information contained within the available program description schemes 18 that may be of interest to the user.”); [0097] (“[T]he user description scheme may be augmented or any technique can be used to compare the information contained in the user description scheme 20 to the available information contained in the program description scheme 18 to make selections. The user description scheme provides a technique for holding user preferences ranging from program categories to program views, as well as usage history.”); [0097] (“[T]he user description scheme is modular and portable so that users can carry or port it from one device to another, such as with a handheld electronic device or smart card or transported over a network connecting multiple devices.”); [0407] (“The filter agent 600 also receives user preference descriptions 608, typically from a user agent 610, which may be hierarchical if desired. The filter agent 600 based upon the user preference descriptions 608 and the program descriptions 602 provides selected

audiovisual 606, which may be rated if desired. The selected audiovisual materials may be provided to a user agent 610 and then to the user, if desired.”).

355. As previously reviewed above in the context of ground 1, element 1[e], the intrinsic evidence does not limit a “template” to a particular type of data structure. The user description scheme is a structured collection of data in XML format, which, in my opinion a POSITA would have understood to be a claimed “template.” *Id.* at [0196] (XML format); [0273] (“A description scheme is a data model of descriptions.”). It is further my opinion that the templates comprise a plurality of template variables, namely the data contained in the user description scheme. *Id.* at [0196]; [0273].

356. ***Wherein said one or more first templates comprise a plurality of variables corresponding to (a) information about the user:*** The user description scheme template includes data about the user. Ferman at [0108] (“[T]he system can store the user history and create entries in the user description scheme based on the’s audio and video viewing habits. In this manner, the user would never need to program the viewing information to obtain desired information. In a sense, the user descriptor scheme enables modeling of the user by providing a central storage for the user’s listening, viewing, browsing preferences, and the user’s behavior.”); [0333] (“This provides the user with a ranked output that they may select from. Also, user preferences may be ‘softened’ to pass programs that are near matches to the

specific preference criteria. This fuzzy approach may allow the user preference description to more clearly model the user's intended criteria.”); [0334] (“After the individual preferences are interpreted into individual preference tests, these tests may be combined into a single test that models the user's preferences. The preferred technique includes combining the individual preference tests according to their hierarchy. Each parent test becomes the combination of its children tests, and this continues up to the root preference, yielding in effect one composite test. The combination of ‘children’ tests within a single ‘parent’ may be broken down into the combination of similar tests and the combination of dissimilar tests. Similar tests may have the same name or otherwise be associated in the Mapping Table such as by being on the same row. Also, dissimilar tests may have different entries in the Mapping Table.”); [0394] (“The multiple User Preference elements may contain a ranking attribute. Such ranking attributes may be applied at each comparison test and each combinatorial operation, to yield a composite ranking score, this may be utilized to provide a sorted test result for the user.”); [0395] (“Referring to FIG. 48, the user preference hierarchy of (name:value) pairs may be supplemented with attributes regarding the intended combination of individual preference tests. As shown in FIG. 19, the supplementation may indicate that the two tests of type A should be AND'd together and this result OR'd with the other tests under this parent. The primary advantage of this supplemental enhancement is that the user preference

description may override the default behavior for specific tests. This makes interpretation of the user preference description more flexible so that it may be tailored to the user's specific preference criteria.”); [0407] (“Referring to FIG. 51, the input to a filter agent 600 may include program descriptions 602 relating to one or more audio, video, or audiovisual (collectively referred to as audiovisual without limitation) materials 604. The filter agent 600 also receives user preference descriptions 608, typically from a user agent 610, which may be hierarchical if desired. The filter agent 600 based upon the user preference descriptions 608 and the program descriptions 602 provides selected audiovisual materials 606, which may be rated if desired. The selected audiovisual materials may be provided to a user agent 610 and then to the user, if desired. The user agent 610 may create the user preference profile. Alternatively, the user or other party may utilize an authoring tool or automatic profile generator to create the user preference profile.”); [0410] (“The user preference description may include a hierarchy of preferences, many of which include preference value attributes. When each ‘individual preference’ is evaluated against the corresponding information from the program descriptions a score is calculated for that individual preference. In one embodiment, the hierarchy of preferences for an individual hierarchy may be evaluated by creating a composite score from the aggregation of individual scores within the hierarchy. The resulting composite score is then compared against other composite scores for other program

descriptions to determine a relative ranking. The composite score for a particular program description may be determined free from consideration of other program descriptions, if desired.”); cl. 1 (“(a) receiving user attribute information corresponding to user preferences wherein said user attribute information includes preferences”).

357. *Wherein said one or more first templates comprise a plurality of variables corresponding to . . . (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other:* The user description scheme includes viewing scenarios of multiple clips in a temporal relationship. Ferman at [0267] (“In the case of video, for example, the browsing preferences may include user’s preference for continuous playback of the entire program versus visualizing a short summary of the program. Various summary types may be described in the program descriptions describing multiple different views of programs where these descriptions are utilized by the device to facilitate rapid non-linear browsing, viewing, and navigation.”).

358. Moreover, other templates in Ferman’s system, including the segment description scheme and the semantic structure description scheme, include temporal viewing scenarios for clips. *Id.* at [0252] (“The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video. A segment description scheme 450 may

contain another segment description scheme 450 and/or shot description scheme to form a segment tree. Such a segment tree may be used to defined a temporal structure of a video program. Multiple segment trees may be created and thereby create multiple table of contents. For example, a video program may be segmented into story units, scenes, and shots, from which the segment description scheme 450 may contain such information as a table of contents.”); [0255] (“Referring again to FIG. 13, the semantic structure description scheme 404 is used to specify semantic features of a video program (or audio), e.g. semantic events. In a similar manner to the syntactic structure description scheme, the semantic structure description scheme 404 preferably includes three modules, namely an event description scheme 480, an object description scheme 482, and an event/objection relation graph description scheme 484. The event description scheme 480 may be used to form relationships between different events of the video normally consisting of multiple frames of the video. An event description scheme 480 may contain another event description scheme 480 to form a segment tree. Such an event segment tree may be used to define a semantic index table for a video program. Multiple event trees may be created and thereby creating multiple index tables. For example, a video program may include multiple events, such as a basketball dunk, a fast break, and a free throw, and the event description scheme may contain such information as an index table. The event description scheme may also contain references which link the event to

the corresponding segments and/or regions specified in the syntactic structure description scheme.”).

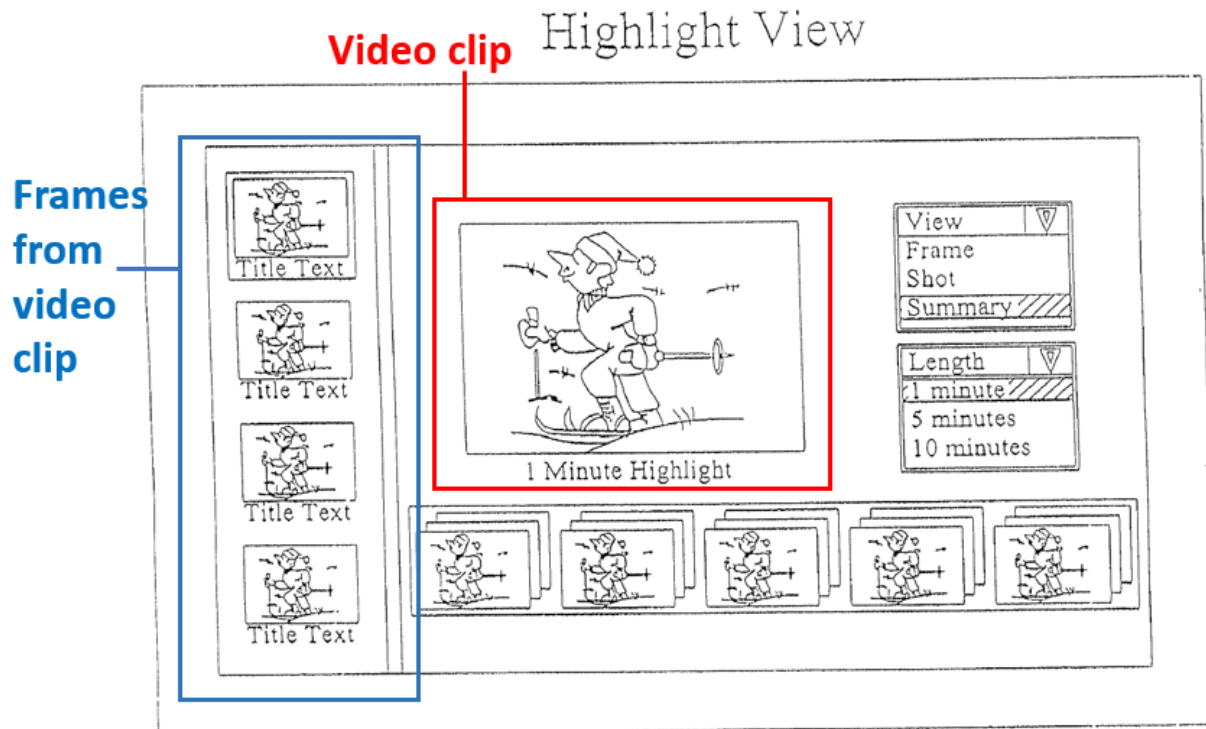
359. *Wherein said one or more first templates comprise a plurality of variables corresponding to . . . (c) style requirements:* The user description scheme template includes style requirements. *Id.* at [0318] (“[T]he User Preferences may describe the desired configuration for presentation, such as volume or any other functionality.”). Additionally, other templates such as the device preferences description also include style requirements. *Id.* at [0266] (“The device preferences description 534 relates to the preferences for setting up the device in relation to the type of content being presented, e.g. romance, drama, action, violence, evening, morning, day, weekend, weekday, and/or the available presentation devices. For example, presentation devices may include stereo sound, mono sound, surround sound, multiple potential displays, multiple different sets of audio speakers, AC-3, and Dolby Digital.”); [0268] (“favorite actors, directors”); *id.* (“favorite program category”); [0269] (“The device preferences description 534 contains descriptors describing preferred audio and/or video rendering settings, such as volume, balance, bass, treble, brightness, contrast, closed captioning, AC-3, Dolby digital, which display device of several, type of display device, etc. The settings of the device relate to how the user browses and consumes the audio and/or video content.”); [282]

(“the preferred theme, or point of view, of a segment...annotated with its theme or emphasis point”).

7. Element 1[f]: “presenting said first video programming in a first area of a display associated with one of said computing devices;”

360. In my opinion, Ferman discloses this element.

361. *Presenting said first video programming in a first area of a display associated with one of said computing devices:* In Ferman’s system, a playing video may be presented next to multiple frame images from the program, for example when a user desires to see the highlight view or event view associated with a program. Ferman at [0250] (“The key frame view description scheme 412 preferably includes a level indicator 426 and a time reference 428. The level indicator 426 accommodates the presentation of a different number of key frames for the same video portion depending on the user's preference. The highlight view description scheme 414 includes a length indicator 430 and a time reference 432. The length indicator 430 accommodates the presentation of a different highlight duration of a video depending on the user's preference. The event view description scheme 416 preferably includes an event indicator 434 for the selection of the desired event and a time reference 436.”). For example, at playing video in the “Highlight View” screen is shown next to frames from that video:



Id. at Fig. 10 (annotated); [0124] (“The thumbnail views are short video segments that provide an indication of the content of the respective actual program that it corresponds with. . . . [A] frame view of particular programs may be displayed, if desired.”); *id.* (“a set of representative frames of different programs in the left hand column”); [0119] (“[T]he description scheme will enable audiovisual browsing and searching of programs and enable filtering within a particular program by supporting multiple program views such as the highlight view.”). The event view has similar functionality. *Id.* at Fig. 11; [0124].

362. Moreover, to the extent PO argues there is any ambiguity in the Ferman’s disclosure on this point, it is my opinion that it would have been obvious

to implement the UI shown in Ferman's Fig. 10 to concurrently play the selected highlight video in the main UI element and display frames from the highlight videos to allow the user to play other versions of that highlight video. Ferman discloses playing highlight videos and displaying frames corresponding to the videos, as well as a user interface (*e.g.*, Fig. 10) that includes areas for both of those UI elements.

363. In my opinion, a POSITA would have been motivated to make the change because it would facilitate efficient browsing of highlight videos, since a user would be able to watch an initial version of the highlight video and then select frames to play other highlight videos of the same program that might focus on different aspects of the program. For example, in Ferman's example of highlight videos of a basketball game (*id.* at [0107]), the user might initially watch a highlight video of the most exciting scoring events and be presented with frames from that video corresponding to different types of scoring events for play, *e.g.*, a highlight video of dunks vs. a highlight video of 3-point shots. A POSITA would have recognized that this would increase the interactivity and pertinence of the videos generated by Ferman's system.

364. In my opinion, a POSITA would have had a reasonable expectation of success in any needed modification because Ferman already implements the necessary UI components and interaction paradigms. *Id.* at Figs. 10, 11. And Ferman already has the ability to generate highlight and event reel sequences of

video programming based on user interactions. *Id.* at [0107]. The change would be entirely in software. No hardware would need to be modified.

8. **Element 1[g]: “concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display;”**

365. *Concurrently with said presenting said first video programming, presenting one or more frames from one or more displayed first video clips from the first video programming in a second area of said display:* As discussed in the previous element, Ferman’s system displays running video alongside frames from the video, in distinct areas of the display.

9. **Element 1[h]: “receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display;”**

366. *Receiving, by one or more of said computing devices, a user selection corresponding to at least one of the frames of at least one first video clip presented in the second area of the display:* In Ferman’s user interface, a user can click on a frame to make a selection. *Id.* at [0304] (“The user may click and start playing back the video starting from that frame. The frame frequency value descriptor provides similar functionality in terms of shots of the video.”); [0309] (“The user may manually peruse program listing to select the desired material.”); [0105] (“The selections of the desired program(s) to be retrieved, stored, and/or viewed may be

programmed, at least in part, through a graphical user interface.”); [0107] (“The user starts interacting with the system with a pointer or voice commands to indicate a desire to view recorded sporting programs. . . . The user is interested in basketball games and indicates a desire to view games. A set of title frames is presented on the display that captures an important moment of each game. The user selects the Chicago Bulls game and indicates a desire to view a 5 minute highlight of the game. The system automatically generates highlights.”); [0114] (“[T]he program description scheme and system description scheme in an advanced VCR or other system will enable the user to browse, search, and filter audiovisual programs.”); [0152]; [0154]; [0250] (“The event view description scheme 416 preferably includes an event indicator 434 for the selection of the desired event and a time reference 436.”); Fig. 10.

367. To the extent that PO argues that Ferman does not explicitly describe receiving a user selection of a frame and using that selection to select a plurality of second video clips, it is my opinion that it would have been obvious to a POSITA to modify Ferman to implement that functionality for the reasons above in 1[f].

10. **Element 1[i]: “selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at least one first video clip indicated by said user selection;”**

368. *Selecting, by one or more of said computing devices, a second video clip based on at least one stored association between the second video clip and at*

least one first video clip indicated by said user selection: The user selections are used to populate the user description scheme and therefore to search for and select additional video clips. Ferman at [0097] (“The user description scheme 20 preferably includes the user’s personal preferences, and information regarding the user’s viewing history such as for example browsing history, filtering history, searching history, and device setting history. . . . [T]he user description scheme 20 may be augmented or any technique can be used to compare the information contained in the user description scheme 20 to the available information contained in the program description scheme 18 to make selections.”). A POSITA would have recognized that this selection would necessarily be based on associations between the clips because the search results conform to the user description scheme, which contains information about the first clip.

369. Additionally, the user interaction with a frame can also directly trigger the generation of new video programming, such as a highlight reel or event reel. *Id.* at [0304] (“The frames provide a visual summary. Depending on the browser, they may also provide clickable entry points to the video. The user may click and start playing back the video starting from that frame.”); [0107] (“The user starts interacting with the system with a pointer or voice commands to indicate a desire to view recorded sporting programs. . . . The user is interested in basketball games and indicates a desire to view games. A set of title frames is presented on the display

that captures an important moment of each game. The user selects the Chicago Bulls game and indicates a desire to view a 5 minute highlight of the game. The system automatically generates highlights.”); [0112] (“The user description scheme is generated by direct user input, and by using a software that watches the user to determine his/her usage pattern and usage history. User description scheme can be updated in a dynamic fashion by the user or automatically.”). This selection is based on associations between the clips because the second video clips are a highlight or event reel based on the selected first video clip.

370. To the extent that PO argues that Ferman does not explicitly disclose this element, it is my opinion that it would have been obvious to a POSITA to modify Ferman to implement that functionality in the manner and for the reasons given above in 1[f].

11. **Element 1[j]: “selecting, by one or more of said computing devices, a plurality of video clips including the second video clip, and determining a second temporal sequence for concatenating the selected plurality of second video clips, wherein said selecting, determining, and concatenating is performed in accordance with one or more first or second templates, to define second video programming;”**

371. In my opinion, Ferman discloses this element.

372. *Selecting, by one or more of said computing devices, a plurality of video clips including the second video clip, and determining a second temporal sequence for concatenating the selected plurality of second video clips, wherein*

said selecting, determining, and concatenating is performed in accordance with one or more first or second templates, to define second video programming: The same systems, data (in particular, templates), and algorithms that Ferman's system used to select a plurality of first video clips, determine a temporal sequence for them, and concatenate them are used in the second search. Therefore, all of the evidence and analysis I reviewed above in the context of elements 1[b], 1[d], 1[e], and 1[i] applies to this element as well.

373. In particular, in response to a user selection, Ferman's system can generate a highlight reel with a plurality of video clips in a temporal order, using Ferman's templates. *Id.* at [0107] ("The user selects the Chicago Bulls game and indicates a desire to view a 5 minute highlight of the game. The system automatically generates highlights. The highlights may be generated by audio or video analysis, or the program description scheme includes data indicating the frames that are presented for a 5 minute highlight.").

- 12. Element 1[k]: "wherein said one or more first or second templates comprises a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements; and"**

374. In my opinion, Ferman discloses this element.

375. *Wherein said one or more first or second templates comprises a plurality of variables corresponding to (a) information about the user, (b) viewing*

scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements: The templates used in follow-up searches in Ferman's system can be the same or different as the initial search (discussed in the context of claim element 1[e] above). And for the reasons I discussed above in the context of element 1[e], those templates contain "a plurality of variables corresponding to (a) information about the user, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) style requirements."

13. Element 1[l]: "presenting said second video programming on a display associated with one or more of said computing devices."

376. In my opinion, Ferman discloses this element.

377. *Presenting said second video programming on a display associated with one or more of said computing devices:* After performing a follow-up search and defining second video programming, the programming is displayed on the computing device's display, just like the first video programming. Additional detail is provided in the context of claim element 1[f] above.

B. Claim 2: "A method according to claim 1, wherein at least one of said video clips of said first or second video programming is a transitional clip."

378. In my opinion, Ferman discloses this claim.

379. *Wherein at least one of said video clips of said first or second video programming is a transitional clip:* As described above in the context of claim element 1[c] above, Ferman's system includes transitional clips in the first video programming. And as described in the context of element 1[j], the same types of transitional clips are also included in the second video programming. Additional details are discussed above in the context of claim elements 1[b], 1[d], and 1[e].

C. Claim 3: "A method according to claim 2, wherein at least one of said video clips of each of said first and second video programming is a transitional clip."

380. In my opinion, Ferman discloses this claim.

381. *Wherein at least one of said video clips of each of said first and second video programming is a transitional clip:* Transitional clips will be present in the dynamic video programming produced by Ferman's system, either for an initial search or a follow-up search. Therefore, for the same reasons detailed above in the context of claim 2, claim 3 is also disclosed by Ferman.

D. Claim 4: "A method according to claim 3, wherein said transitional clip is a hosting clip consistent with at least one template variable."

382. In my opinion, Ferman discloses or renders obvious this claim.

383. *Wherein said transitional clip is a hosting clip consistent with at least one template variable:* Video clips in Ferman's system can focus on a particular person or another host subject, like a lion:

Character/Object View

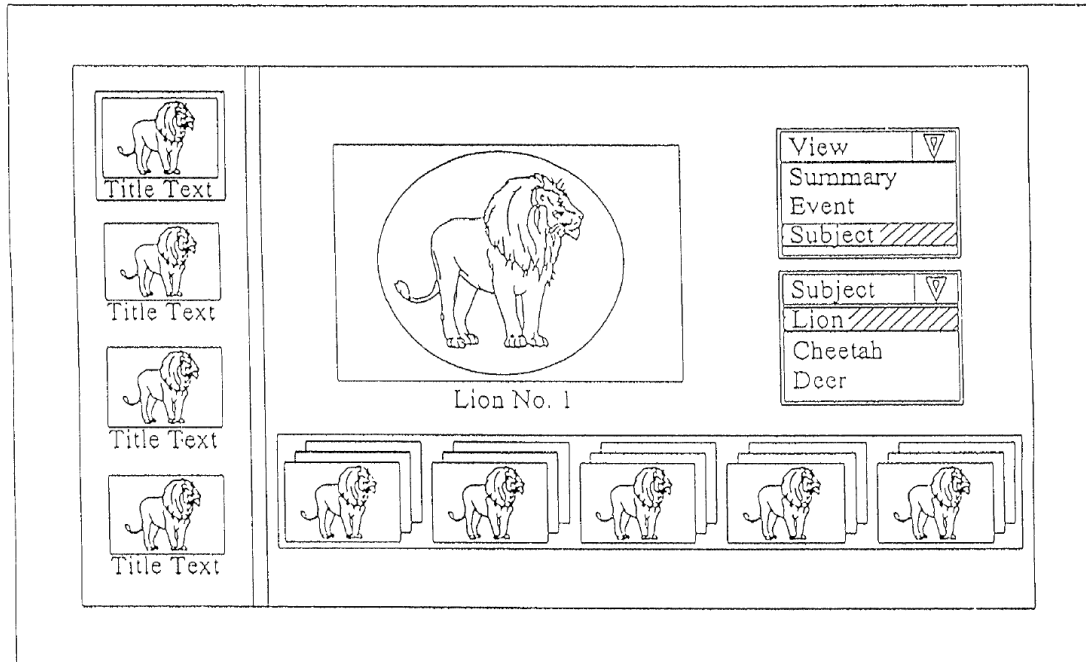


FIG. 12

Ferman at Fig. 12; [0124] (“Referring to FIG. 12, a character/object view may likewise be displayed, as desired.”); [0095] (“Program views contained in the program description scheme are a feature that supports a functionality such as close-up view. In the close-up view, a certain image object, e.g., a famous basketball player such as Michael Jordan, can be viewed up close by playing back a close-up sequence that is separate from the original program. An alternative view can be incorporated in a straightforward manner. Character profile on the other hand may contain spatio-temporal position and size of a rectangular region around the character of interest. This region can be enlarged by the presentation engine, or the presentation engine may darken outside the region to focus the user's attention to the characters spanning

a certain number of frames. Information within the program description scheme may contain data about the initial size or location of the region, movement of the region from one frame to another, and duration and terms of the number of frames featuring the region. The character profile also provides provision for including text annotation and audio annotation about the character as well as web page information, and any other suitable information. Such character profiles may include the audio annotation which is separate from and in addition to the associated audio track of the video.”). As with the “talking head” clips in Marcus156 (ground 1), in my opinion a POSITA would consider these types of clips to be “hosting clips,” as claimed. They feature a “host” that is the focal point of the clip.

384. To the extent that Ferman does not disclose this claim, it is my opinion that it renders it obvious, for the same reasons that Marcus156’s “talking head” clips would also render this claim obvious, discussed in ground 1, claim 4 above.

E. Claim 5: “A method according to claim 2, wherein at least one transitional clip is a hosting clip consistent with at least one template variable.”

385. In my opinion, Ferman discloses or renders obvious this claim.

386. *Wherein at least one transitional clip is a hosting clip consistent with at least one template variable:* Ferman discloses or renders obvious this claim for the same reasons given above in the context of claim 4.

F. Claim 6: “A method according to claim 2, wherein said first or second template variables corresponding to viewing scenarios

comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.”

387. In my opinion, Ferman discloses this claim.

388. *Wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming:* The template variables corresponding to viewing scenarios discussed in the context of element 1[e] above also include sequencing limitations for incorporating video clips at specific locations in the first or second video programming. *Id.* at [0095] (“Program views contained in the program description scheme are a feature that supports a functionality such as close-up view. In the close-up view, a certain image object, e.g., a famous basketball player such as Michael Jordan, can be viewed up close by playing back a close-up sequence that is separate from the original program. An alternative view can be incorporated in a straightforward manner. Character profile on the other hand may contain spatio-temporal position and size of a rectangular region around the character of interest. This region can be enlarged by the presentation engine, or the presentation engine may darken outside the region to focus the user's attention to the characters spanning a certain number of frames. Information within the program description scheme may contain data about the initial size or location of the region, movement of the region from one frame to another, and duration and terms of the number of frames featuring the region. The

character profile also provides provision for including text annotation and audio annotation about the character as well as web page information, and any other suitable information. Such character profiles may include the audio annotation which is separate from and in addition to the associated audio track of the video.”); [0252] (“The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video. A segment description scheme 450 may contain another segment description scheme 450 and/or shot description scheme to form a segment tree. Such a segment tree may be used to defined a temporal structure of a video program. Multiple segment trees may be created and thereby create multiple table of contents. For example, a video program may be segmented into story units, scenes, and shots, from which the segment description scheme 450 may contain such information as a table of contents.”); [0255] (“Referring again to FIG. 13, the semantic structure description scheme 404 is used to specify semantic features of a video program (or audio), e.g. semantic events. In a similar manner to the syntactic structure description scheme, the semantic structure description scheme 404 preferably includes three modules, namely an event description scheme 480, an object description scheme 482, and an event/objection relation graph description scheme 484. The event description scheme 480 may be used to form relationships between different events of the video normally consisting of multiple frames of the video. An event description scheme

480 may contain another event description scheme 480 to form a segment tree. Such an event segment tree may be used to define a semantic index table for a video program. Multiple event trees may be created and thereby creating multiple index tables. For example, a video program may include multiple events, such as a basketball dunk, a fast break, and a free throw, and the event description scheme may contain such information as an index table. The event description scheme may also contain references which link the event to the corresponding segments and/or regions specified in the syntactic structure description scheme.”); [0614] (compilation of sequencing-related variables). Because these “sequential requirements” place clips in temporal relationships to each other, a POSITA would have recognized that they are “sequencing limitations for incorporating video clips at specific locations” in the programming.

G. Claim 7: “A method according to claim 6, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.”

389. In my opinion, Ferman discloses this claim.

390. *Wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame:* Ferman’s templated sequencing limitations

meet each of these three categories. The limitations may be length requirements or requirements for presentation within a particular time frame. *Id.* at [0252] (“The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video. A segment description scheme 450 may contain another segment description scheme 450 and/or shot description scheme to form a segment tree. Such a segment tree may be used to defined a temporal structure of a video program. Multiple segment trees may be created and thereby create multiple table of contents. For example, a video program may be segmented into story units, scenes, and shots, from which the segment description scheme 450 may contain such information as a table of contents.”); [0255] (“Referring again to FIG. 13, the semantic structure description scheme 404 is used to specify semantic features of a video program (or audio), e.g. semantic events. In a similar manner to the syntactic structure description scheme, the semantic structure description scheme 404 preferably includes three modules, namely an event description scheme 480, an object description scheme 482, and an event/object relation graph description scheme 484. The event description scheme 480 may be used to form relationships between different events of the video normally consisting of multiple frames of the video. An event description scheme 480 may contain another event description scheme 480 to form a segment tree. Such an event segment tree may be used to define a semantic index table for a video

program. Multiple event trees may be created and thereby creating multiple index tables. For example, a video program may include multiple events, such as a basketball dunk, a fast break, and a free throw, and the event description scheme may contain such information as an index table. The event description scheme may also contain references which link the event to the corresponding segments and/or regions specified in the syntactic structure description scheme.”); [0281] (“The MaxSummaryDuration and MinSummaryDuration descriptors specify, respectively, the maximum and minimum temporal duration of an audiovisual highlight summary.”); [0614] (compilation of sequencing-related variables, including the temporal duration of clips).

H. Claim 9: “A method according to claim 6, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.”

391. In my opinion, Ferman discloses this claim.

392. *Wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history:*

Ferman’s templates include demographic information. *Id.* at [0097] (“The user description scheme may also include personal information about the particular user, such as demographic and geographic information, e.g. zip code and age.”); *id.* (“[S]elected information within the program description scheme may be encrypted since at least part of the information may be deemed to be private (e.g.,

demographics).”); [0110] (“Personalization can range from device settings, such as display contrast and volume control, to settings of television channels, radio stations, web stations, web sites, geographic information, and demographic information such as age, zip code, etc.”); [0118] (““The description scheme fulfills the user's desire to have applications that pay attention and are responsive to their viewing and usage habits, preferences, and personal demographics. The proposed user description scheme directly addresses this desire in its selection of fields and interrelationship to other description schemes. Because the description schemes are modular in nature, the user can port his user description scheme from one device to another in order to ‘personalize’ the device.”); [0221] (“The descriptor <Age> specifies the age of a user.”); [0224] (“The descriptor <Gender> specifies the gender of a user.”).

393. The templates further include psychographic information, namely the user’s “emotional state.” Ferman at [0274] (“Multiple user preference descriptions may correspond to, for example, different locations (e.g., at home, at the office, away from home, stationary versus traveling in a vehicle), different situations, different times (e.g., different days, different seasons), different emotional states of the user (e.g., happy mood versus tired or sad), and/or persistence (e.g., temporary usage versus permanent usage).”); [0275] (“In addition to maintaining multiple user preferences for a particular user based on the aforementioned conditions, the present inventors also came to the realization that the different locations, different situations,

different emotional states, different seasons, and/or different terminals (etc.), may likewise be used as the basis for distinguishing between the user preference descriptions.”). A POSITA would have considered an estimated or measured “emotional state” to be “psychographic information,” as claimed.

394. The templates also include a viewing history. Ferman at [0097] (“The user description scheme 20 preferably includes the user's personal preferences, and information regarding the user's viewing history such as for example browsing history, filtering history, searching history, and device setting history.”); [0212] (“The descriptor <BrowsingHistory> captures the history of a user’s browsing related activities.”); [0218] (“The descriptor <DeviceHistory> captures the history of a user’s device related activities.”); [0258] (“Moreover, the user preference description scheme should be readily usable by the system while likewise being suitable for modification based on the user’s historical usage patterns. It is possible to collectively track all users of a particular device to build a database for the historical viewing preferences of the users of the device, and thereafter process the data dynamically to determine which content the users would likely desire.”); [0259] (“The usage history description scheme 502, described in detail later, includes a description scheme of the user’s historical audio and/or video activity, such as browsing, device settings, viewing, and selection. The usage history description scheme 502 describes one or more of the following, depending on the particular

implementation, (a) browsing history, (b) filtering history, (c) searching history, and (d) device usage history.”); [0259] (“The data contained in the usage history description scheme 502 may be considered generally ‘unprocessed’, at least in comparison to the data contained in the usage preferences description scheme 500 because it generally contains the historical usage data of the audio and/or video content of the viewer.”); [0271] (“The usage history description 502 may contain additional descriptors therein (or description if desired) that describe the time and/or time duration of information contained therein. The time refers to the duration of consuming a particular audio and/or video program. The duration of time that a particular program has been viewed provides information that may be used to determine user preferences. For example, if a user only watches a show for 5 minutes then it may not be a suitable preference for inclusion the usage preference description 500.”).

I. Claim 10: “A method according to claim 6, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.”

395. In my opinion, Ferman discloses this claim.

396. *Wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates:* Templates are used to select video clips based on associations with other video clips, including by using a tag or attribute of one clip

as a necessary requirement for another clip. Ferman at [0105] (“A searching, filtering, and browsing (SFB) module 52 implements the description scheme technique by parsing and extracting information contained within the description scheme. The SFB module 52 may perform filtering, searching, and browsing of the programs 38, on the basis of the information contained in the description schemes. . . . In this manner, desirable content may be provided to the user through a display 80.”); [0270] (“It is preferable that media content descriptions are consistent with descriptions of preferences of users consuming the media.”).

J. Claim 11: “A method according to claim 6, wherein said stored associations between said video clips comprise a plurality of media tags.”

397. In my opinion, Ferman discloses this claim.

398. *Wherein said stored associations between said video clips comprise a plurality of media tags:* Ferman’s system uses media tags associated with clips to categorize clips and to associate them with other clips. *Id.* at [0094] (“The program profiles define distinctive characteristics of the content of the program, such as actors, stars, rating, director, release date, time stamps, keyword identification, trigger profile, still profile, event profile, character profile, object profile, color profile, texture profile, shape profile, motion profile, and categories. The program profiles are especially suitable to facilitate filtering and searching of the audio and video information.”); [0315] (“Each of the selected programs include a rich set of

associated data which may then be applied against user criteria at block 618 to further refine the data by extracting desirable sub-portions of each program.”); [0407] (“Referring to FIG. 51, the input to a filter agent 600 may include program descriptions 602 relating to one or more audio, video, or audiovisual (collectively referred to as audiovisual without limitation) materials.”); cl. 1 (“(b) receiving program attribute information corresponding to said at least one of said audio and video, where said program attribute information corresponds with said user preferences, where at least one of said user attribute information and said program attribute information includes usage history”). In my opinion, these are “media tags” as described in the appendix to the ’597 provisional and according to the plain and ordinary meaning. Ex. 1003 at 61 (“Described here are tagging considerations Media clips will normally bear taxonomic data allowing the clip to be retrieved and sorted by clip content.”).

K. Claim 12: “A method according to claim 6, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.”

399. In my opinion, Ferman discloses this claim.

400. *Wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history:* Ferman’s media

tags include context and subject matter data. Ferman at [0278] (“Another aspect of the present invention permits the user preferences (and history) to be based upon the media type, media source, or content (e.g., music versus video, radio versus television broadcast, and/or sports video versus home video.”); [0282] (“A Segment Theme descriptor(s) may describe the preferred theme, or point of view, of a segment, e.g., a video or audio clip, annotated with its theme or emphasis point. For example, the theme may specify characteristics of the content of the theme. Such characterization may include a goal from your favorite team, 3-point shots from your favorite player, etc. Specifying these descriptor(s) and also ranking them enables a client application or a server to provide to the user segments according to preferred themes (and/or their ranking) matching to the their labels or descriptors at the segment level, or provide users with pre-assembled highlights composed of segments with labels matching the SegmentTheme preference.”); [0303] (“(D) The SegmentTheme descriptor(s) may be included, which describes the theme or point of view of a segment, e.g., a video or audio clip annotated with its theme or emphasis point.”); [0601] (“The Creation history 744, may provide a history of the user’s choices with respect to creation descriptions of content, such as title 746, location 748, creator 750, date 752, and actor/director.”).

401. The media tags also include source information. Ferman at [0283] (“[T]he present inventors came to the further realization that the identification of the

source of the material, in contrast to merely its content, provides beneficial information for the processing and presentation of the audiovisual materials. For example, the source of the content may be from terrestrial sources, digital video disc, cable television, analog broadcast television, digital broadcast television, analog radio broadcasts, and digital radio broadcasts. The inclusion of this information permits the user to select among these different sources and increase effectiveness by narrowing down the choices to those sources that are available to the user, such as terrestrial broadcast which is more widely available than satellite broadcast. For example, user may describe user's preference for 'Star Trek' episodes that are available from terrestrial broadcast channels only."); [0284] ("This source distinction and identification may be performed by including a source preferences description scheme under the filtering and search preferences description scheme (or other description scheme). Accordingly, the search and preferences description scheme may include from zero or one (or more if desired) source preferences description scheme. The source preferences description scheme may be derived from the Media Format description scheme or Publication Description Scheme specified in ISO/IEC JTC1/SC29/WG11 N3247, MPEG-7 Multimedia Description Schemes, Experimentation Model (v2.0) Noordwijkerhout, March 2000."); [0305] ("The source preference description describes the preferred source of multimedia information, such as the broadcast or storage medium type (e.g., terrestrial, satellite,

DVD), broadcast channel identifier, etc.”); [0602] (“The Source history 754, may provide a history of the user’s choices with respect to source (such as its publisher or distributor) and format (such as coding type) descriptions of the content.”).

402. Moreover, the media tags include data about region and language. Ferman at [0306] (“The filtering and search preferences description includes at least one of the descriptors of preferred program title, genre, language, actor, creator of the program. An example description where user’s preference is for news programs in English is given below. Such description may be included in user’s smart card when he travels to Japan, for example. Note that this particular preference description is identified as being specific to Japan and differentiated by choosing an appropriate user name.”); [0328] (“Common names, such as ‘Country’ used at multiple locations, may be distinguished by including all or part of the ancestry path. For example, the following two preference tests have the same “leaf” name, but it may be desirable to have different tests for each. This may be done by specifying more of the ancestry in the Name field (column 1) of the mapping table: ‘/FilteringAndSearchPreferences/CreationPreferences/CreationLocation/Country’, and ‘/FilteringAndSearchPreferences/ClassificationPreferences/Country’. To distinguish between the two, the following names may be used: ‘/CreationLocation/Country’ and ‘/ClassificationPreferences/Country’. In addition the preference tests may be associated with multiple entries in the Mapping Table.

This permits a single test to be performed on more than one location in the Program Description.”); [0329] (“The Location field may include various wildcards to expand or restrict the target paths to be evaluated in the Program Description. For example, a ‘*’ wildcard implies that there may be multiple instances of the given location under one parent, e.g., /Creation/*Creator implies that there may be multiple Creators under the Creation parent. A “#xxx” wildcard restricts the target path to the xxx instance of the given location under its parent, e.g., /Creation/#002Creator restricts the target path to the second instance of Creator under Creation. A double forward slash “//” indicates a node of the target path which may be used as a base path for groups of tests which must be constrained to evaluate from the same common location. In particular, this is useful for Constrained-AND operations, described later. The preference paths may be used to build target locations in the program. These preference paths may also allow preference paths to be interpreted as locations. Composite values may be defined for these preference path locations.”); [0584] (“The system also allows usage history information to be captured at different levels of detail. For example, a usage history description can contain a detailed list of all the actions the user has performed over a period of time; or basic statistical information according to certain categorizations of content, such as language, country of origin, ratings, actors, etc.; or both. This flexibility enables even systems with limited resources (e.g. in terms of available storage space) to generate

standardized, exchangeable usage histories.”); [0600] (“The Classification history 742, may provide a history of the user’s choices with respect to classification descriptions of content; such as genre, language, and country of origin.”).

L. Claim 13: “A method according to claim 1, wherein said first and second video programming are each presented as a stream of information to the user.”

403. In my opinion, Ferman discloses this claim.

404. *Wherein said first and second video programming are each presented as a stream of information to the user:* The finalized first and second video programming is assembled and presented to the user as a unified video stream. The user sees a sequence of videos where the clips and transitions proceed sequentially, one after the other and without further user input. *Id.* at [0607] (“PlayStream Play content from input stream”). The system generates a sequence of video programming of a given length, composed of a concatenation of highlight or event clips. *Id.* at [0091] (“In light of the vast amount of audio, image, and video information, the present inventors came to the realization that a system that records and presents to the user audio and video information based upon the user’s prior viewing and listening habits, preferences, and personal characteristics, generally referred to as user information, is desirable. In addition, the system may present such information based on the capabilities of the system device.”); [0107] (“The user selects the Chicago Bulls game and indicates a desire to view a 5 minute highlight

of the game. The system automatically generates highlights.”); *id.* (“The user plugs in the tape into his VCR, that is connected to the system, and invokes the summarization functionality of the system to scan the tape and prepare a summary. The user can then view the summary the next morning to quickly discover the baby’s looks, and playback segments between the keyframes of the summary to catch a glimpse of the crying baby.”); [0267] (15-minute highlight stream); Fig. 10.

M. Claim 14: “A method according to claim 1, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.”

405. In my opinion, Ferman discloses this claim for the same reasons given above in the context of claim 6, which uses similar language.

N. Claim 15: “A method according to claim 14, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.”

406. In my opinion, Ferman discloses this claim for the same reasons given above in the context of claim 7, which uses similar language.

O. Claim 17: “A method according to claim 1, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.”

407. In my opinion, Ferman discloses this claim for the same reasons given above in the context of claim 9, which uses similar language.

- P. Claim 18: “A method according to claim 1, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.”**

408. In my opinion, Ferman discloses this claim for the same reasons given above in the context of claim 10, which uses similar language.

- Q. Claim 19: “A method according to claim 1, wherein said stored associations between said video clips comprise a plurality of media tags.”**

409. In my opinion, Ferman discloses this claim for the same reasons given above in the context of claim 11, which uses similar language.

- R. Claim 20: “A method according to claim 1, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.”**

410. In my opinion, Ferman discloses this claim for the same reasons given above in the context of claim 12, which uses similar language.

XIII. Ground 9: Ferman and Byers Render Claims 8 and 16 Obvious

411. In my opinion, Ferman in combination with Byers renders claims 8 and 16 obvious.

A. Specific Proposed Modification to Ferman

412. Ferman’s system would be modified according to the teachings of Byers to include advertising clips in the video programming, including by dynamically editing other clips. Part of this modification would entail adding

advertising material to the program description scheme and tracking additional advertising-related parameters in both the program description scheme and the user description scheme.

B. Motivation to Combine

413. *Commercial benefit to advertising*: Because insertion of advertising into video programming would enable a new revenue stream for the video creator, a POSITA would have had an economic motivation to do it. *E.g.*, Byers at 2:51-52 (“Such placement can lead to a boon in sales and publicity for the products so advertised.”); Ferman at [0088-92].

414. *Commercial benefit to user-targeted advertising*: The user-targeted recommendation techniques disclosed in Ferman allow improved advertising that is customized to the demographics, interests, or other facets of the user, thus increasing the pertinence and, ultimately, the click-through or conversion rate of the targeted advertising. Byers at 2:57-67 (“One problem with product placement is that the placement is not tailored for an individual user. . . . Utilizing the present invention alleviates this problem by facilitating the placement of replacement images that are related to a user profile . . .”).

415. *Same field of art*: Like Ferman and the ’185 patent, Byers is directed to creation and presentation of dynamic video programming. Byers at 1:40-43 (“Therefore, a need exists for a method and apparatus for dynamically altering

digital video images to allow customization of a digital video stream based upon characteristics associated with a user.”); Ferman at [0090] (“[T]he presentation of the information should be in a manner that is time efficient, especially in light of the limited time frequently available for the presentation of such information. In addition, the user should be able to access only that portion of all of the available information that the user is interested in, while skipping the remainder of the information.”); [0091] (“In light of the vast amount of audio, image, and video information, the present inventors came to the realization that a system that records and presents to the user audio and video information based upon the user’s prior viewing and listening habits, preferences, and personal characteristics, generally referred to as user information, is desirable.”); [0102].

C. Reasonable Expectation of Success

416. *Software modification would be predictable*: This change would be accomplished in software and would entail predictable changes to Ferman’s system. Because the template infrastructure already exists to add advertising-related information, the change would only require minor modifications, such as adding a few extra fields, to Ferman’s data structures.

417. *Byers provides explicit instructions on how to modify clips to insert advertising*: Byers discloses underlying mechanisms to insert the advertising content. *E.g.*, Byers at 3:67-4:5 (“[S]ervice node 103 receives the digital video

stream, determines a replacement image to insert in the digital video stream, performs target geometry transforming and lighting model, digitally alters the digital video stream by inserting the replacement image, and transmits the altered video image to broadband access network 105.”); 5:60-66 (“DSP 209 preferably accomplishes the replacement of the original element by calculating the geometry of the replacement image, determining the lighting present in the original image, determining the transparency of the replacement area in the original image, determining the object order and motion, and overwriting selected portions of the frame buffer that include the original video image.”). The hardware used in this process is similar to Ferman’s disclosed hardware. Byers at 5:45-67; Ferman at Fig. 2. Ferman’s disclosed hardware could perform Byers’ video editing algorithms, including because it is capable of performing similar editing algorithms described in Ferman, such as highlight generation.

D. Claims 8 and 16: Ferman and Byers Render Obvious “One or More Video Clips Comprise Advertising”

418. Byers discloses an “advertiser database” that contains advertising-related information. Byers at 4:19-25 (“Advertiser database 113 includes advertising information that relates to advertisers. Such information is preferably organized to allow the retrieval of digital images from image database 117 that can be inserted into the digital video stream. For example, advertiser database can include information relating to thresholds that advertisers are willing to utilize.”).

419. The information in the database is used to dynamically insert ads into a video stream. *Id.* at 2:48-51; 3:1-4 (“A replacement image for a product, such as a product that an actor is using in a movie or television show, can be altered for each desired user based upon factors associated with the intended recipient of the video stream.”).

420. Moreover, Ferman’s system explicitly tracks advertising clips. For example, Ferman mentions the capability to intentionally remove advertising. *Id.* at [0103] (“The knowledge based system 90 may invoke a commercial filter 92 to remove commercials and a slow motion detector 54 to assist in creating the video summary.”). A POSITA would have understood that in order to remove commercials, the system must be able to identify commercials in clips, and the clips in the database must include commercials. If the commercial-removing functionality is not activated, then commercials will remain in the video programming and be displayed to the user.

XIV. No Secondary Considerations of Non-Obviousness

421. I understand that so-called “secondary considerations” are legally relevant to an obviousness analysis. But I am unaware of any evidence of such considerations that would suggest the non-obviousness of any of the claims of the ’185 patent addressed in the accompanying *inter partes* review petition. If such evidence is presented to me, I will consider it; however, I doubt that any commercial

success of the claimed invention or other objective evidence would be attributable to the merits of the claimed invention given my analysis of the lack of novelty of the '185 patent.

XV. Conclusion

422. In signing this declaration, I recognize that the declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to cross-examination in the case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

* * *

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on the information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Respectfully submitted,

Date: August 23, 2024



Kevin C. Almeroth, Ph.D.