

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TIKTOK INC.,
Petitioner,

v.

NTECH PROPERTIES, INC.,
Patent Owner.

Case No. IPR2024-01338

U.S. Patent No. 8,875,185

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 8,875,185**

TABLE OF CONTENTS

I.	Introduction.....	1
II.	Mandatory Notices.....	2
	A. Real Party-in-Interest	2
	B. Related Matters.....	2
	C. Counsel and Service Information.....	2
	D. Payment of Fees	3
	E. Requirements For Inter Partes Review	3
III.	Grounds.....	3
IV.	'185 Patent Overview	4
V.	Level of Ordinary Skill in the Art	5
VI.	Claim Construction.....	5
VII.	Prior Art Summary	5
	A. Marcus156 (Ex. 1009).....	5
	B. Rosen (Ex. 1011).....	6
	C. Sull (Ex. 1010)	7
	D. Hosken (Ex. 1013).....	8
	E. Byers (Ex. 1014).....	8
	F. Ferman (Ex. 1012).....	9
VIII.	Ground 1: Marcus156 Renders Claims 1-20 Obvious	10

A.	Claim 1	10
1.	Element 1[pre].....	10
2.	Element 1[a]	10
3.	Element 1[b].....	11
4.	Element 1[c]	12
5.	Element 1[d].....	14
6.	Element 1[e]	16
7.	Element 1[f]	21
8.	Element 1[g].....	22
9.	Element 1[h].....	24
10.	Element 1[i].....	25
11.	Element 1[j].....	26
12.	Element 1[k].....	27
13.	Element 1[l].....	27
B.	Claim 2	27
C.	Claim 3	28
D.	Claim 4	28
E.	Claim 5	30
F.	Claim 6	30
G.	Claim 7	30
H.	Claim 8	30

I.	Claim 9	31
J.	Claim 10	31
K.	Claim 11	31
L.	Claim 12	32
M.	Claim 13	32
N.	Claims 14-20	33
IX.	Ground 2: Marcus156 and Rosen Render Claims 1-20 Obvious	33
A.	Proposed Modification	33
B.	Motivation to Combine	34
C.	Reasonable Expectation of Success	35
D.	All Claims and Claim Elements Other Than 1[g], 1[h], and 1[i]	36
E.	Element 1[g]	37
F.	Element 1[h]	38
G.	Element 1[i]	38
X.	Ground 3: Marcus156 and Sull Render Claims 1-20 Obvious.....	38
A.	Proposed Modification	38
B.	Motivation to Combine	39
C.	Reasonable Expectation of Success	39
D.	All Claims and Claim Elements Other Than 1[g], 1[h], and 1[i]	40
E.	Element 1[g]	40
F.	Element 1[h]	41

G.	Element 1[i]	41
XI.	Ground 4: Marcus156, Rosen, and Hosken Render Claims 1-20 Obvious...	42
A.	Proposed Modification	42
B.	Motivation to Combine	42
C.	Reasonable Expectation of Success	43
D.	All Claims and Claim Elements Other Than 1[i].....	44
E.	Element 1[i]	44
XII.	Ground 5: Marcus156, Sull, and Hosken Render Claims 1-20 Obvious	45
XIII.	Ground 6: Marcus156, Rosen, Hosken, and Byers Render Claims 8 and 16 Obvious.....	45
A.	Proposed Modification	45
B.	Motivation to Combine	45
C.	Reasonable Expectation of Success	46
D.	Byers Discloses and Renders Obvious “One or More Video Clips Comprise Advertising”	47
XIV.	Ground 7: Marcus156, Sull, Hosken, and Byers Render Claims 8 and 16 Obvious.....	47
XV.	Ground 8: Ferman Renders Claims 1-7, 9-15, and 17-20 Obvious.....	48
A.	Claim 1	48
1.	Element 1[pre].....	48
2.	Element 1[a]	48
3.	Element 1[b].....	50
4.	Element 1[c]	52

5.	Element 1[d].....	53
6.	Element 1[e].....	55
7.	Element 1[f]	57
8.	Element 1[g].....	60
9.	Element 1[h].....	60
10.	Element 1[i].....	61
11.	Element 1[j].....	62
12.	Element 1[k].....	62
13.	Element 1[l].....	62
B.	Claim 2	63
C.	Claim 3	63
D.	Claim 4	63
E.	Claim 5	64
F.	Claim 6	65
G.	Claim 7	65
H.	Claim 9	65
I.	Claim 10	66
J.	Claim 11	67
K.	Claim 12	67
L.	Claim 13	68
M.	Claims 14-15 and 17-20	69

XVI. Ground 9: Ferman and Byers Render Claims 8 and 16 Obvious	69
A. Proposed Modification	69
B. Motivation to Combine	69
C. Reasonable Expectation of Success	70
D. Claims 8 and 16: Ferman and Byers Render Obvious “One or More Video Clips Comprise Advertising”	71
XVII. Instituting This IPR Would Be Equitable.....	71
A. 35 U.S.C. § 325(d) Supports Institution.....	71
B. <i>Fintiv</i> Supports Institution.....	74
1. Issue Overlap.....	74
2. Likelihood of Stay.....	75
3. Proximity of Trial to FWD	75
4. Investment in Parallel Proceeding	75
5. Identity of Parties.....	76
6. Other Circumstances.....	76
XVIII. Conclusion.....	76

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

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

LIST OF CHALLENGED CLAIMS

Claim No.	Claim Language
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 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;

Claim No.	Claim Language
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;
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.

Claim No.	Claim Language
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.
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

Claim No.	Claim Language
	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.
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.

I. Introduction

The disclosure of the challenged '185 patent is limited and superficial. The '185 patent has three figures and under four columns of specification, of which one column is a description of the prior art. The '185 patent incorporates by reference the contents of the Marcus156 patent, Ex. 1006, which names the same inventor, Dwight Marcus, as the '185 patent. Marcus156 had already issued more than three years before the first provisional application that led to the '185 patent was filed. Despite incorporating it by reference, the '185 patent does not claim priority to Marcus156 or to any patent applications in the same family as Marcus156, rendering Marcus156 prior art to the '185 patent under 35 U.S.C. § 102(b).

The issued claims of the '185 patent, however, are based primarily on material contained in the prior art Marcus156 patent, rather than anything in the specification of the '185 patent itself. For example, claim 1 of the '185 patent requires assembling video streams comprising clips and transitions, using templates containing “(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.” '185 patent at cl. 1. These requirements are discussed at length in Marcus156 using the exact same terminology claimed in the '185 patent. On the other hand, the specification of the '185 patent itself is mostly silent on these topics. Even where the '185 patent specification goes beyond Marcus156, for example by providing more detail on a

particular user interface that displays a video stream and a frame from that stream at the same time, this additional concept would have been obvious in light of Marcus156, alone or in view of other prior art. Thus, the '185 patent is an improper attempt by the named inventor and assignee to extend the life of Marcus156 by claiming it again—or at least a minor and obvious modification to it.

Accordingly, Petitioner respectfully requests that the Board institute this IPR and find all claims of the '185 patent invalid.

II. Mandatory Notices

A. Real Party-in-Interest

The real parties-in-interest for Petitioner are TikTok Inc., Bytedance Ltd., Bytedance Pte. Ltd., and TikTok Pte. Ltd.

B. Related Matters

The '185 patent has only been asserted in *NTECH Properties, Inc. v. ByteDance Ltd. et al.*, Case No. 2:24-cv-00130 (E.D. Tex., filed Feb. 23, 2024).

The '185 patent is in the same family as U.S. Patent No. 9,317,597. Petitioner is concurrently filing an IPR petition against that patent.

C. Counsel and Service Information

Electronic service may be made on the email addresses of counsel identified below and in the accompanying Power of Attorney.

Lead Counsel	Backup Counsel
James M. Glass (Reg. No. 46729) jimglass@quinnemanuel.com QUINN EMANUEL URQUHART & SULLIVAN, LLP 51 Madison Avenue, 22 nd Floor New York, NY 10010 Tel: (212) 849-7000	Ognjen Zivojnovic (Reg. No. 69,516) ogizivojnovic@quinnemanuel.com QUINN EMANUEL URQUHART & SULLIVAN, LLP 50 California Street, 22nd Floor San Francisco, California 9411 Tel: (415) 875-6600 Gavin Snyder (<i>pro hac vice</i> to be requested upon grant authorization) gavinsnyder@quinnemanuel.com QUINN EMANUEL URQUHART & SULLIVAN, LLP 1109 First Avenue, Suite 210 Seattle, Washington 98101 Tel: (206) 905-7000

D. Payment of Fees

The Office is authorized to charge the fee required for this Petition (and any additional fees) to Deposit Account No. 50-5708.

E. Requirements For Inter Partes Review

Petitioner certifies that the '185 patent is available for *inter partes* review, that Petitioner is not barred or estopped, and that 35 U.S.C. §§315(a)-(b) is inapplicable.

III. Grounds

Petitioner presents the following grounds of challenge.

Ground	Basis	Prior Art	Claims
1	§103	Marcus156	1-20

Ground	Basis	Prior Art	Claims
2	§103	Marcus156 and Rosen	1-20
3	§103	Marcus156 and Sull	1-20
4	§103	Marcus156, Rosen, and Hosken	1-20
5	§103	Marcus156, Sull, and Hosken	1-20
6	§103	Marcus156, Rosen, Hosken, and Byers	8, 16
7	§103	Marcus156, Sull, Hosken, and Byers	8, 16
8	§103	Ferman	1-7, 9-15, 17-20
9	§103	Ferman and Byers	8, 16

IV. '185 Patent Overview

The '185 patent claims priority to the '597 provisional, filed June 24, 2003. The specification of the '185 patent is exceptionally short. It includes three figures and less than four columns of text. The “Background of the Invention” contrasts the alleged invention with three prior art video delivery systems: television, video on demand, and the world wide web. '185 patent at 1:26-2:2. With television, “a user has very little control over what information is presented at a specific time.” *Id.* at 1:47-48. Video on demand used mundane, list-based user interfaces. *Id.* at 1:58-61. The world wide web offered more user control but was “inefficient and non-entertaining.” *Id.* at 2:1-2.

The '185 patent purports to address these deficiencies by automatically customizing video programming based on metadata, which the '185 patent calls “templates” and “headers.” *Id.* at 2:6-18; 2:67-3:8.

V. Level of Ordinary Skill in the Art

A person of ordinary skill in the art at the time of the '185 patent ("POSITA") 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. More education can supplement for less practical experience, and vice versa.

Petitioner's expert exceeded this level by the priority date.

VI. Claim Construction

The claim terms should be given their plain and ordinary meaning under *Phillips*, and "construed only to the extent necessary to resolve the controversy." *Wellman, Inc. v. Eastman Chem. Co.*, 642 F.3d 1355, 1361 (Fed. Cir. 2011). Here, no express constructions are required to find the '185 patent claims unpatentable.

VII. Prior Art Summary

A. Marcus156 (Ex. 1009)

Marcus156 issued on February 29, 2000. It is 35 U.S.C §102(b) prior art to the '185 patent. It names the same inventor, Dwight Marcus, and describes the same technology as the '185 patent. There is no exception to pre-AIA 35 U.S.C. §102(b) based on commonality of inventorship between the prior art and the challenged patent.

Just like 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. Marcus156, Abstract.

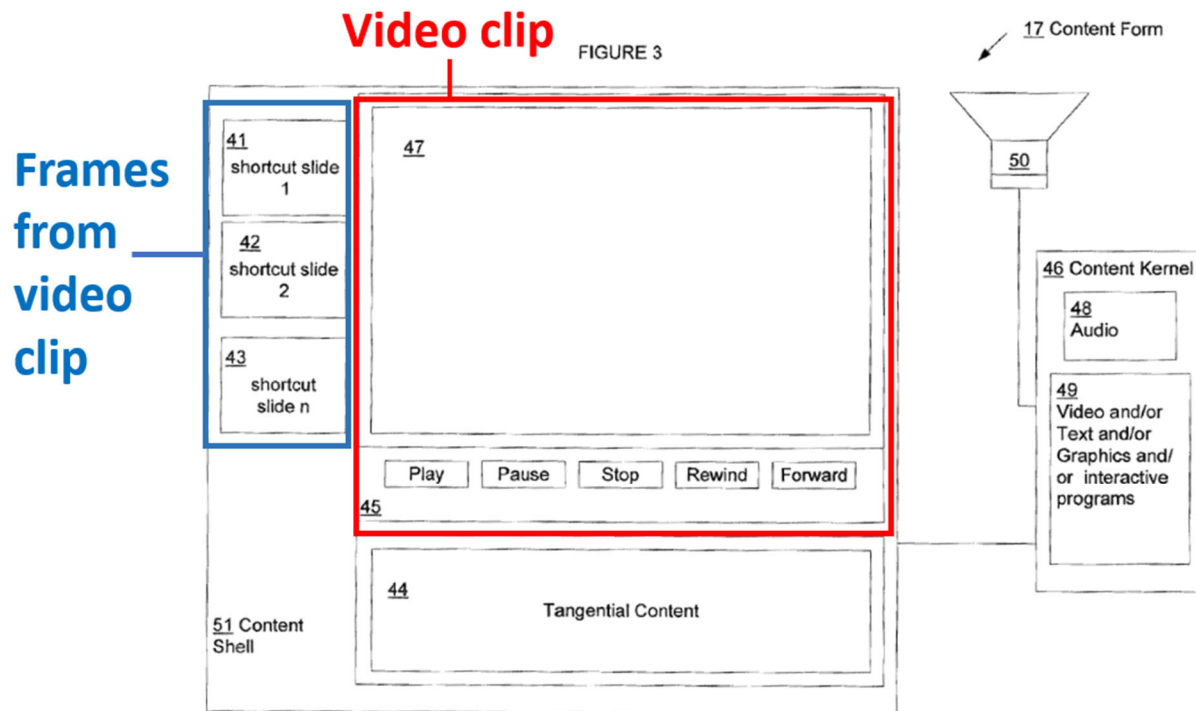
During prosecution, the examiner raised Marcus156 as a secondary reference in obviousness rejections (Ex. 1005 at 219-23) but never as a base reference. The applicant acquiesced to these rejections, canceled all pending claims, and added the set of new claims that became issued claims 1-20. *Id.* at 232-42. The '185 patent issued without any further office action rejecting these new claims. *Id.* at 253-61, 294-303, 328. In allowing the new claims, the examiner appears to have overlooked several aspects of Marcus156. This petition highlights and explains those overlooked aspects.

Marcus156 was also cited by the examiner during the prosecution of other patents in the same family as the '185 patent, and each time applicant acquiesced to these rejections by rewriting the claims. In the instances where the Examiner considered Marcus156 as a base reference (rather than just a secondary reference), applicant was further forced to add additional, substantive limitations not found in the '185 patent claims.

B. Rosen (Ex. 1011)

Rosen was filed on November 13, 2001 and published on October 14, 2004. Rosen is 35 U.S.C. §102(e) prior art. Rosen discloses displaying **video** alongside **slides** and “tangential content.” Rosen at Abstract; [0001]. **A video plays** “in a

main window 47”:



Id. [0023]; Fig. 3 (annotated). **Shortcut slides 41-43** are displayed next to **the playing video**. *Id.* at Fig. 3, [0025]. These contain **frames of the playing video**. *Id.* [0027]. The user selects **a shortcut slide** to initiate an action. *Id.* [0023], [0027].

C. Sull (Ex. 1010)

Sull is a patent publication published on June 6, 2002. Sull is 35 U.S.C. §102(b) prior art. Sull discloses a system for “tagging, indexing, searching, retrieving, manipulating, and editing” videos. Sull at Abstract. **“Multimedia bookmarks”** correspond to a specific location in **a playing video** and are displayed as **a frame**:

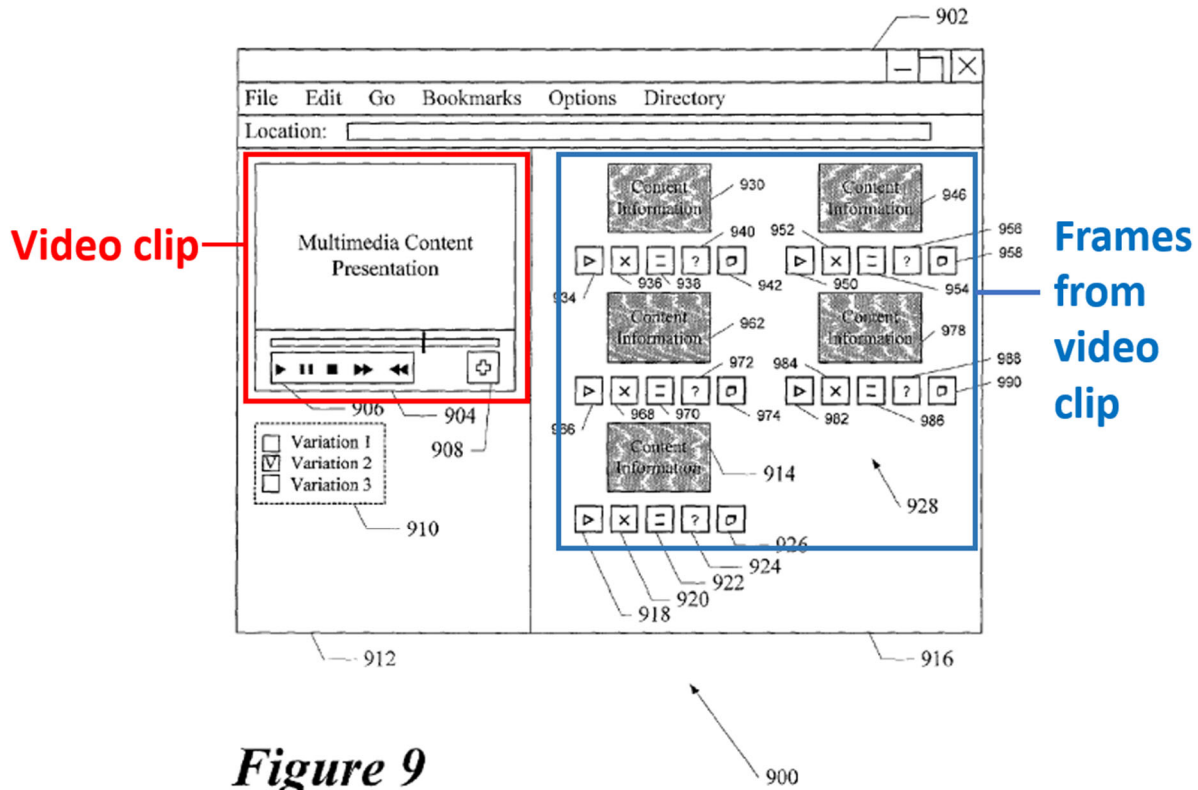


Figure 9

Id. at Fig. 9; [0205]; [0220]. A user selects a bookmark frame to search for compatible videos. *Id.* [0225]; [0221]; [0024]; [0060].

D. Hosken (Ex. 1013)

Hosken is a PCT publication published on January 25, 2001. Hosken is 35 U.S.C. §102(b) prior art. Hosken discloses content recommendation algorithms. Hosken at 1:10-14. Hosken describes how items from a first search result set can be selected by the user and used to perform successive searches. *Id.* at 2:9-11; 9:4-8.

E. Byers (Ex. 1014)

Byers is a U.S. Patent that was filed on April 26, 2000 and issued February 19, 2008. Byers is prior art under 35 U.S.C. §102(e). Byers describes dynamically

inserting advertising into a video. Byers at 2:48-52.

F. Ferman (Ex. 1012)

Ferman is a patent publication published on May 16, 2002. Ferman is 35 U.S.C §102(b) prior art. Ferman uses metadata templates to match user preferences with a database of video clips to create interactive video programming. Ferman [0001]; [0089]; [0091]; [0093]. Ferman's user interface displays **a video clip** alongside **frames from the same**:

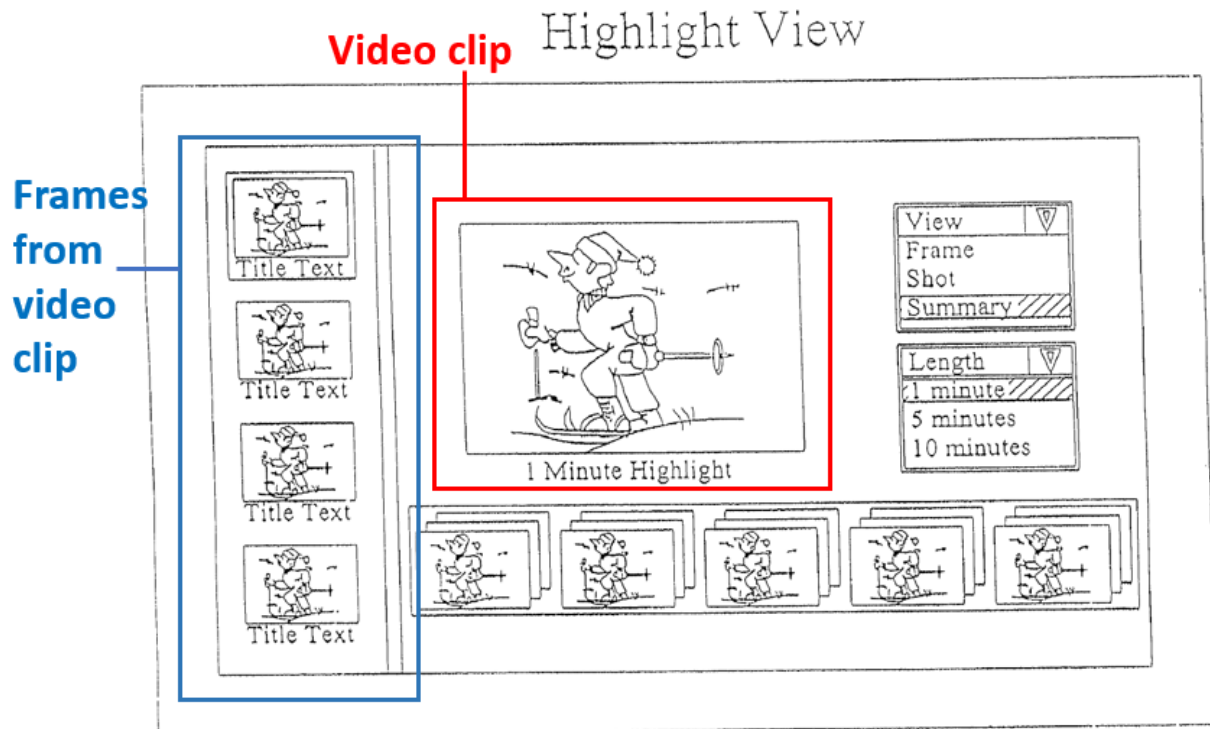


FIG. 10

Id. at Fig. 10 (annotated).

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

A. Claim 1

1. Element 1[pre]

Video programming method: Marcus156's system "deliver[s] information via customized video programming." Marcus156, Abstract; 2:10-11; 10:23-27; 14:46-58; Almeroth ¶¶149-150.

2. Element 1[a]

Databases of video clips: Marcus156 discloses a database 100 of video clips. Marcus156, 1:63-65 ("A database is provided which contains selected information about each of a large number of media elements."). The database includes metadata about clips. *Id.* at 2:67-3:2; 2:35-37; 5:40-43; Fig. 2; cl. 1; Almeroth ¶¶152-53.

Marcus156 also discloses a library 70 storing the underlying video data for those video clips. Marcus156, 2:56-58; 2:61-64. To the extent the claimed "database of video clips" must include raw video data, library 70, alone or when combined with database 100, constitutes the claimed "one or more databases of video clips." To the extent PO argues that library 70 is not a "database," it would have been obvious to integrate the raw video data in library 70 into database 100 (thereby eliminating need to separately access library 70 after selecting video clips based on database 100) or to implement library 70 as a database (thereby speeding up the lookup and retrieval of video data from library 70). Almeroth ¶¶154-56.

Searching in response to a user request: The databases are searched in response to a user request (also called a “query”). Marcus156, 1:67-2:2 (“selecting some of those media elements in response to a request for media programming”); 2:25-28 (“from a database containing information concerning the audiovisual media elements, certain audiovisual media elements are selected”); 4:60-61; 9:9-11; 13:51-52; 14:33-34. The user may be using either Marcus156’s “List Editor/Previewer” (7:10-15) or its “interactive system” (14:35-67). Almeroth ¶¶157-58.

Computing devices: The search is performed by computer 20. “Computer 20 is also suitably associated with a database 100.” Marcus156, 2:66-67. An editing program including an automatic expert system, running on computer 20, can perform the search. *Id.* at 4:52-5:2; 9:11-12; 8:13-29; cl. 1; Almeroth ¶159.

Frames: Marcus156’s clips are, like any video, comprised of a plurality of frames. Marcus156, 9:48-51 (MPEG-compressed video contains frames); 2:61-62; 3:39-41 (transition time measured in amount of frames); 7:54-57 (transitions computed on frame-by-frame basis); Almeroth ¶160.

3. Element 1[b]

Selecting video clips: Computer 20 selects clips that satisfy the request. Marcus156, 2:6-7 (“selects and orders the media elements according to the data in the request”); 1:67-2:2; 2:25-28; 4:60-61; 8:29-32; 9:9-11; 13:51-52; 14:33-34; cl. 1; Almeroth ¶162.

Clip associations: Metadata, including tags and templates, associates each clip in the database with other clips. Any selected clip (“first video clip”) is associated with other video clips (“media content comprising one or more second video clips”) because they share the same or similar tags, or are identified as similar according to templates. Marcus156, 3:25-30 (“The system by which the content information is defined is preferably organized to provide for relationships among type of information.”); 3:13-14; Almeroth ¶163.

Hierarchical categorizations also cause the clips to be associated. Marcus156, 2:67-3:2; 3:31-33; 3:60-65. Finally, associations between clips will cause them to be concatenated into the same output video stream. Marcus156, 4:9-14; 8:3-5; Almeroth ¶¶164-65.

4. Element 1[c]

Selecting transitional clips: Once initial clips are selected, Marcus156’s system selects accompanying transitions. Marcus156, 2:29-31 (“Transitions between the media elements are automatically selected.”); 14:13-17; 14:43-45. The transitions can be standalone clips (a “bridge” or “floater”) inserted between two other clips (for example, inserting a percussion clip between the two clips lacking percussion). *Id.* at 8:51-56; 4:28-29; 4:34-38. Or they can be created by modifying the start or end of other clips. *Id.* at 8:33-35 (“***transition selections*** at the beginning and end of the audio and video portions of each clip”) (emphasis added); Almeroth

¶¶167-69.¹

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

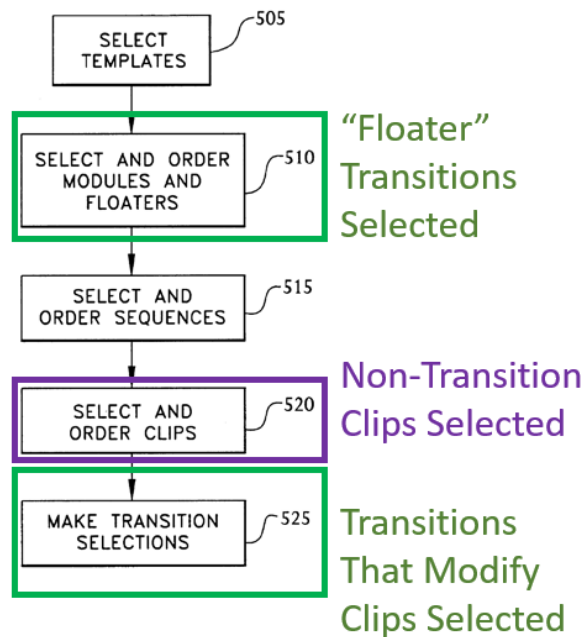


FIG. 5

Id. at Fig. 5 (annotated); 2:42-43; 7:3-7; Almeroth ¶170.

Computing devices: The transitions are selected by computer 20. *Id.* 14:18-

¹ The '185 patent specification does not mention “transitional clips,” but the appendix to the '597 patent's provisional application confirms that modification of existing clips qualifies as the claimed “transitional clips.” Ex. 1003 at 34 (“types of transitions (hard cut, fade-to-black, white-burst, and so-on)”).

21; 2:29-31; 9:40-41; 8:36-38; 8:48-51; Almeroth ¶171.

Transition associations: Each transition is associated with the clips that it transitions between. Marcus156, 1:39-50; 4:26-38; 4:48-51; 5:27-35; 7:47-49. Transitions are also associated with other clips based on aesthetic requirements derived from the content and format of the clips that are being transitioned between. *Id.* at 8:48-51 (“determine whether a direct transition between any two clips meets aesthetic requirements”); 3:36-39; Almeroth ¶172. The transitions are also associated with selected clips because they meet the system’s transition criteria or share tags. Marcus156, 4:26-38; 4:48-51; 5:27-33. Transitions that fulfill the criteria for a preferred or required transition of another video clip (and then are slotted next to the other clip in the final video programming) are “associated with” that clip. Almeroth ¶173.

5. Element 1[d]

Temporal sequence: Marcus156’s system determines a temporal order for the clips and transitions. Marcus156, 1:67-2:3; 2:28-31 (“A temporal sequence for the selected elements is designated.”); 4:22-34. Control tags for clips give a temporal structure to the concatenated video programming. *Id.* at cl. 8; Almeroth ¶175-76.

Concatenating: Marcus156 describes assembling clips and transitions into a sequenced final video as “concatenating.” Marcus156, 14:13-17 (“create a concatenated stream of audiovisual programming that has transitions”); 3:44-46

(“placement of media elements in the concatenated stream”); 5:35-39; 7:5-9; 9:23-32; Almeroth ¶177.

PO may argue that claim 1 requires selecting transitions before determining a temporal sequence. Marcus156 nevertheless discloses or renders obvious this claim element.

First, this is a method claim, and method steps are ordinarily not construed to require an order. *Interactive Gift Exp., Inc. v. Compuserve Inc.*, 256 F.3d 1323, 1342 (Fed. Cir. 2001). Nothing in the nature of these steps requires a particular order. An appropriate transition can be selected by knowing the order of the other clips first. Nothing in the intrinsic evidence suggests that the invention requires selecting transitions *before* sequencing the video programming. Claim 1 contains an express timing requirement elsewhere, where it specifies “*concurrently*,” so the applicant knew how to require a strict order when it wanted. Almeroth ¶179.

Second, Marcus156 discloses the claimed steps even if they must be performed in order. Floater/bridge transitions are selected before a sequence is determined:

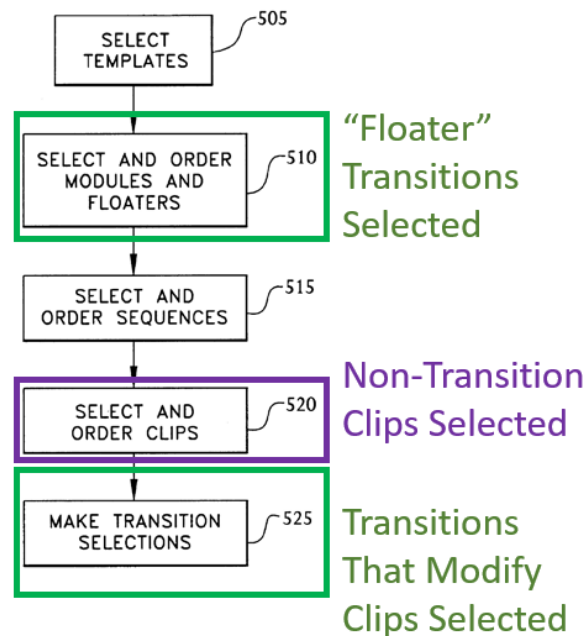


FIG. 5

Marcus156, Fig. 5 (annotated); 14:17-21. Additionally, after transitions based on modifying clips are selected, generating those transitional clips further sequences those transitions relative to other clips. Almeroth ¶180.

6. Element 1[e]

Defining video programming: An assembly program creates a stream of video from the selected video clips and transitional clips, thus “defin[ing] first video programming.” Marcus156, 4:13-15; 9:22-39; 14:7-11; Almeroth ¶182.

According to templates: A “template layer” is consulted during the selecting, determining, and concatenating steps (which are thus “performed according to” the templates):

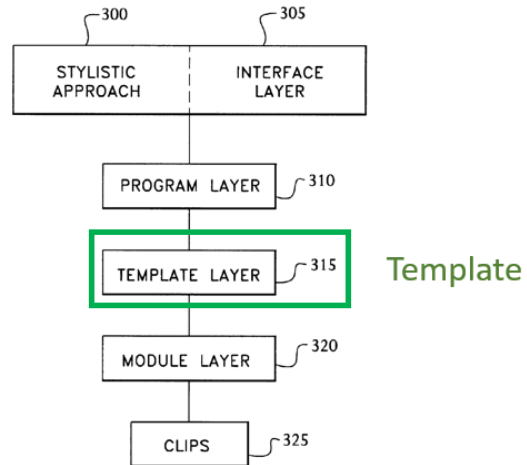


FIG. 3

Marcus156, Fig. 3 (annotated); 8:13-15 (“one or more suitable templates”); 9:9-21; Abstract. Templates are selected before selecting, ordering, and concatenating video clips and transitional clips:

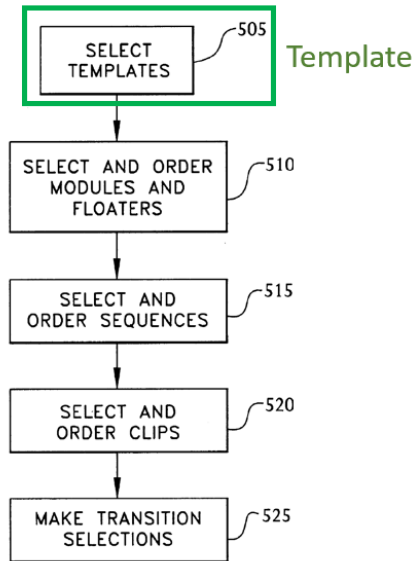


FIG. 5

Id. at Fig. 5 (annotated). The subsequent steps all comply with (“performed according to”) the requirements of applicable templates. *Id.* at 3:67-4:9; 8:19-25; 8:39-43; 9:2-6; 14:22-24; Almeroth ¶¶183-84.

Plurality of template variables: The templates comprise a plurality of template variables, namely the data contained in the templates. Marcus156, 3:67-4:3 (“a family of defining values”); Almeroth ¶184.

Information about the user: Marcus156’s templates include demographic and psychographic parameters that are used to customize the video stream to the viewer characteristics. Marcus156, 3:66-4:5 (template layer can target viewers by “demographic and psychographic terms.”); 6:33-40; 13:47-51 (parameters for “demographic or psychographic characteristics”). This is “information about the

user” at least because it restricts the intended audience for the video programming based on user characteristics. *Id.* at 4:9-14; 14:30-34; Almeroth ¶¶185-86.

Additionally, Marcus156’s system tracks information about individual users in its “database of viewer information” and in its “patient profile.” Marcus156, 10:12-20; 6:49-52. That information is used to select clips and transitions, and to sequence video programming. *Id.* at 10:16-20; 4:56-59; Almeroth ¶187.

While Marcus156 does not explicitly describe the “database of viewer information” or the “patient profile” as part of the template layer, a POSITA would not restrict the scope of the “templates” claimed in the ’185 patent to the specific data structures in the template layer in Marcus156. The plain and ordinary meaning of “template” does not require any particular data structure. The intrinsic evidence does not define “template.” 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. The appendix contemplates multiple templates addressed to different types of metadata. *Id.* at 35. PO has accused TikTok’s video recommendation algorithms as a whole of being a claimed “template.” Ex. 1018 at 68. Applying the same interpretation of “template” that PO uses for infringement purposes, a POSITA would have considered the “database of viewer information” or “patient profile” in Marcus156 to also be a “template.” Almeroth ¶¶188-91.

To the extent a specific data format is required to qualify as a “template,” it

would have been obvious to include Marcus156's "database of viewer information" or "patient profile" in its template layer. Among other things, 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 create video programming, and to simplify the system. A POSITA would have had a reasonable expectation of success in any such modification at least because it would entail a routine conversion of information in one data format to another. Additionally, Marcus156's system already uses data in a template format and includes software to import new clip and user data (Marcus156, 6:15-7:9), which would reduce the amount of time and effort to make the modification. Almeroth ¶¶192-93.

Viewing scenarios: Marcus156 includes "temporal organizational templates." Marcus156, 4:5-7. These include viewing scenarios of two or more selected clips in a temporal relationship to each other. *Id.* at 4:7-9 ("preferred temporal modes of presentation"); 4:14-21; 8:39-43 (clips sequenced by templates); 14:22-25 ("template that imposes sequential requirements"). They can describe an expected audience and subject matter for a sequence of temporally-ordered clips – a "scenario" for how the video programming is ultimately viewed. Marcus156, 4:9-14. The '185 patent specification does not contain the term "viewing scenario," and it only shows up once in the appendix to the '597 provisional application, where its meaning is not restricted. Ex. 1003 at 25. The plain and ordinary meaning of

“viewing scenario” encompasses this functionality in Marcus156. Almeroth ¶¶194-96.

Style requirements: Marcus156’s templates constrain the video style. Marcus156, 9:2-6 (“a specific style of editing”); 14:22-29; 3:41-46 (“aesthetic or programmatic choices”); Almeroth ¶197.

7. Element 1[f]

Presenting first video programming: Marcus156’s computer 20’s display device 50 is the claimed “display associated with one or more of said computing devices.” Marcus156, Fig. 1; 2:51-54. The same computer 20 that performs the searching and selecting can also present the video programming on its display. When the first video programming has been assembled, it is displayed on a display device 50 of computing device 20. There are several ways this happens.

First, the first video programming is presented as part of the user interface for a “List Editor/Previewer” program. Marcus156, 8:2 (“Programs are created from clips by an editor program.”); 6:3-7 (“decision/transition list editor 430 with built-in transition previewer”); 6:9-12; 7:11-15. Using that program, a video editor sees the assembled first video programming in a first area of a display associated with computer 20. Marcus156, 7:51-52; Almeroth ¶¶199-200.

Second, the first video programming is assembled by a “production player” or “assembly program” and then output as a video stream to be displayed, which can

be on the computer 20. Marcus156, 7:36-41; 7:59-60; 9:30-32; Almeroth ¶201.

Third, Marcus156 discusses users “receiving the programming online through a modem or Internet link” and then “playing of the video sequence.” Marcus156, 11:4-10; 14:35-40. This would necessarily be on a display. Almeroth ¶202.

Any of these three scenarios independently disclose or render obvious this element.

8. Element 1[g]

Concurrent frame display: Marcus156 describes creating media streams that include both video and still elements. Marcus156, 13:63-67; cl. 3. Marcus156’s system tracks different areas of the display screen that is playing a video, allowing the user to select a particular area of the display screen to request additional video programming. *Id.* at 14:46-52. Marcus156’s system varies the contents of the displayed media in different portions of the screen. *Id.* at 14:61-67. A companion window with user interface controls may appear alongside a playing video. *Id.* at 11:4-10. Marcus156’s “List Editor/Previewer” program allows the user to preview the video. *Id.* at 7:11-15; 6:3-5; 6:11-14; Almeroth ¶¶204-08.

Given this guidance, a POSITA would have found it obvious to modify Marcus156’s user interface to add a concurrent display of a frame. For example, to request “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,”

rather than being the video itself, could naturally be presented as a still frame of the portion of the video in another, concurrently-displayed area of the screen. As another example, the content in the companion window “which remains open and active during the playing of the video sequence” would naturally include a still frame of the playing video that included the superimposed strings. As yet a further 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. Almeroth ¶¶209-10.

A POSITA would have been motivated to add this interface to Marcus156. Concurrent display allows the user to view more relevant information at once. Showing frames next to the video allows more efficient navigation through the playing video stream. Displaying frames gives a user more time to click on the screen to get more information on what is being displayed. Enhancing the interface follows Marcus156’s instructions that “[t]he user interface is preferably...graphical in nature.” Marcus156, 6:11-12. The similar concepts in Marcus156, above, would have also motivated a POSITA to improve Marcus156’s system in a similar manner. Adding a frame display would have improved the user interface and enabled new capabilities. Almeroth ¶211.

A POSITA would have had a reasonable expectation of success in making such a software modification because Marcus156 already discloses software

techniques to extract a frame from a video to prepare transitions, and the same technique could be used to generate frames for concurrent display. Marcus156, 7:54-57. The techniques required to implement this software modification were routine and well-understood. Almeroth ¶¶212-13.

9. Element 1[h]

Receiving user frame selection: A user selects displayed elements by mouse click. Marcus156, 14:35-43; 14:46-52. Because the “List Editor/Previewer” provides “the ability to change, add to and preview the EDL,” it entails selecting UI elements. *Id.* at 7:11-15. And Marcus156’s system supports “viewer input” to search for video. *Id.* at 14:59-67; Almeroth ¶¶215-18.

Marcus156’s system, as modified above, would track a user selection of a frame and for computer 20 to receive data about that selection. Almeroth ¶219. A user could click on a frame in the same manner that Marcus discloses a user clicking on a playing video to request “more detailed information.” Marcus156, 14:35-58.

A POSITA would have been motivated to make this modification. Adding interactive feedback to frame selections improves the system’s functionality and responsiveness. Marcus156 already discloses a user clicking on a frame in a running 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, and doing so would give the user more time to click on the frame, which may have already

passed in the main video but is still visible in the concurrently-displayed frame. *Id.* ¶220.

A POSITA would have had a reasonable expectation of success in making such a modification. Determining what displayed user interface element a user interaction corresponds to is a routine software function from the first graphical user interfaces such as the Xerox Alto from 1973. The existing graphical interface code used by Marcus156's "List Editor/Previewer" or "interactive system" would include such code, or it could be added with minimal programming effort, as could any necessary code to transmit the selection. *Id.* ¶221.

10. Element 1[i]

Selecting second video clips: When the user interacts with video programming that was already generated by the system, Marcus156's system performs additional interactive searches and video selections, including the generation of a new stream of video programming, which entails the selection of multiple second video clips. Marcus156, 14:40-45 (new search performed based on displayed information from first search); 14:46-55 (user click on screen of first search results triggers second search); 7:11-15 (List Editor/Previewer 430 can change video programming in response to user interaction); Almeroth ¶223.

Although column 14 mentions "a predetermined program," Marcus156's disclosure is not so limited. Marcus156 describes a user interacting with video

programming and generating new programming by selecting a portion of the screen. It is agnostic about whether that programming is “predetermined” or 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. The same benefits of user interaction with a “predetermined” program to search for “particular information” about that program equally apply to searching for additional information about an interactive or targeted program. There would have also a reasonable likelihood of success, since the same selection mechanics can be applied to either type of video. Almeroth ¶224.

Based on associations: The selection of a second video clip is based on the stored associations between the prior, selected clip and the new, second clip, because both videos share the same or similar tags, including at least at higher levels of Marcus156’s tag hierarchy. Marcus156, 9:16-21; 3:26-30; 14:40-55. 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; Almeroth ¶225.

11. Element 1[j]

Selecting, sequencing, and concatenating second video clips using

templates: The same systems, data (in particular, 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 to perform the same actions on a plurality of second video clips. The material in 1[b], 1[d], 1[e], and 1[i] applies here too. Almeroth ¶227.

12. Element 1[k]

Templates: The templates used in follow-up searches in Marcus156's system can be the same as or different from the initial search (discussed in 1[e]). While searching, the system "selects and orders one or more suitable templates" first. Marcus156, 8:13-15. The interactive searches at 14:35-67 use templates, just as any initial search does. As in 1[e], those templates contain the required (a), (b), and (c) information. Almeroth ¶229.

13. Element 1[l]

Presenting second video programming: 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. Marcus156, 14:52-55; 14:61-67; Almeroth ¶231. Additional detail is in 1[f].

B. Claim 2

As in 1[c], Marcus156 includes transitional clips in the first video programming. As in 1[j], the same types of transitional clips are also included in the

second video programming. Additional details are in 1[b], 1[d], and 1[e]. Almeroth ¶223.

C. Claim 3

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, claim 3 is disclosed by Marcus156 for the same reasons claim 2 is. Almeroth ¶225.

D. Claim 4

Marcus156 discloses transitional clips that are hosting clips. Video clips in database 100 can be tagged with their actors. Marcus156, 3:21-25; 14:30-34. A media tag “TALKING HEAD” can be used to select transitional clips for inclusion in the video. *Id.* at 4:39-48. A talking head clip has a person—a host—talking. The appendix to the ’597 provisional describes such talking head clips as hosting clips. Ex. 1005 at 102. The term “hosting clip” is also described broadly in the appendix. *Id.* at 34, 71, 74 (“Any form of entertainment inserts/content or hosting segments can be the subject of this model.”). Clips with actors or talking head clips can be “bridge” or “floater” transitional clips. Marcus156, 4:22-38; claim 2 above. When such a clip is selected as a linking clip, it is a “hosting clip” that is also a “transitional clip.” Almeroth ¶¶237-38.

Marcus156’s hosting clips are “consistent with at least one template variable.”

Like other transitional clips, the hosting clips must comport with the contents of the selected templates to be included in the video programming. If a hosting clip is inconsistent with the templates then it will not be included. Marcus156, 3:41-46; 8:19-25; Almeroth ¶¶239-40; 1[e] above.

To the extent that Marcus156 does not disclose this claim, it renders it obvious. A POSITA would have been motivated to use hosting clips as transitional clips. Including hosting clips allows the video programming to emulate television shows with hosts, such as nature documentaries, late-night talk shows, or sequences of music videos hosted by a video jockey. Anybody familiar with radio DJs, music video ranking shows, America's Funniest Home Videos, or similar programming at or (well) before the time of invention would understand the ubiquity of using hosting clips as transitions. Using hosting clips 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 (allowing for introductions to video programming users may otherwise be unfamiliar with). Including hosting clips would allow for additional types of transitions for comedic or dramatic effect. Almeroth ¶241.

A POSITA would have had a reasonable expectation of success in making any needed modification to Marcus156. Given Marcus156's existing disclosure of tracking talking head clips, a POSITA would have reused the existing data about

those clips. The existing templates could have been easily modified to use the hosting clips as transitional clips, given the sequencing and clip association capabilities detailed in Marcus156. Almeroth ¶242.

E. Claim 5

Marcus156 discloses or renders obvious this claim for the same reasons given above in claim 4. Almeroth ¶244.

F. Claim 6

The temporal organization templates restrict how clips are placed at specific locations in the programming, as described in 1[e]. The template variables define a “preferred sequence” of clips ordered by content type. Marcus156, 4:14-21; 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. Almeroth ¶246.

G. Claim 7

Marcus156’s templated sequencing limitations meet each of the required three categories. The limitations may be length requirements or requirements for presentation within a particular time frame. Marcus156, 4:65-5:2; 5:27-33. Order of clips can also be limited. *Id.* at 4:14-21; Almeroth ¶248.

H. Claim 8

Video clips in the database (“assets”) may be tagged as advertising or

promotion. Marcus156, 3:52-55. Clips may be classified as “market-driven,” which a POSITA would understand to include advertising. *Id.* at 3:55-57. Marcus156 says that customized advertising programming is valuable, which is an indication that its system included advertising video clips. *Id.* at 1:19-26. 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, one or more clips in Marcus156’s system contain advertising. Almeroth ¶250.

I. Claim 9

As in 1[e], Marcus156’s templates contain demographic and psychographic information. And the viewing history of users is tracked in the “database of viewer information.” Marcus156, 10:12-16 (“the history of items viewed by the viewer”); Almeroth ¶252.

J. Claim 10

Templates are used to select clips based on associations with other clips, including by using a tag of one clip as a requirement for another clip. Marcus156, 3:67-4:5; 4:56-64; 4:66-5:2; Almeroth ¶¶245-55.

K. Claim 11

Media tags (aka “header information”) are attached to clips. Marcus156, 3:4-8; 5:3-4; 5:27-33; Almeroth ¶257.

L. Claim 12

Tags include context, subject matter, and demographics. Marcus156, 3:21-26; 4:41-48. Demographics include source or region. *Id.* at 14:30-33 (“geographic origin”). Tags include psychographic information. *Id.* at 13:47-51 (“demographic or psychographic characteristics”); 4:61-63 (“psychographic requirements”); Almeroth ¶¶259-61.

M. Claim 13

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. Marcus156, 3:44-46 (“media elements in the concatenated stream”); 5:35-37 (“data stream before or after the clips are concatenated”); 7:7-10 (“The transitions are then executed and inserted into the stream....”); 14:13-17 (“concatenated stream of audiovisual programming that has transitions”); 9:30-32; cl. 1; Almeroth ¶¶263.

During prosecution, the examiner stated: “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. This was clear error by the examiner. As detailed above, Marcus156 discloses expressly refers to its video programming as a “stream.” Almeroth ¶¶264.

N. Claims 14-20

Marcus156 renders claims 14-20 obvious for the same reasons as discussed in corresponding claims. Almeroth ¶¶265-271.

Claim	Corresponding Claim
14	6
15	7
16	8
17	9
18	10
19	11
20	12

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

This ground relies on Marcus156 as a base reference, with modifications according to the teachings of Rosen.

A. Proposed Modification

For either the “interactive system” described in Marcus156, 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 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 modified 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

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

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. *Id.* ¶275.

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. *Id.* ¶276.

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, 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. Almeroth ¶277.

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. *Id.* ¶278.

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 would have been within a realm of minor design decisions that a POSITA would naturally make during the coding of such a system. *Id.* ¶279.

C. Reasonable Expectation of Success

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 reasons.

Software modification would be predictable: The modification can be implemented entirely in software and requires no hardware changes. 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 [0035]-[0161]; Almeroth ¶281.

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, 7:54-57. And it teaches that its techniques can be used with still images (like frames). *Id.* at 13:62-65; cl. 13. Marcus156's system also already processes user selections of displayed elements on a screen. *Id.* at 14:37-43. 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. Almeroth ¶282.

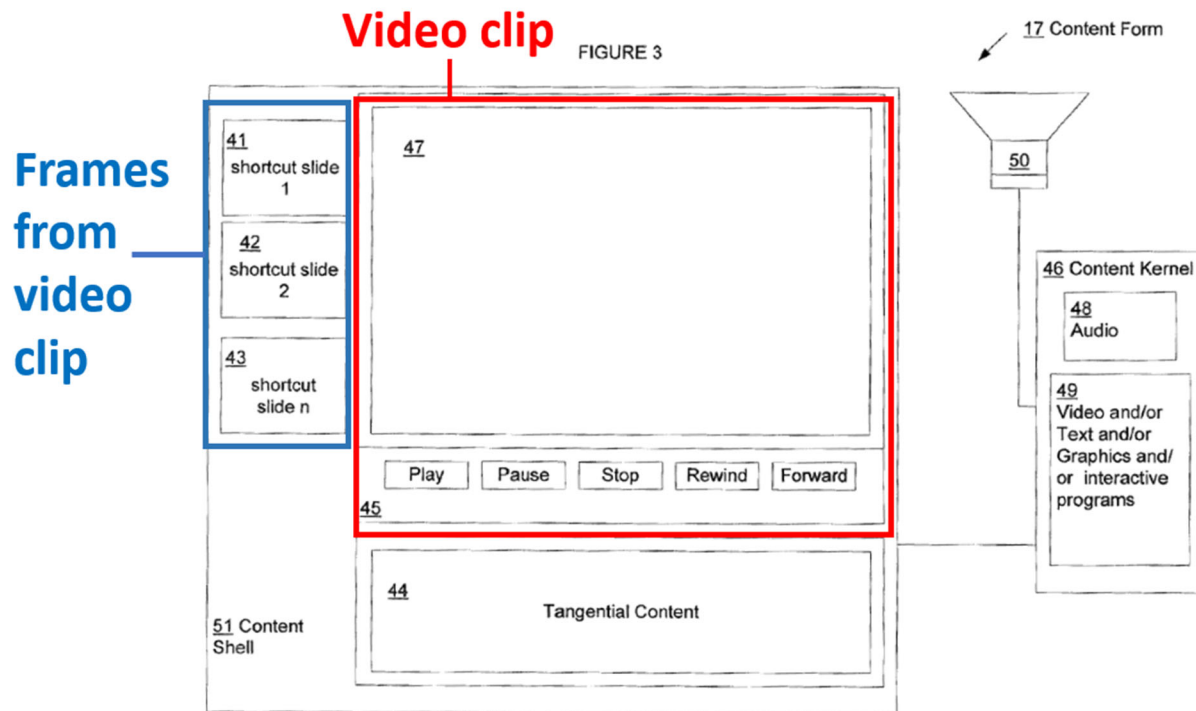
D. All Claims and Claim Elements Other Than 1[g], 1[h], and 1[i]

All claims 1-20 and claim elements other than element 1[g] , 1[h], and 1[i] are disclosed or rendered obvious by Marcus156 for the reasons given above in ground

1.

E. Element 1[g]

Rosen 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.

The **shortcut slides** contain **frames of the playing video**. *Id.* [0027] (“At any desired point in the video, the author may choose and extract a playback point P0 (or **frame**) from the video 64.”) (emphasis added); [0028]. Therefore, Rosen discloses the concurrent display of **a playing video** and **a frame from that video**, as required by the claim language. Almeroth ¶¶285-86.

F. Element 1[h]

In Rosen, the user may select a **shortcut slide** to initiate an action. Rosen [0023]; [0027]. Rosen discloses the computing device receiving a user selection corresponding to **a frame** of the **playing video**. Almeroth ¶286.

G. Element 1[i]

Rosen processes a user selection of a **shortcut slide**, which can trigger a wide variety of actions. Rosen [0031] (“A specific event is then chosen 114 for the shortcut box.”). Applied to Marcus156, Rosen’s teachings would allow an additional supplemental search to be performed by the system when a **frame** is selected by the user. Almeroth ¶288.

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

This ground relies on Marcus156 as a base reference, with modifications according to the teachings of Sull. This ground is an alternative to ground 2, using Sull instead of Rosen for teachings relevant to elements 1[g], 1[h], and 1[i].

A. Proposed Modification

Similar to ground 2, for either the “interactive system” described in Marcus156, 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 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

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

(1) Showing frame content provides more information to the user; (2) showing frame content time-shifts the display of information to prevent a user from missing critical information; (3) showing a frame, allowing a user selection, and performing a second search allows the video programming to be progressively refined; (4) design choice: Described above in ground 2.

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. Almeroth ¶295.

C. Reasonable Expectation of Success

A POSITA would have had a reasonable expectation of success in modifying Marcus156's system according to the teachings of Sull because:

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

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.

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

All claims 1-20 and claim elements other than element 1[g], 1[h], and 1[i] are disclosed and rendered obvious by Marcus156 for the reasons given above in ground 1.

E. Element 1[g]

Sull displays **a playing video** next to **frames from that video** that represent **multimedia bookmarks**:

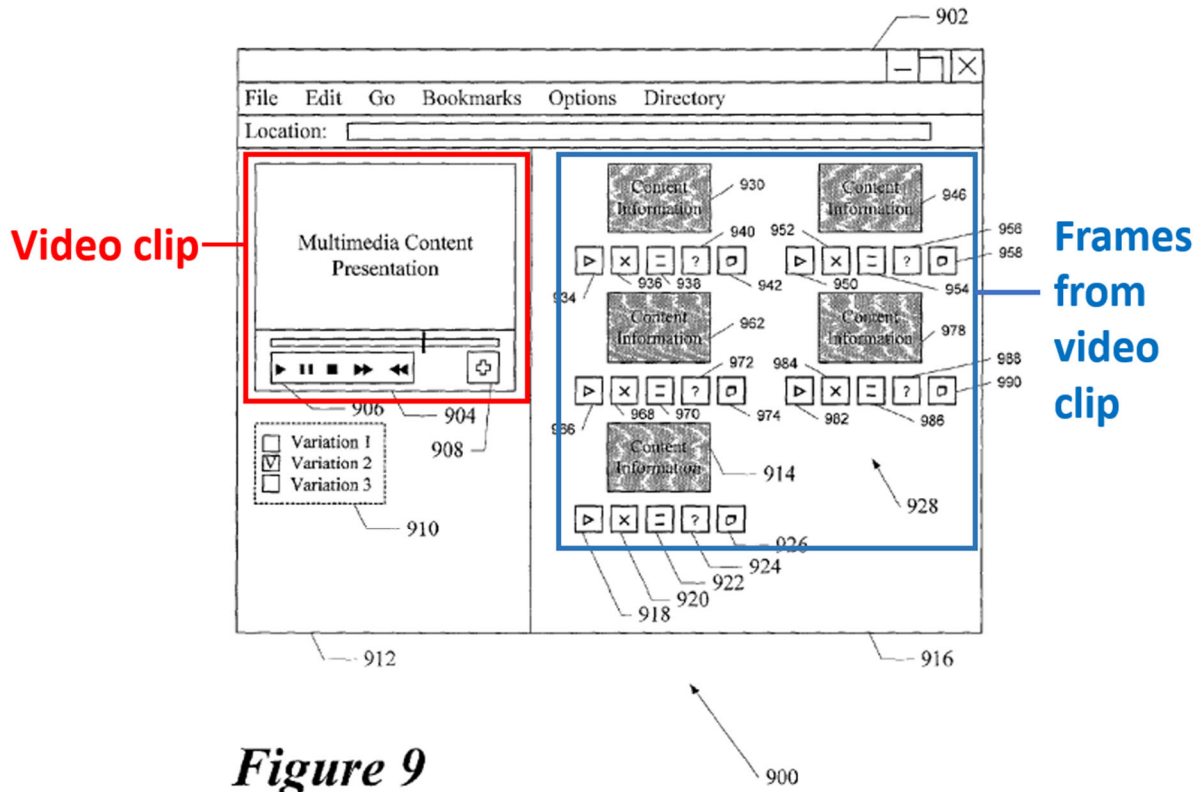


Figure 9

Sull at Fig. 9 (annotated); [0205]; [0220]; Almeroth ¶302.

F. Element 1[h]

A user selects one of the displayed **multimedia bookmarks (frames)** and that selection is received at a computing device. Sull [0194]; [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.”); Almeroth ¶303.

G. Element 1[i]

The user selection of the **multimedia bookmark (frame)** is communicated to a computing device, which searches for additional clips based on the selection. Sull [0225] (quoted above); [0221]; [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.”); [0060]; [0214] (“[A] frame-based video search may be employed to find other video segments similar to the particular video segments.”); Almeroth ¶304.

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

A. Proposed Modification

To the extent that the combined system of Marcus156 and Rosen somehow does not fully render obvious a progressive video search in 1[i], the combined system of Marcus156 and Rosen would be further modified to perform additional searches for video clips (using the existing capabilities of Marcus156 and Rosen, plus the additional teachings of Hosken) based on the user selection of a displayed frame.

B. Motivation to Combine

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.

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 (advantage of personalized recommendations).

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. Almeroth ¶308.

Similar field of art: The '185 patent, Marcus156, and Hosken all pertain to automated video recommendation systems. 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.”); Almeroth ¶309.

C. Reasonable Expectation of Success

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:

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

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 to add progressive search refinement to Marcus156 because all of the building blocks are already present. Marcus156, 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.”); Almeroth ¶312.

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

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.

E. Element 1[i]

A user selection of one item in search results is used in combination with “associations” between content items, to produce another search results set. Hosken at 9:4-8; 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. 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.”); Almeroth ¶314.

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

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. Nothing in Sull’s techniques would preclude also using Hosken’s techniques. Almeroth ¶315.

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

A. Proposed Modification

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.

B. Motivation to Combine

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. Byers at 2:51-52; Marcus156, 3:52-57; Almeroth ¶318.

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.”); Almeroth ¶319.

Same field of art: Like Marcus156 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.”); Almeroth ¶320.

C. Reasonable Expectation of Success

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. Almeroth ¶321.

Byers provides explicit instructions on how to modify clips to insert advertising: Byers discloses underlying mechanisms to insert the advertising

content. Byers at 3:67-4:5; 5:60-66. The hardware used in this process is similar to Marcus156's disclosed hardware. Byers at 5:45-67; Marcus156, 2:51-56. Marcus156's disclosed hardware could perform Byers' video editing algorithms. Almeroth ¶322.

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, 7:47-63. The same technique could be used to add advertising information. Almeroth ¶323.

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

Byers discloses an “advertiser database” that contains advertising-related information. Byers at 4:19-25. The information in the database is used to dynamically insert ads into a video stream. *Id.* at 2:48-51; 3:1-4; Almeroth ¶¶324-26.

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

For the same reasons given above in the context of grounds 5 and 6, 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. Nothing in Sull or Hosken’s techniques would preclude also using Byers’ techniques. Almeroth ¶327.

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

This ground relies on Ferman as a single-reference obviousness ground for all claims except 8 and 16. Unless otherwise noted, Ferman discloses each claim and claim element.

A. Claim 1

1. Element 1[pre]

Ferman’s system provides customized interactive video programming to a user. Ferman [0092]; [0106]; Almeroth ¶330.

2. Element 1[a]

Databases of video clips: Ferman’s system includes a program description scheme, which is a database of video clips. Ferman [0094]; [0114]; [0117] (“The search engine uses the information contained in the program description scheme to retrieve programs on the basis of their content.”); [0264] (“program descriptions...describing...video content available from which a filtering agent 520 selects the appropriate content for the particular user(s).”); [0266]; [0315]; Almeroth ¶332.

Searching in response to a user request: Ferman’s system searches the program description scheme in response to a user request. Ferman [0114] (“enable

the user to browse, search, and filter audiovisual programs.”); [0264] (“[A] a filtering agent 520 selects the appropriate content for the particular user(s).”). The user request may be triggered by an interaction or may be a user preference description or a set of filtering preferences. *Id.* [0315]; [0266] (“[P]referred programs can be determined as a result of filtering program descriptions according to user’s filtering preferences.”); Almeroth ¶334.

Computing devices: The search is performed by a computing device (system 16) running a software agent. Ferman [0094] (“a software agent that in turn performs a search and filtering on behalf of the user”); [0092]; Fig. 1. Components of the computing device are shown in Figure 2:

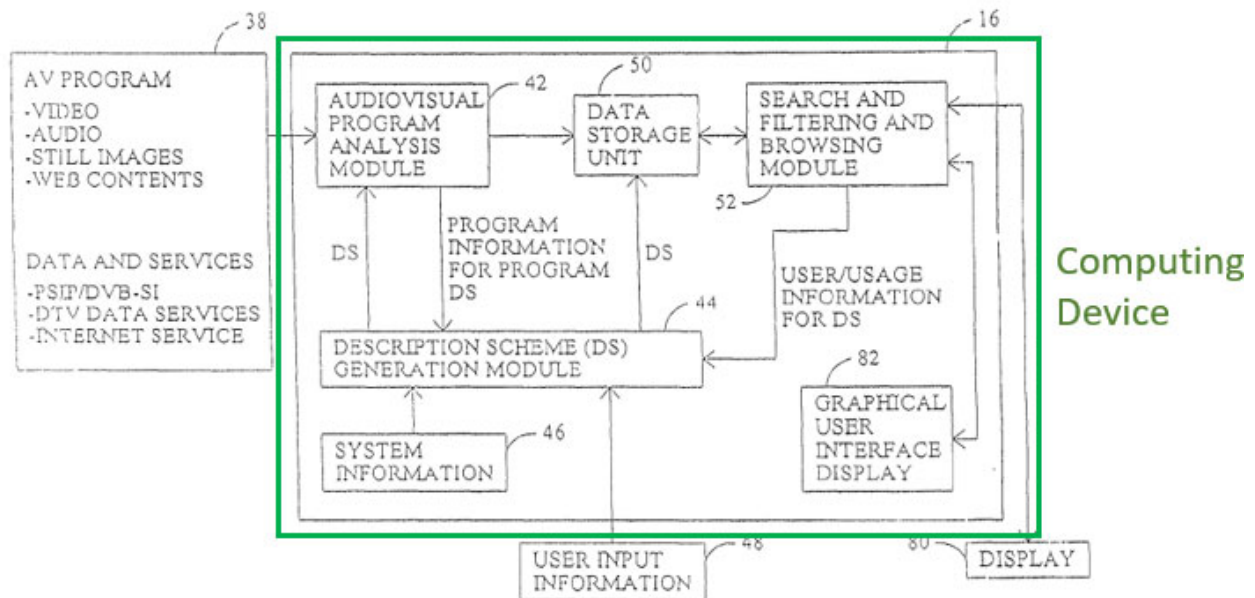


FIG. 2

Id. at Fig. 2 (annotated); [0007]; [0102] (“The system 16 may include any device(s)

suitable to receive any one or more of such programs 38[.]”); Almeroth ¶335.

Frames: The clips in the database are comprised of frames, like any other video. Ferman [0094]; [0107]; [0252]; Almeroth ¶336.

3. Element 1[b]

Selecting video clips: Ferman’s system selects an initial set of video clips (“preferred programs”) that fulfill the user request. Ferman [0117] (“The search engine uses the information contained in the program description scheme to retrieve programs on the basis of their content.”); [0264] (“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]; Fig. 28 (“FILTERING AND SEARCH” component leads to “PREFERRED PROGRAMS”); [407]; Almeroth ¶338.

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.* [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). Ferman refers to this functionality of breaking down a long-form program into subsequences as “filtering.” *Id.* [0119].

The result of these operations is a set of recommended video clips that are presented to the user. Ferman 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.”); Almeroth ¶¶339-40. This is the claimed “plurality of first video clips” in Ferman.

Clip associations: Each first video clip 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. Ferman [0103]; [0104] (“generating the preferred views and storing associated data in the program description scheme.”); [0107]. These are the claimed “media content comprising one or more second video clips” in Ferman. Almeroth ¶341.

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); Almeroth ¶342.

4. Element 1[c]

Selecting transitional clips: Ferman’s system filters and segments a video program “into story units, scenes, and shots.” Ferman [0252]. The set of segments includes multiple major events in the video clip, including events which transition between other clips. *Id.* [0094]; [0252] (syntactic structure description scheme segments clips via a “segment tree [which] may be used to define[] a temporal structure of a video program”); [255] (semantic description scheme segments clips into events, reflected in an event description scheme). In particular, 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.* [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. A POSITA would consider these video segments, which are thematically and structurally related to other surrounding video clips, to be “transitional clips.” Almeroth ¶343.

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. 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.” Ferman [0089]. Moreover, doing so would provide more compelling and cohesive video programming. A POSITA would have had a reasonable expectation of success in making any necessary changes to Ferman's system to more explicitly include transitions, including because Ferman already categorizes sub-sections of a video program by their contextual relationship to other sections. *Id.* [0094]; Almeroth ¶344.

Computing devices: For the same reasons as 1[a], the selecting is done by the same computing device that performs the searching step.

Transition associations: The transitional clips are associated with the first video clips via at least the segment description scheme. Ferman [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, they are associated at least via the program description scheme and program view descriptions, for example by sharing similar media tags. *Id.* [094]; [0104]; Almeroth ¶346.

5. Element 1[d]

Temporal sequence: The selection of video clips and transitional clips in turn determines a temporal sequence for the concatenated video programming. Ferman [0252] (“Such a segment tree may be used to define[] a temporal structure of a video program.”). Additionally, the video programming is sequenced based on

information in the program description scheme and/or event description scheme. *Id.* [0095] (sequencing program views using the program description scheme); [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.”); Almeroth ¶348.

Concatenating: The responsive video clip and transitions are concatenated and presented as a stream of video programming to the user. Ferman [0607] (“PlayStream Play content from input stream”); [0107] (“summarization functionality of the system...prepare[s] a summary”); [0111]; [0114] (“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.* [0267] (15-minute highlight); Fig. 10; [0107] (5-minute highlight). The act of generating a single video highlight from multiple previously separate clips qualifies as “concatenating.” Almeroth ¶349.

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. A POSITA would have been motivated to make the modification because it would

allow its system to achieve better VCR-like functionality. Moreover, concatenating clips would lead to a more coherent viewing experience corresponding, for example, to a single highlight video rather than disparate clips. A POSITA would also have had a reasonable expectation of success in making the modification, because Ferman's system already has video recording capability. Ferman [0107]; Almeroth ¶350.

6. Element 1[e]

Defining video programming: As above, Ferman's system selects, determines, and concatenates clips to form a final video stream ("first video programming"). Ferman [0102]; [0105] (filtered video programming output by SFB module); [0107] (summarized highlight video programming output); [0111]; [0114]. To the extent PO contends that Ferman's system itself fails to define video programming consisting of video data, 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 video data (e.g., via transcoding) were well-known at the time of the invention. Almeroth ¶352-53.

According to templates: These steps are performed according to at least a template referred to in Ferman as the “user description scheme.” Ferman [0097] (contents of user description scheme); [0097] (“[C]ompare the information contained in the user description scheme 20 to the available information contained in the program description scheme 18 to make selections.”); [0407]; Almeroth ¶354.

As discussed in 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 a POSITA would have understood to be a claimed “template.” Ferman [0196] (XML format); [0273] (“A description scheme is a data model of descriptions.”); Almeroth ¶355.

Plurality of template values: The templates comprise a plurality of template variables, namely the data contained in the user description scheme. Ferman [0196]; [0273]; Almeroth ¶355.

Information about the user: The user description scheme template includes information about the user. Ferman [0108] (“user history” and “video viewing habits”); [0112]; [0333]; [0334] (“user’s preferences”); [0394]; [0395]; [0407]; [0410]; cl. 1 (“user attribute information”); Almeroth ¶356.

Viewing scenarios: The user description scheme includes viewing scenarios of multiple clips in a temporal relationship. Ferman [0267] (“user’s preference for continuous playback of the entire program versus visualizing a short summary of the

program”); Almeroth ¶357.

Moreover, other templates in Ferman’s system, including the segment description scheme and the semantic structure description scheme, include temporal viewing scenarios for clips. Ferman [0252] (“[A] segment tree may be used to defined a temporal structure of a video 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.”); Almeroth ¶358.

Style requirements: The user description scheme template includes style requirements. Ferman [0318] (“the desired configuration for presentation, such as volume or any other functionality”). Additionally, other templates that are used in the process of generating video programming also include style requirements. *Id.* [0266] (“preferences for setting up the device in relation to the type of content being presented”); [0268] (“favorite actors, directors”); *id.* (“favorite program category”); [0269] (“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”); Almeroth ¶359.

7. Element 1[f]

Presenting first video programming: 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 [0250]. For example, **playing video** in the “Highlight View” screen is shown next to **frames from that video**:

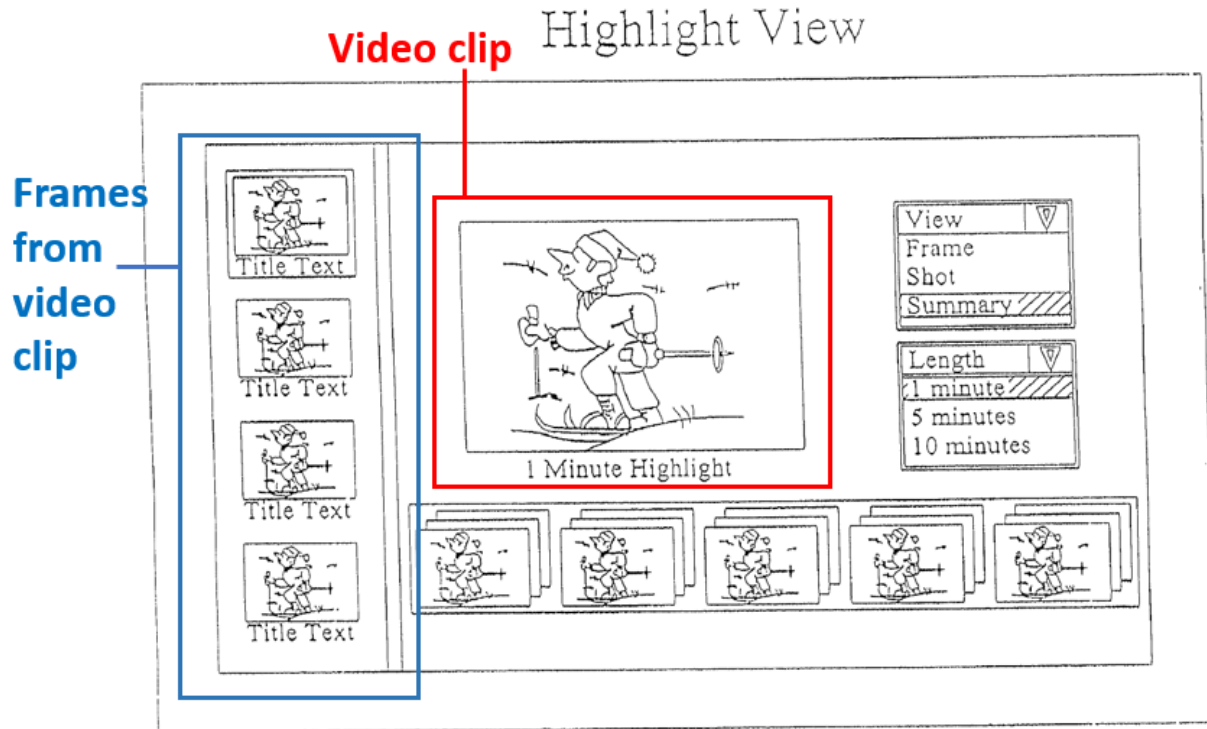


FIG. 10

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]. The event view has similar functionality. *Id.* at Fig. 11; [0124]; Almeroth ¶361.

Moreover, to the extent PO argues there is any ambiguity in the Ferman’s disclosure, 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 [0124]. Ferman discloses playing highlight videos and displaying frames corresponding to the videos, as well as a UI (Fig. 10) that includes areas for both of those UI elements. Almeroth ¶362.

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.* [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. This would increase the interactivity and pertinence of the videos generated by Ferman's system. Almeroth ¶363.

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. Ferman 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.* [0107]. The change would be entirely in software. No hardware

would need to be modified. Almeroth ¶364.

8. Element 1[g]

Concurrent frame display: As discussed immediately above, Ferman’s system displays running video alongside frames from the video, in distinct areas of the display, in at least the highlight view and event view.

9. Element 1[h]

Receiving user frame selection: In Ferman’s user interface, a user can click on a frame to make a selection, including one that triggers the generation and display of a highlight reel or an event reel. Ferman [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....”); [0105]; [0112]; [0114]; [0152]; [0154]; [0250]; [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.”); [0309]; Fig. 10; Almeroth ¶366.

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 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 second video clips based on associations: The user selections are used to populate the user description scheme and therefore to search for and select additional video clips. Ferman [0097] (user description scheme “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”). This selection is based on associations between the clips because the search results conform to the user description scheme, which contains information about the first clip. Almeroth ¶368.

Additionally, the user interaction with a frame can also directly trigger the generation of new video, such as another highlight reel or event reel. *Id.* [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.”); [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.”). 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. Almeroth ¶369.

To the extent that PO argues that Ferman does not explicitly disclose this element, 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, sequencing, and concatenating second video clips using templates: 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 also used in the second search to perform the same actions on a plurality of second video clips. Therefore, all of the evidence and analysis reviewed above in 1[b], 1[d], 1[e], and 1[i] applies here too. In response to a user selection, Ferman's system can generate a highlight reel or event reel with a plurality of video clips in a temporal order, using Ferman's templates. Ferman [0107] (generation of highlights). Almeroth ¶¶372-73.

12. Element 1[k]

Templates: The templates used in follow-up searches in Ferman's system can be the same or different as the initial search. And for the reasons in 1[e], those templates contain the claimed (a), (b), and (c) information.

13. Element 1[l]

Presenting second video programming: After defining second video programming, the programming is presented on the computing device's display, just

like first video programming. *See* 1[f] above.

B. Claim 2

As described in 1[c] above, Ferman's system includes transitional clips in the first video programming. As in 1[j], the same types of transitional clips are also included in the second video programming. *See* 1[b], 1[d], and 1[e] above.

C. Claim 3

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 as claim 2, claim 3 is also disclosed by Ferman.

D. Claim 4

Video clips in Ferman's system can focus on a particular person or another host subject, like a lion:

Character/Object View

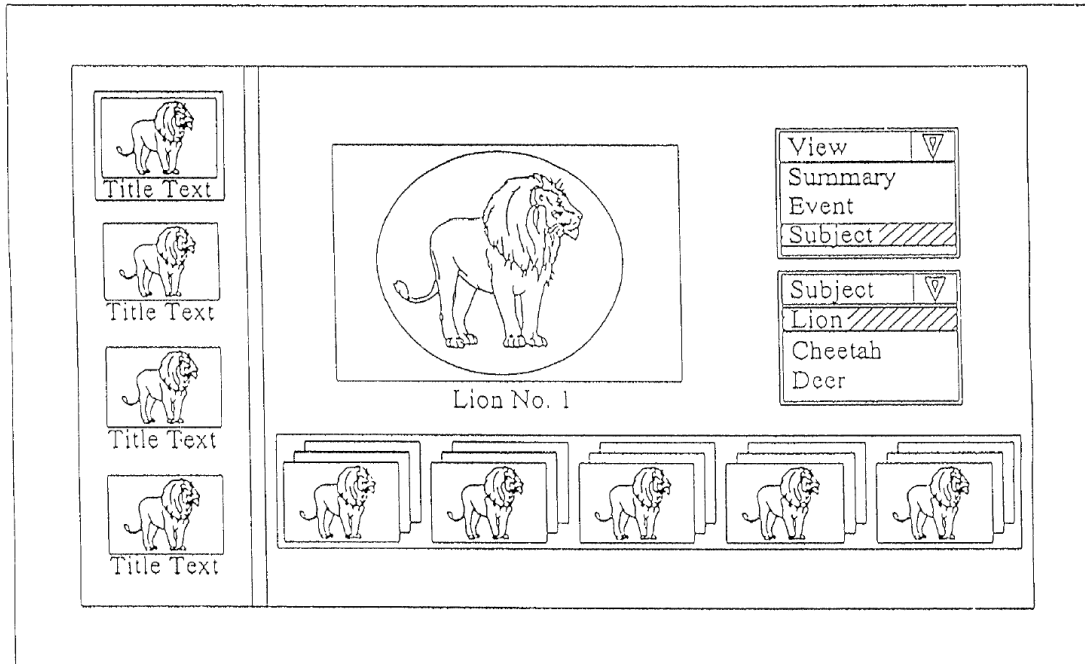


FIG. 12

Ferman at Fig. 12; [0124] (character/object view); [0095] (“In the close-up view, a certain image object, e.g., a famous basketball player such as Michael Jordan, can be viewed up close....”). As with the “talking head” clips in Marcus156 (ground 1), a POSITA would consider these to be “hosting clips.” Almeroth ¶383.

To the extent that Ferman does not disclose this claim, 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. *Id.* ¶384.

E. Claim 5

Ferman discloses or renders obvious this claim for the same reasons as claim 4.

F. Claim 6

The template variables corresponding to viewing scenarios discussed in 1[e] also include sequencing limitations for incorporating video clips at specific locations in the first or second video programming. Ferman [0095]; [0252] (“[A] segment tree may be used to defined a temporal structure of a video 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.”); [0614] (compilation of sequencing-related variables). 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. Almeroth ¶388.

G. Claim 7

Ferman’s templated sequencing limitations meet each of the three categories. The limitations may be length, order, or time frame requirements. Ferman [0252]; [0255]; [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); Almeroth ¶390.

H. Claim 9

Ferman’s templates include demographic information. Ferman [0097]

(“personal information about the particular user, such as demographic and geographic information, e.g. zip code and age”); *id.* (“demographics”); [0110] (“demographic information such as age, zip code, etc.”); [0118]; [0221] (“The descriptor <Age> specifies the age of a user.”); [0224] (“The descriptor <Gender> specifies the gender of a user.”); Almeroth ¶392.

The templates include psychographic information, namely the user’s “emotional state.” Ferman [0274] (“different emotional states of the user (e.g., happy mood versus tired or sad”); [0275] (“different emotional states...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. Almeroth ¶393.

The templates also include a viewing history. Ferman [0097] (“information regarding the user's viewing history such as for example browsing history, filtering history, searching history, and device setting history”); [0212] (“the history of a user’s browsing related activities”); [0218] (“the history of a user’s device related activities”); Ferman [0258]; [0259]; [0271]; Almeroth ¶394.

I. Claim 10

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 [0105] (“SFB module 52 may perform filtering, searching, and

browsing of the programs 38, on the basis of the information contained in the description schemes.”); [0270] (“It is preferable that media content descriptions are consistent with descriptions of preferences of users consuming the media.”); Almeroth ¶396.

J. Claim 11

Ferman’s system uses media tags associated with clips in the program profile to categorize clips and to associate them with other clips. Ferman [0094] (program profiles include “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”); [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]; cl. 1 (program attribute information). These are “media tags” as described in the appendix to the ’597 provisional. 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.”); Almeroth ¶398.

K. Claim 12

Ferman’s media tags include context and subject matter data. Ferman [0278]

(“media type, media source, or content”); [0282] (“preferred theme, or point of view, of a segment”); [0303] (same); [0601] (“title 746, location 748, creator 750, date 752, and actor/director”); Almeroth ¶400.

The media tags also include source information. Ferman [0283] (“source of the content”); [0284] (“source distinction and identification”); [0305] (“source preference description describes the preferred source of multimedia information”); [0602] (history of sources); Almeroth ¶401.

Moreover, the media tags include data about region and language. Ferman [0306] (“preferred program title, genre, language, actor, creator of the program”); [0328]; [0329] (“Location”); [0584]; [0600] (“genre, language, and country of origin”); Almeroth ¶402.

L. Claim 13

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. Ferman [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. [0091]; [0107] (5-minute highlight stream); [0267] (15-minute highlight stream); Fig. 10; Almeroth ¶404.

M. Claims 14-15 and 17-20

Ferman renders claims 14-15 and 17-20 obvious for the same reasons as discussed in corresponding claims:

Claim	Corresponding Claim
14	6
15	7
17	9
18	10
19	11
20	12

XVI. Ground 9: Ferman and Byers Render Claims 8 and 16 Obvious

A. Proposed Modification

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. Advertising-related data would also be added to the program description scheme and the user description scheme.

B. Motivation to Combine

Commercial benefit to advertising: Because adding advertising to videos earns money, a POSITA would have had an economic motivation to do it. *E.g.*, Byers at 2:51-52; Ferman [0088-92]; Almeroth ¶413.

Commercial benefit to user-targeted advertising: Ferman’s user-targeted recommendations facilitate advertising that is customized to the demographics, interests, or other facets of the user, thus increasing the advertising’s pertinence and click-through or conversion rate. Byers at 2:57-67 (“One problem with product placement is that the placement is not tailored for an individual user....”); Almeroth ¶414.

Same field of art: Like Ferman and the ’185 patent, Byers concerns dynamic video programming. Byers at 1:40-43 (“[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 [090]; [091]; [0102]; Almeroth ¶415.

C. Reasonable Expectation of Success

Software modification would be predictable: This change would be accomplished in software and would entail predictable changes to Ferman’s system. Almeroth ¶416.

Byers provides explicit instructions on how to modify clips to insert advertising: Byers discloses underlying mechanisms to insert the advertising content. Byers at 3:67-4:5; 5:60-66. 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. Almeroth ¶417.

D. Claims 8 and 16: Ferman and Byers Render Obvious “One or More Video Clips Comprise Advertising”

Byers discloses an “advertiser database” that contains advertising-related information. Byers at 4:19-25. The information in the database is used to dynamically insert advertising into a video stream. *Id.* at 2:48-51; 3:1-4; Almeroth ¶¶418-19.

Moreover, Ferman’s system tracks advertising clips. Ferman’s system can remove advertising. Ferman [0103] (“commercial filter 92 to remove commercials”). A POSITA would have understood that in order to remove commercials, the system must be able to identify them first, and so some 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. Almeroth ¶420.

XVII. No Secondary Considerations

PO has not identified any secondary considerations sufficient to overcome Petitioner’s strong showing of obviousness. Almeroth ¶421.

XVIII. Instituting This IPR Would Be Equitable

A. 35 U.S.C. § 325(d) Supports Institution

The Board should not exercise its discretion under § 325(d). *Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6 at 7 (PTAB Feb. 13, 2020). The petition does not present the same or similar

grounds previously before the Office, and all six *Becton* factors support institution.

Becton, Dickinson & Co. v. B. Braun Melsungen AG, IPR2017-01586, Paper 8 at 24 (Dec. 15, 2017).

Becton factors (a) and (b): None of the references relied upon in this petition except for Marcus156 were considered by the examiner. Grounds 8 and 9 do not rely on Marcus156 and instead rely on Ferman and Byers. These references were not before the examiner, so any rationale for discretionary denial based on the previous citation of Marcus156 is not applicable to the full set of grounds. And grounds 1-7 combine Marcus156 with additional references not considered by the examiner. For example, to gain allowance, PO argued that none of the cited art discussed “a second area of the video display [that] present[s] one or more use-selectable frames from recently-displayed video clips.” Ex. 1005 at 232-42. In concluding that the claims are allowable, however, the examiner did not have the benefit of Sull or Rosen, which expressly disclose this claim limitation. Thus, the PTAB is presented with new, non-cumulative obviousness grounds that were never considered during prosecution of the ’185 patent.

Becton factors (c) and (d): Marcus156 was presented differently and for different purposes by the examiner during prosecution than here. When Marcus156 was cited by the examiner during the prosecution of the ’185 patent, it was only used as a secondary reference for some specific claim limitations. Ex. 1005 at 217-26.

In contrast, the Petition relies on Marcus156 as the base reference in grounds 1-7. Additionally, when Marcus156 was cited as a primary reference during the prosecution of other patents in the same family, the wording of the then-pending claims was significantly different from the challenged claims of the '185 patent. Ex. 1007 at 290-96, 320-39; Ex. 1008 at 51-57.

Prosecution of later family members of the '185 patent shows the critical relevance of Marcus156 to the alleged invention. In the subsequent '873 and '374 applications, the examiner rejected claims based on Marcus156 in combination with other prior art, and in both cases PO was unable to successfully distinguish Marcus156 and instead had to further amend the pending claims with limitations not found in the '185 patent. Ex. 1007 at 305-14, 360-67; Ex. 1008 at 78-81. Had the examiner considered Marcus156 as a base reference during prosecution of the '185 patent, the '185 patent claims would not have issued in their current form.

Becton factors (e) and (f): Petitioner raises new arguments and expert analysis concerning the disclosures of Marcus156 and its relationship to the challenged claims. For example, during prosecution, the examiner stated that Marcus156 does not “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 claim 13 in ground 1 above, Petitioner and its expert show how that is simply not the case, and how the examiner overlooked Marcus156’s disclosure of a

stream of video programming. Indeed, Marcus 156 expressly states that its video programming is “streamed.” Petitioner has overcome any presumption of discretionary denial by showing specifically how the examiner erred in evaluating Marcus156 during prosecution, and has submitted additional analysis in the petition and expert declaration that warrants reconsideration of the Marcus156. *Becton* at 24.

Put simply, if there could ever be a case where it would be appropriate to raise references in an IPR that were previously considered during prosecution, this is it. Discretionary denial under § 325(d) is unwarranted. *Pure Storage, Inc. v. Realtime Data LLC*, Case No. IPR2018-00549, Paper 7 at 11 (PTAB July 23, 2018).

B. *Fintiv* Supports Institution

District court litigation has not commenced in earnest. Petitioner is filing this IPR petition a month before its answer is due. The six *Fintiv* factors thus strongly favor institution.

1. Issue Overlap

While PO has not yet identified the set of asserted claims, this Petition challenges all ’185 patent claims. Accordingly, no matter what claims PO asserts in district court, all those claims will also be at issue here. On the other hand, Petitioner has not yet served its invalidity contentions. Thus, there is no basis for concluding that there is substantive overlap in invalidity theories.

2. Likelihood of Stay

Neither party has requested a stay in the district court in part because the district court case is not yet underway. The likelihood of a stay is increased here, however, by the fact that Petitioner is filing this IPR petition before answering the complaint and is filing concurrent IPR petitions against all asserted patents.

3. Proximity of Trial to FWD

No trial date has been set. No case schedule has been entered. Nor has a Rule 26(f) conference been held. At Judge Gilstrap's most-recent, July 2024 scheduling conferences, the Court assigned trial dates from January to April 2026. *Avant Location Techs., LLC v. Samsung Elec. Co. Ltd.*, No. 2:24-cv-00133-JRC, Dkt. 36 (E.D. Tex. Aug. 5, 2024); *Alto Dynamics, LLC v. Gucci Am., Inc.*, No. 2:24-cv-00375, Dkt. 29 (E.D. Tex. Aug. 2, 2024). This case is scheduled for a case management conference on September 18, 2024. *NTECH Properties, Inc. v. ByteDance Ltd.*, No. 2:24-cv-00130-JRG, Dkt. 13 (E.D. Tex. Aug. 8, 2024). By then, Judge Gilstrap will presumptively be assigning trial dates starting in May 2026. If that is the case, then the PTAB will issue its final written decision ***before*** any district court trial.

4. Investment in Parallel Proceeding

Although PO filed a district court complaint on February 23, 2024, investment in the parallel proceeding has been nonexistent. PO did not serve the complaint until

June 14, 2024. The parties then stipulated to an answer deadline of September 16, 2024. There have been no other case developments and no discovery. The district court has also invested zero time in assessing the parties' claims and defenses.

Moreover, Petitioner is filing this IPR a mere two months after being served with the complaint—almost 10 months before its statutory bar date (June 14, 2025).

5. Identity of Parties

The parties are the same in this IPR and in the Litigation.

6. Other Circumstances

This is the Board's opportunity to reassess a facially invalid patent that improperly attempts to extend the lifespan of PO's significantly earlier Marcus¹⁵⁶ patent. The merits of the Petition are particularly strong and outweigh any countervailing considerations of efficiency.

XIX. Conclusion

For the reasons above, *inter partes* review is requested.

Date: August 23, 2024

Respectfully submitted,

By: /s/ James M. Glass

James M. Glass

Counsel for Petitioner

CERTIFICATION UNDER 37 C.F.R. § 42.24

Under the provisions of 37 C.F.R. § 42.24, the undersigned hereby certifies that the word count for the foregoing Petition for *inter partes* review (excluding the table of contents, table of authorities, mandatory notices, certificate of service or word count, and appendix of exhibits or claim listing) totals 13,891 words, which is within the word limit allowed under 37 C.F.R. § 42.24(a)(i).

Date: August 23, 2024

/s/ James Glass

James Glass (Reg. No. 46729)

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. §§ 42.6(e), 42.105(a), the undersigned hereby certifies service on the PO of a copy of this Petition and its respective exhibits at the official correspondence address for the attorneys of record for the '185 patent as shown in USPTO PAIR via FedEx:

Louis DelJuidice
Robert Schaffer
Troutman Pepper Hamilton Sanders LLP
600 Peachtree St. NE, Suite 3000
Atlanta, GA

Courtesy copies were also sent via electronic mail to Patent Owner's counsel of record in the related district court proceeding:

Elizabeth L. DeRieux
ederieux@capshawlaw.com
Capshaw DeRieux LLP
114 E. Commerce St.
Gladewater, TX 75647

Paul J. Andre
pandrew@kramerlevin.com
Lisa Kobialka
lkobialka@kramerlevin.com
James Hannah
jhannah@kramerlevin.com
Kris Kastens
kkastens@kramerlevin.com
Kramer Levin Naftalis & Frankel LLP
333 Twin Dolphin Drive
Redwood City, CA 94065

Petition for *Inter Partes* Review

U.S. Patent No. 8,875,185

Date: August 23, 2024

/s/ *James Glass*

James Glass (Reg. No. 46729)