

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC,

Petitioner,

v.

SONOS, INC,

Patent Owner.

Case No. IPR2026-00020
U.S. Patent No. 11,080,001

DECLARATION OF DR. DAVID ANDERSON

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	SUMMARY OF OPINIONS	1
III.	QUALIFICATIONS AND BACKGROUND	2
IV.	MATERIALS CONSIDERED	6
V.	LEGAL STANDARDS.....	8
A.	Level of Ordinary Skill.....	9
B.	Prior Art.....	10
C.	Anticipation	10
D.	Obviousness.....	12
VI.	THE '001 PATENT	17
VII.	CLAIM CONSTRUCTION.....	19
VIII.	LEVEL OF ORDINARY SKILL	21
IX.	OVERVIEW OF PRIOR ART	22
X.	Ground 1: Claims 1-2, 6-13, 17-24, 28-33 Are Obvious Over Janevski, Kawamura, and Okamura.....	23
A.	Janevski (Ex. 1005).....	23
B.	Kawamura (Ex. 1006).....	27
C.	Okamura (Ex. 1007)	29
D.	Independent Claims	30
1.	Claim 1	31
a.	[1preamble]: “A method performed by a first zone player, the method comprising:”	31

b.	[1a]: receiving, via a network interface at the first zone player, a request to engage in synchronous playback of audio content as part of a synchrony group that includes at least a second zone player that is communicatively coupled to the first zone player via at least one data network;	31
c.	[1b] “after receiving the request to enter into the synchrony group:”	41
d.	[1ci] “detecting an indication that the first zone player is to operate in (a) one of a control-master mode or a control-slave mode for the synchrony group”	41
e.	[1cii] “and (b) one of an audio-master mode or an audio-slave mode for the synchrony group;”	43
f.	[1d] “beginning to operate in the synchrony group in accordance with the indication;”	45
g.	[1e] “wherein, while operating in the control-master mode for the synchrony group, the first zone player is configured to: receive, via the network interface, first control information for the synchrony group from a network device that is communicatively coupled to the first zone player; and”	46
h.	[1f] “based on the first control information, cause, via the network interface, at least one playback action to be applied in the synchrony group;”	50
i.	[1g] “wherein, while operating in the control-slave mode for the synchrony group, the first zone player is configured to: receive, via the network interface, second control information from another zone player; and”	52
j.	[1h] “perform one or more playback actions in accordance with the second control information;”	53

k.	[1i] “wherein, while operating in the audio-master mode for the synchrony group, the first zone player is configured to: obtain audio information that is representative of the audio content;”	53
l.	[1j] “generate playback timing information associated with the obtained audio information that is indicative of at least one future time relative to a reference clock time that denotes a time at which at least the first and second zone players are to engage in synchronous playback of a corresponding portion of the obtained audio information; and”	56
m.	[1k] “transmit, via the network interface, the obtained audio information and the generated playback timing information to the second zone player; and”	65
n.	[1l] “wherein, while operating in the audio-slave mode for the synchrony group, the first zone player is configured to: receive, via the network interface, audio information and playback timing information associated with the received audio information from another zone player; and”	67
o.	[1m] “engage in synchronous playback of the received audio information with at least the second zone player based on the received playback timing information associated with the received audio information while a local clock time of the first zone player differs from a local clock time of the second zone player.”	71
2.	Claims 12 and 23	76
E.	Dependent Claims.....	80
1.	Claims 2, 13, and 24: “wherein detecting an indication that the first zone player is to operate in (a) one of a control-master mode or a control-slave mode for the synchrony group and (b) one of an audio-master mode or	

an audio-slave mode for the synchrony group comprises detecting an indication that the first zone player is to operate in (a) the control-master mode for the synchrony group and (b) the audio-master mode for the synchrony group.”80

2. Claims 6, 17, and 28: “wherein beginning to operate in the synchrony group in accordance with the indication comprises either (a) transitioning from operating in the audio-master mode to operating in the audio-slave mode or (b) transitioning from operating in the audio-slave mode to operating in the audio-master mode.”81

3. Claims 7, 18 and 29: “wherein the first control information identifies particular audio content to be played back by the synchrony group that is available at an audio source outside of the at least one data network, and wherein causing the at least one playback action to be applied in the synchrony group comprises causing a zone player operating in the audio-master mode to obtain audio information that is representative of the particular audio content.”83

4. Claims 8, 19, and 30: “wherein the at least one future time relative to the reference clock time comprise at least one first future time that is determined based on a local clock of a zone player other than the first zone player.”85

5. Claims 9, 20 and 31: “wherein the at least one future time relative to the reference clock time comprise at least one first future time that is determined based on a local clock of the first zone player.”86

6. Claims 10, 21, and 32: “wherein the second control information comprises information indicative of a volume adjustment, and wherein performing one or more playback actions in accordance with the second control information comprises adjusting a playback volume of the first zone player.”88

7. Claims 11, 22 and 3391

a.	11, 22, and 33 (pre): “wherein the playback timing information that is received while operating in the audio-slave mode comprises at least one future time at which at least the first and second zone players are to engage in synchronous playback of a corresponding portion of the received audio information, and wherein operating in the audio-slave mode to engage in synchronous playback of the received audio information with at least the second zone player comprises:”	91
b.	11, 22, and 33 (a): “updating the at least one future time to account for a differential between the local clock time of the first zone player and a local clock time of another zone player; and”	91
c.	11, 22, and 33 (b): “when the local clock time of the first zone player reaches the updated at least one future time, engaging in synchronous playback of the corresponding portion of the received audio information with at least the second zone player.”	92
XI.	Ground 2: Claims 3, 14, and 25 Are Rendered Obvious by Janevski, Kawamura, Okamura, and Kono.....	93
A.	Kono (Ex. 1008)	93
B.	Claims 3, 14, and 25: “wherein the obtained audio information comprises a beginning of the obtained audio information, and wherein the playback timing further comprises a future time relative to the reference clock time that denotes a time at which at least the first and second zone players are to initiate synchronous playback at the beginning of the obtained audio information.”	96
XII.	CONCLUSION.....	102

I. INTRODUCTION

1. I have been retained by Google LLC (“Petitioner”) as an independent expert consultant in this proceeding before the United States Patent and Trademark Office.

2. I am over 21 years of age and, if I am called upon to do so, I would be competent to testify as to the matters set forth herein.

3. I have been asked by Petitioner to provide my opinion on the validity of certain claims of the U.S. Patent No. 11,080,001 from the perspective of a person of ordinary skill in the art as of the priority date of the ’001 patent. In performing my analysis, I have been asked to assume that the priority date of the ’001 patent is April 1, 2004. My opinions are set forth below.

4. Throughout this declaration, I refer to specific paragraphs, pages, figures, and/or line numbers of various exhibits. These citations are illustrative and are not intended to suggest that they are the only support for the propositions for which they are cited.

II. SUMMARY OF OPINIONS

5. In preparing this declaration, I have considered claims 1-3, 6-14, 17-25, 28-33 of the ’001 patent. Below, I set forth the opinions I have formed, the conclusions I have reached, and the bases for these opinions and conclusions. The

statements contained in this declaration are true and correct to the best of my knowledge.

6. Based on my experience and knowledge of the art at the time of April 1, 2004, it is my opinion that claims 1-3, 6-14, 17-25, 28-33 of the '001 patent would have been obvious based on the asserted grounds that I discuss below.

Ground	Claim(s) Challenged	35 U.S.C. §	Reference(s)
1	1-2, 6-13, 17-24, 28-33	103	Janevski, Kawamura, and Okamura
2	3, 14, 25	103	Janevski, Kawamura, Okamura, and Kono

7. I am being compensated for my time at my normal hourly rate of \$600. My compensation is in no way contingent upon the nature of my findings, the presentation of my findings in testimony, or the outcome of this proceeding.

III. QUALIFICATIONS AND BACKGROUND

8. I am well qualified to serve as a technical expert in this matter based upon my qualifications, discussed in detail below.

9. I am competent to write this declaration.

10. I am an independent consultant. All of my opinions stated in this declaration are based on my own personal knowledge and professional judgment. In forming my opinions, I have relied on my knowledge and experience in the fields of

networking and network-based systems, such as consumer audio systems, and on the documents and information referenced in this declaration. I am over 18 years of age and, if I am called upon to do so, I am competent to testify as to the matters set forth herein.

11. In formulating my opinions, I have relied upon my training, knowledge, and experience in the relevant art. Provided as Exhibit 1004 is a copy of my current curriculum vitae, which details my education and experience, and a list of all other cases in which, during the previous four years, I testified as an expert at trial or by deposition. The following provides a brief overview of some of my experience that is relevant to the matters set forth in this declaration.

12. I am a professor in the School of Electrical and Computer Engineering at the Georgia Institute of Technology (“Georgia Tech”) in Atlanta, Georgia. I have been a professor at Georgia Tech since 1999. In 2009 I served as a visiting professor in the Department of Computer Science at Korea University in Seoul, South Korea.

13. I received my Ph.D. in Electrical and Computer Engineering from Georgia Tech in 1999. I received my B.S. and M.S. in Electrical Engineering from Brigham Young University in 1993 and 1994, respectively.

14. In my employment prior to Georgia Tech as well as in my subsequent studies and research, I have worked extensively in areas related to the research,

design, and implementation of speech and audio processing systems. In 1983-1987 and again in 1989-1991 I worked as an embedded systems developer for industrial control equipment and my responsibilities included developing low-level code and protocols for communicating between interconnected systems. Prior to my employment at Georgia Tech, I also developed prototype real-time audio systems for Sonic Innovations in 1993-1994, and various real-time audio processing systems for consulting clients from 1995-1999.

15. As a professor at Georgia Tech, I have taught graduate and undergraduate level courses at Georgia Tech on the implementation of signal processing and embedded systems. For example, I have taught courses on multimedia processing and systems (covering topics including standards and methods for multi-channel and networked multimedia), software design, computer architecture and operating systems, real-time signal processing systems, applications of signal processing (covering topics in audio processing and communications systems), speech recognition systems, statistical machine learning, machine learning for speech, pattern recognition, and adaptive signal processing systems. I also designed and teach a course on signal processing in the context of human perception which covers audio standards and audio system design considerations. These courses and

my research have covered many topics relevant to the subject matter of the '001 patent and the prior art cited therein.

16. I have served as principal investigator or co-principal investigator in numerous multi-disciplinary research projects including “Blind Source Separation for Audio,” “Audio Classification,” “Auditory Scene Analysis,” “Hearing Aid Audio Processing,” “Speaker Driver Sound Enhancement,” “I-Vector Based Voice Quality,” “Analysis of Voice Exercise Using Signal Processing,” and “Smart Homes for Effective and Safe Remote Work During a Pandemic and Beyond.”

17. I also have extensive experience with the practical implementation of signal processing algorithms, information theory, signal detection, and related topics through my research and consulting. I have published over 200 papers in reviewed journals and conferences and books and book chapters.

18. I am a named inventor on eight patents, including “Speech activity detector for use in noise reduction system, and methods therefor” (U.S. Patent No. 6,351,731), and “Analog audio signal enhancement system using a noise suppression algorithm” (U.S. Patent No. 7,590,250).

19. I was also a co-founder of Ratrix, LLC, a company that developed efficient algorithms for reliable next-generation wireless communications systems.

20. I am a Senior Member of the Institute of Electrical and Electronics Engineers (“IEEE”) and have been a Member since 1991. I am also a Member of the IEEE Signal Processing Society. From 1994 to 2016, I was also a member of the Acoustical Society of America. In 2003, I served as the Co-Chair for the NSF Symposium on Next Generation Automatic Speech Recognition. In 2004, I received the Presidential Early Career Award for Scientists and Engineers, presented by then-President George W. Bush, for my work on ultra-low-power signal processing system design.

IV. MATERIALS CONSIDERED

21. In forming my opinions, I have reviewed the following documents and any other document cited in this declaration.

Exhibit	Description
Ex. 1001	U.S. Patent No. 11,080,001 (“the ’001 patent”)
Ex. 1002	File History of U.S. Patent No. 11,080,001
Ex. 1005	U.S. Patent No. 7,269,338 to Janevski (“Janevski”)
Ex. 1006	U.S. Patent Publication No. 2004/0048569 (“Kawamura”)
Ex. 1007	U.S. Patent No. 6,751,228 to Okamura (“Okamura”)
Ex. 1008	JP2003323186A to Kono (“Kono”) (with certified translation)
Ex. 1009	U.S. Patent No. 5,010,399 to Goodman et al. (“Goodman”)

Exhibit	Description
Ex. 1012	Sonos's Corrected Disclosure of Asserted Claims and Infringement Contentions in <i>Sonos, Inc. v. Google LLC</i> , No. 2:20-cv-00169 (C.D. Cal.)
Ex. 1013	U.S. Patent No. 7,643,894 to Braithwaite et al. (“Braithwaite”)
Ex. 1014	Gary C. Kessler, <i>Overview of TCP/IP and the Internet</i> , 1994, available at https://garykessler.net/library/tcpip.html (“Kessler”)
Ex. 1015	Agreed Claim Constructions in <i>Sonos, Inc. v. Google LLC</i> , No. 2:20-cv-00169 (C.D. Cal.)
Ex. 1017	U.S. Patent No. 5,369, 724 to Lim (“Lim”)

22. I have also relied on my education, experience, research, training, and knowledge in the relevant art, and my understanding of any applicable legal principles described in this declaration.

23. All the opinions contained in this declaration are based on the documents I reviewed and my knowledge and professional judgment. My opinions have also been guided by my understanding of how a person of ordinary skill in the art would have understood the claims of the '001 patent.

24. I reserve the right to supplement and amend any of my opinions in this declaration based on documents, testimony, and other information that becomes available to me after the date of this declaration.

V. LEGAL STANDARDS

25. I am not a lawyer. My understanding of the legal standards to apply in reaching the conclusions in this declaration is based on discussions with counsel for Petitioner, my experience applying similar standards in other patent-related matters, and my reading of the documents submitted in this proceeding. In preparing this declaration, I have tried to faithfully apply these legal standards to the challenged claims.

26. I have been informed that there are two ways in which prior art may render a patent claim unpatentable. First, I have been informed that the prior art can “anticipate” a claim. Second, I have been informed that the prior art can render a claim “obvious” to a person of ordinary skill in the art. I understand that a claim is patentable if it was not anticipated and would not have been obvious in view of the prior art as of the effective filing date of the patent.

27. I have been informed that a dependent claim is a patent claim that refers back to another patent claim. I have been informed that a dependent claim includes all the limitations of the claim to which it refers.

28. I have been asked to provide my opinion on the validity of claims 1-3, 6-14, 17-25, 28-33 of the '001 patent from the perspective of a person of ordinary

skill in the art at the '001 patent's priority date, which I have been informed to assume is April 1, 2004, as described in more detail below.

29. I have been informed that in *inter partes* review proceedings, such as this one, the party challenging the patent bears the burden of proving unpatentability by a preponderance of the evidence. I understand that a preponderance of the evidence means "more likely than not."

30. For the purposes of this declaration, I have been asked to provide my opinions on issues regarding unpatentability. I have been informed of the following legal standards, which I have applied in forming my opinions.

A. Level of Ordinary Skill

31. I have been informed that a person of ordinary skill in the art is determined by considering several factors, including the (i) type of problems encountered in the art; (ii) prior art solutions to those problems; (iii) rapidity with which innovations are made; (iv) sophistication of the technology; and (v) educational level of active workers in the field.

32. I have been instructed to assume that a person of ordinary skill in the art is not a specific real individual but rather a hypothetical individual having the qualities reflected by the factors discussed above.

B. Prior Art

33. I have been advised and understand that the information used to evaluate whether an invention was novel and not obvious when made is generally referred to as “prior art.” I understand that prior art includes patents and printed publications that existed before the priority date of the patent. I have been informed and understand that a patent or published patent application is prior art if it was filed before the priority date of the claimed invention and that a printed publication is prior art if it was publicly available before the priority date. I have been informed to assume that the priority date of the '001 patent is April 1, 2004.

C. Anticipation

34. I have been informed that under 35 U.S.C. § 102, a patent claim is unpatentable for anticipation if the claimed subject matter was patented or described in a printed publication before the effective filing date of the claimed invention. I have been informed that this is referred to as unpatentability by anticipation. I have been informed that a patent claim is anticipated under § 102 if a single prior art reference discloses all limitations of the claimed invention.

35. I understand that a patent claim is anticipated only if each and every element as set forth in the claim is described in a single prior art reference. I also understand that, to be considered anticipatory, the prior art reference does not need

to use the same words as the claims or provide as many details as the patent's specification; rather, it must only be enabling and describe the claimed invention sufficiently to have placed it in possession of a person of ordinary skill in the field of the invention. I further understand that to be an anticipatory reference, a prior art reference must disclose all elements of the claim arranged or combined in the same way stated in the claim.

36. I also understand that in order for a patent claim to be anticipated, each and every element set forth in the claim must be found, either expressly or inherently, in a single prior art reference, such that a person of ordinary skill in the art could practice the invention without undue experimentation. A claim limitation that is not expressly found in a prior art reference is 'inherent' if the prior art necessarily functions in accordance with, or includes, the limitation. It is acceptable to consider evidence outside the prior art reference to determine whether a limitation is necessarily present.

37. I understand that a dependent claim incorporates by reference all elements of the claim from which it depends. For this reason, I understand that if a reference fails to disclose one or more elements of an independent claim, it cannot anticipate any claims that depend from the independent claim.

D. Obviousness

38. I have been informed that for obviousness under 35 U.S.C. § 103, a patent claim is unpatentable if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious to a person having ordinary skill in the art to which said subject matter pertains at the time the invention was made. I have been informed that this is referred to as unpatentability by obviousness.

39. I have been informed that an obviousness analysis includes the following considerations:

- a. Determining the scope and content of the prior art;
- b. Ascertaining the differences between the prior art and the claims at issue;
- c. Resolving the level of ordinary skill in the pertinent art; and
- d. Considering evidence of secondary indicia of nonobviousness (if available).

40. I have been informed that the relevant time for considering whether a claim would have been obvious to a person of ordinary skill in the art is the time of invention. For my obviousness analysis, counsel for Petitioner instructed me to

assume that the date of invention for the challenged claims is April 1, 2004. My opinions would not change if I assumed a later date of invention.

41. I have been informed that a reference may be modified or combined with other references or with a person of ordinary skill in the art's own knowledge if the person would have found the modification or combination obvious. I have also been informed that a person of ordinary skill in the art is presumed to know all the relevant prior art, and the obviousness analysis may take into account the inferences and creative steps that a person of ordinary skill in the art would employ.

42. I have been informed that whether a prior art reference renders a patent claim obvious is determined from the perspective of a person of ordinary skill in the art. I have also been informed that, while there is no requirement that the prior art contain an express suggestion to combine known elements to achieve the claimed invention, and while a suggestion to combine known elements to achieve the claimed invention may come from the prior art as a whole or individually and may consider the inferences and creative steps a person of ordinary skill in the art would employ, as filtered through the knowledge of one skilled in the art, obviousness grounds cannot be sustained by mere conclusory statements and must include some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness.

43. In determining whether a prior art reference could have been combined with another prior art reference or other information known to a person having ordinary skill in the art, I have been informed that the following principles may be considered:

- a. A combination of familiar elements according to known methods is likely to be obvious if it yields predictable results;
- b. The substitution of one known element for another is likely to be obvious if it yields predictable results;
- c. The use of a known technique to improve similar items or methods in the same way is likely to be obvious if it yields predictable results;
- d. The application of a known technique to a prior art reference that is ready for improvement to yield predictable results;
- e. Any need or problem known in the field and addressed by the reference can provide a reason for combining the elements in the manner claimed;
- f. A person of ordinary skill often will be able to fit the teachings of multiple references together like a puzzle; and
- g. The proper analysis of obviousness requires a determination of whether a person of ordinary skill in the art would have a “reasonable

expectation of success”—but not “absolute predictability” of success—in achieving the claimed invention by combining prior art references.

44. I have been informed that, when a work is available in one field, design alternatives and other market forces can prompt variations of it, either in the same field or in another. I have been informed that if a person of ordinary skill in the art could have implemented a predictable variation and would have seen the benefit of doing so, that variation is likely to have been obvious. I have been informed that, in many fields, there may be little discussion of obvious combinations, and in these fields, market demand—not scientific literature—may drive design trends. I have been informed that, when there was a design need or market pressure and there are a finite number of predictable solutions, a person of ordinary skill in the art would have had a good reason to pursue those known options.

45. I have been informed that the law permits the application of “common sense” in examining whether a claimed invention would have been obvious to a person skilled in the art. For example, I have been informed that combining familiar elements according to known methods and in a predictable way may suggest obviousness when such a combination would yield nothing more than predictable results. I understand, however, that a claim is not obvious merely because every claim element is disclosed in the prior art, and that a party asserting obviousness

must still provide a specific motivation to combine or modify the references as recited in the claims and explain why one skilled in the art would have reasonably expected to succeed in doing so.

46. I have been informed that there is no rigid rule that a reference or combination of references must contain a “teaching, suggestion, or motivation” to combine references. But I also understand that the “teaching, suggestion, or motivation” test can be a useful guide in establishing a rationale for combining elements of the prior art. I have been informed that this test poses the question as to whether there is an express or implied teaching, suggestion, or motivation to combine prior art elements in a way that realizes the claimed invention, and that it seeks to counter impermissible hindsight analysis.

47. I am not aware of any evidence of secondary considerations that would support a determination of nonobviousness of the claimed subject matter in the '001 patent.

48. I have been informed that, in an obviousness analysis, prior art must be analogous to the patent being considered. I have been informed that a prior art reference is considered to be analogous, or in the same field of art, if the reference is either (1) in the same field of endeavor as the challenged patent, regardless of the

problems the challenged patent and the prior art address; or (2) reasonably pertinent to the particular problem being solved by the challenged patent.

VI. THE '001 PATENT

49. The '001 patent relates to “synchronizing operations among a plurality of independently-clocked digital data processing devices.” Ex. 1001, 1:25-29. Those devices form a “synchrony group 20 [that] includes a master device 21 and zero or more slave devices 22(1) through 22(g) ... all of which synchronously play an audio program.” Ex. 1001, 8:2-7. “The master device 21 ... controls the operations of the slave devices 22(g) in the synchrony group 20.” Ex. 1001, 8:27-35.

50. “The master device 21 receives control information from the user interface module 13 for controlling the synchrony group.” Ex. 1001, 9:5-9. The control information “can enable the synchrony group 20 to terminate playing of a current track of the audio program and skip to the next track, and to re-order tracks in a play list of tracks defining the audio program that is to be played by the synchrony group 20.” Ex. 1001, 9:31-35.

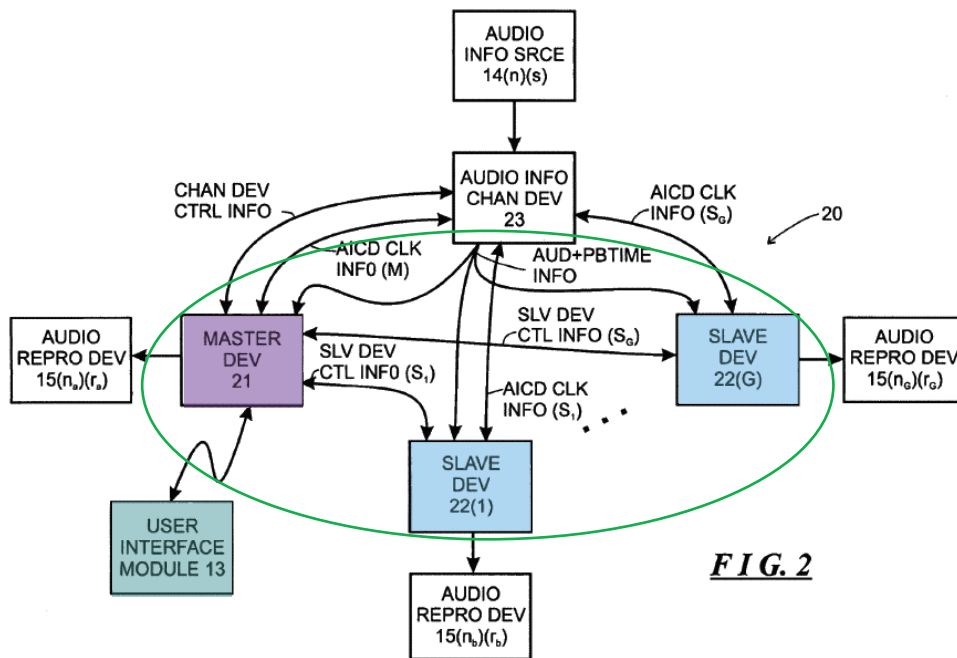


FIG. 2

Ex. 1001, FIG. 2.

51. “Each of the master device 21, slave devices 22(g) and audio information channel device 23 utilizes a zone player 11(n) depicted in FIG. 1.” Ex. 1001, 8:7-13.

52. The '001 patent describes “zone players” as devices that “receive[] audio information from an audio information source 14(n)(s) that is connected thereto” and “provide[] playback and/or forward the audio information, along with playback timing information, over the local network 12 to other zone players for playback.” Ex. 1001, 5:13–18. A zone player may also “transmit the audio information that it receives ... to selected ones of the other zone players ... for

playback,” and those **zone players** “synchronize their playback ... so that the **zone players** ... provide the same audio program at the same time.” Ex. 1001, 5:21–39.

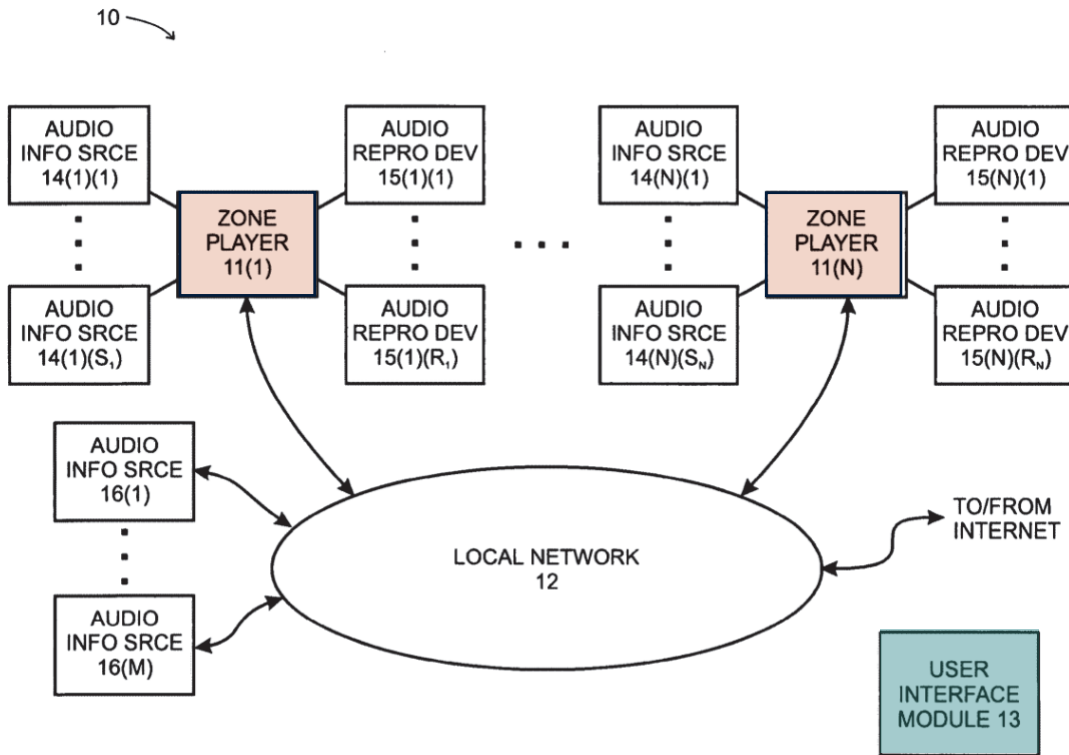


FIG. 1

Ex. 1001, FIG. 1.

VII. CLAIM CONSTRUCTION

53. I have been instructed that the “intrinsic record” includes the ’001 patent itself, including the claims, description, and figures, and the ’001 patent’s prosecution history, which I understand is the record of proceedings at the U.S. Patent and Trademark Office (“Patent Office”) concerning the patent. I understand that, like the claims and written description, the prosecution history provides

evidence to a person of ordinary skill in the art of how the inventor intended their patent to be understood and how the Patent Office understood the patent. I understand that the inventor is permitted to apply a special definition to the terms or to limit the scope of claim terms in their patent claims, which may differ from the term's plain and ordinary meaning. That special definition or limitation on scope may be provided in the patent's written description, the patent's prosecution history, or both.

54. I have been informed that the parties have agreed to the following claim constructions. Ex. 1015.

“zone player”	“data network device configured to process and output audio”
“network interface”	“physical component of a device that provides an interconnection with a data network”
“playback timing information”	“information indicating when the audio [information/content] is to be played back”

“synchrony group”	“a set of two or more zone players that are to play the same audio program synchronously”
-------------------	-------------------------------------------------------------------------------------------

55. I have applied these agreed constructions and the plain and ordinary meaning of all other the claim terms as they would have been understood by a person of ordinary skill in the art at the priority date of the '001 patent. It is my opinion that a person of ordinary skill in the art would have found the claims obvious over the prior art references I have considered, applying the agreed constructions and the plain and ordinary meaning of the claim terms.

VIII. LEVEL OF ORDINARY SKILL

56. In rendering the opinions set forth in this declaration, I have been asked to consider the '001 patent's claims and the prior art through the eyes of a person of ordinary skill in the art (which I may also refer to as “one skilled in the art,” “skilled artisan,” “POSITA,” or similar variations). I have considered factors such as the educational level and years of experience of those working in the pertinent art, the types of problems encountered in the art, the teachings of the prior art, patents and publications of other persons or companies, and the sophistication of the technology.

57. Taking the factors described in the Legal Standards – Level of Ordinary Skill section into consideration, it is my opinion that a person of ordinary skill in the

art (“POSITA”) at the time of the purported invention would have had the equivalent of a four-year degree from an accredited institution in computer science, computer engineering, electrical engineering, or an equivalent thereof, and approximately four years of professional experience in the fields of networking and network-based systems or applications, such as consumer audio systems, or an equivalent level of skill, knowledge, and experience. Additional education could substitute for professional experience, and significant work experience could substitute for formal education.

58. I met the criteria for a POSITA at the priority date of the ’001 patent of April 1, 2004.

IX. OVERVIEW OF PRIOR ART

59. In forming my opinions, I found the following prior art relevant when considering the validity of certain claims of the ’001 patent from the perspective of a person of ordinary skill in the art.

Exhibit	Description
Ex. 1005	U.S. Patent No. 7,269,338 to Janevski (“Janevski”)
Ex. 1006	U.S. Patent Publication No. 2004/0048569 (“Kawamura”)
Ex. 1007	U.S. Patent No. 6,751,228 to Okamura (“Okamura”)
Ex. 1008	JP2003323186A to Kono (“Kono”) (with certified translation)
Ex. 1009	U.S. Patent No. 5,010,399 to Goodman et al. (“Goodman”)

Exhibit	Description
Ex. 1013	U.S. Patent No. 7,643,894 to Braithwaite et al. (“Braithwaite”)
Ex. 1014	Gary C. Kessler, <i>Overview of TCP/IP and the Internet</i> , 1994, available at https://garykessler.net/library/tcpip.html (“Kessler”)
Ex. 1017	U.S. Patent No. 5,369,724 to Lim (“Lim”)

60. As explained below, in my opinion, claims 1-3, 6-14, 17-25, 28-33 of the '001 patent would have been obvious based on the grounds that I discuss below.

X. Ground 1: Claims 1-2, 6-13, 17-24, 28-33 Are Obvious Over Janevski, Kawamura, and Okamura.

A. Janevski (Ex. 1005)

61. Janevski relates “synchronizing playback of two or more digital streams based on renderable content of those streams.” Janevski, 1:5-11. Janevski discloses “a **synchronized viewing session**” including “**PVRs 114a, b**.” Janevski, 6:4-18. “[B]roadcasts 112a, b of a television program ... are received by receivers 113a, b housed within each of the respective **PVRs 114a, b**.” *Id.* “The presentation may be ... aural.” Janevski, 16:34-43. And “**remote control 116a, b** is commonly associated with the **personal video recorder 114a, b** to allow the user to operate the **personal video recorder 114a, b** remotely.” Janevski, 6:39-44.

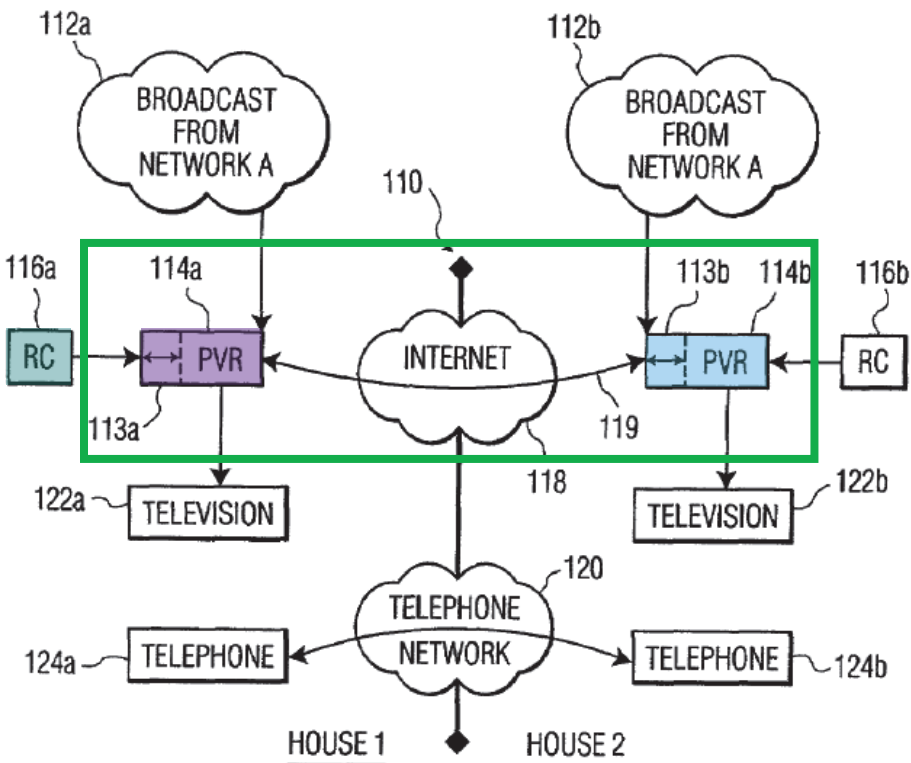


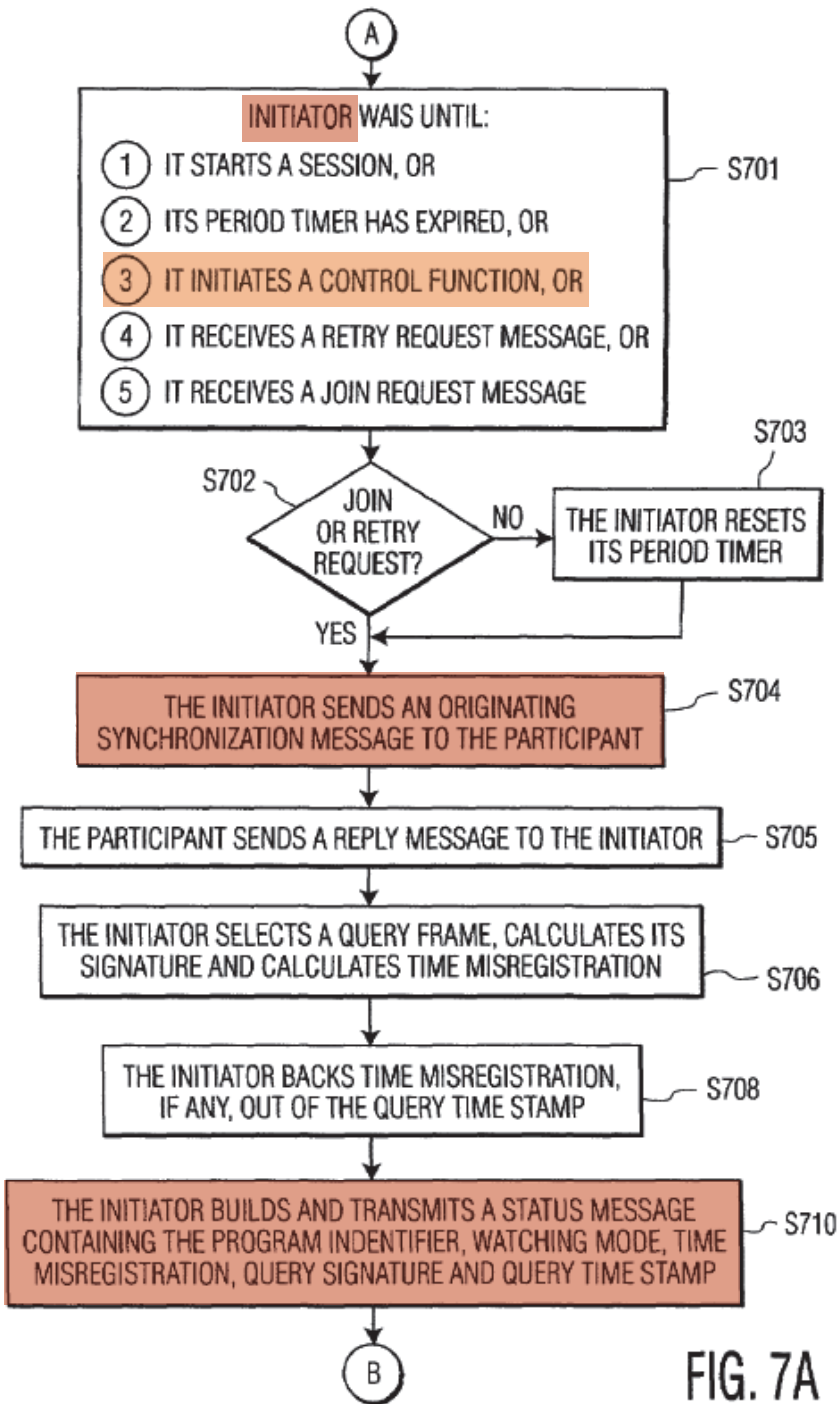
FIG. 1

Janevski, FIG. 1.

62. A session user “performs a **control function** (e.g., **start, pause, rewind, fast forward**).” Janevski, 7:31-36, Abstract. When a user “executes a **control function**,” it “changes the watching mode.” Janevski, 8:48-52.

63. “The **initiator** [PVR] directs all **participants** ... upon execution of each control function, and periodically, **to synchronize their playbacks to that of the initiator**, whereby all playbacks are synchronized and maintain[ed] in synchronization.” Janevski, Abstract. “To ensure that the **PVRs 114a, b** participating

in the session remain synchronous, a status message is sent out periodically by the ‘initiator,’ i.e., the PVR 114a.” Janevski, 7:36-50. When “the role of initiator is handed off,” Janevski explains that “another participant 114b assumes the role of initiator PVR 114a.” Janevski, 6:16-25, 7:51-57, 15:32-42.



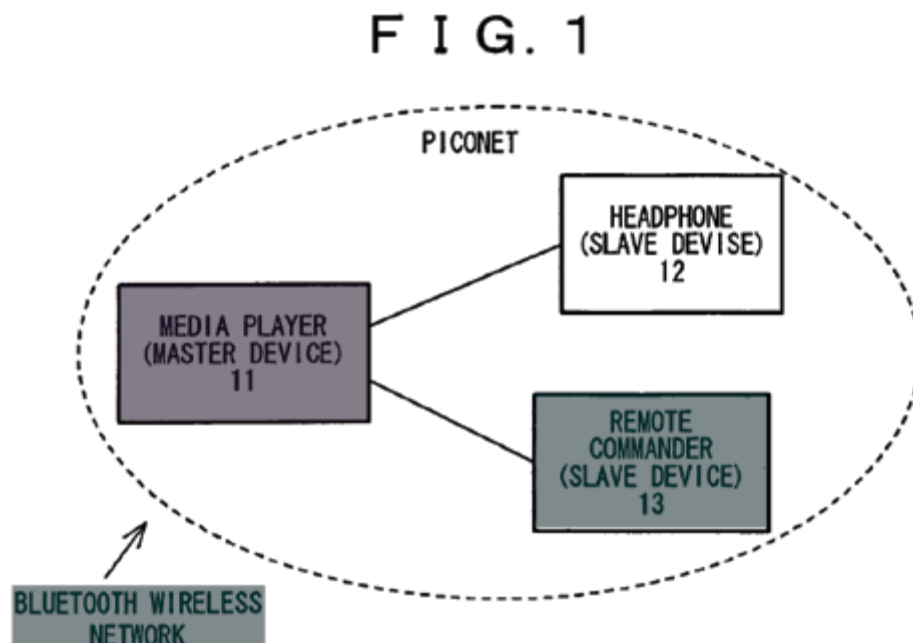
Janevski, FIG. 7A.

64. In my opinion, like the '001 patent, Janevski is in the technical field of controlling synchronous audio playback. *Compare* Ex. 1001, 1:25-45, 2:35-48, 5:33-39, *with* Janevski, 1:5-11, 5:3-43.

B. Kawamura (Ex. 1006)

65. Kawamura “relates to exchanging device operating commands between [the devices within a piconet ... implemented by the Bluetooth™ standard.” Kawamura, [0002].

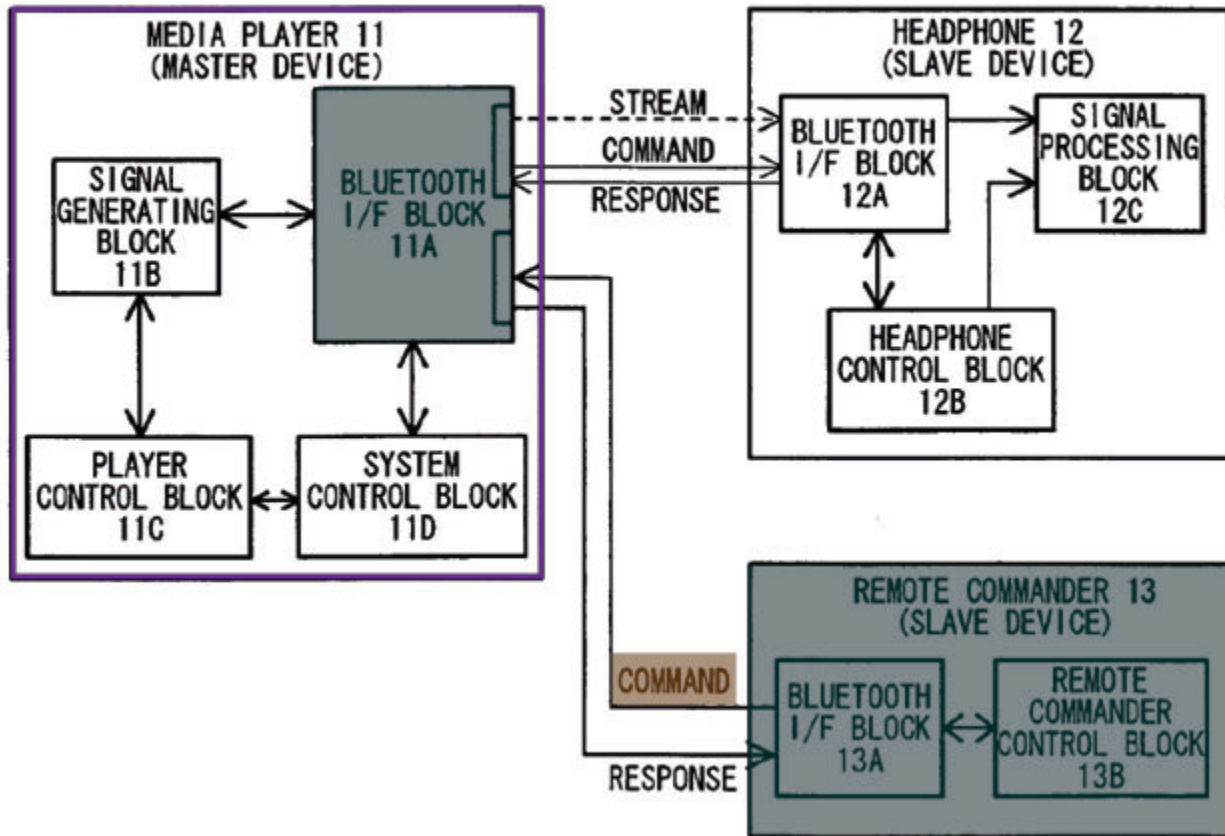
66. Kawamura discloses a “piconet 10 contain[ing] a media player 11 ... and a remote commander 13,” as illustrated in Figure 1. Kawamura, [0088]. “The remote commander 13 remotely controls the media player 11 for executing play control such as play, stop, pause, fast forward and rewind.” *Id.*



Kawamura, FIG. 1.

67. Kawamura's "media player" includes "a Bluetooth interface" which "transfers commands and responses" with the remote commander 13. Kawamura, [0090]-[0091].

F I G . 2

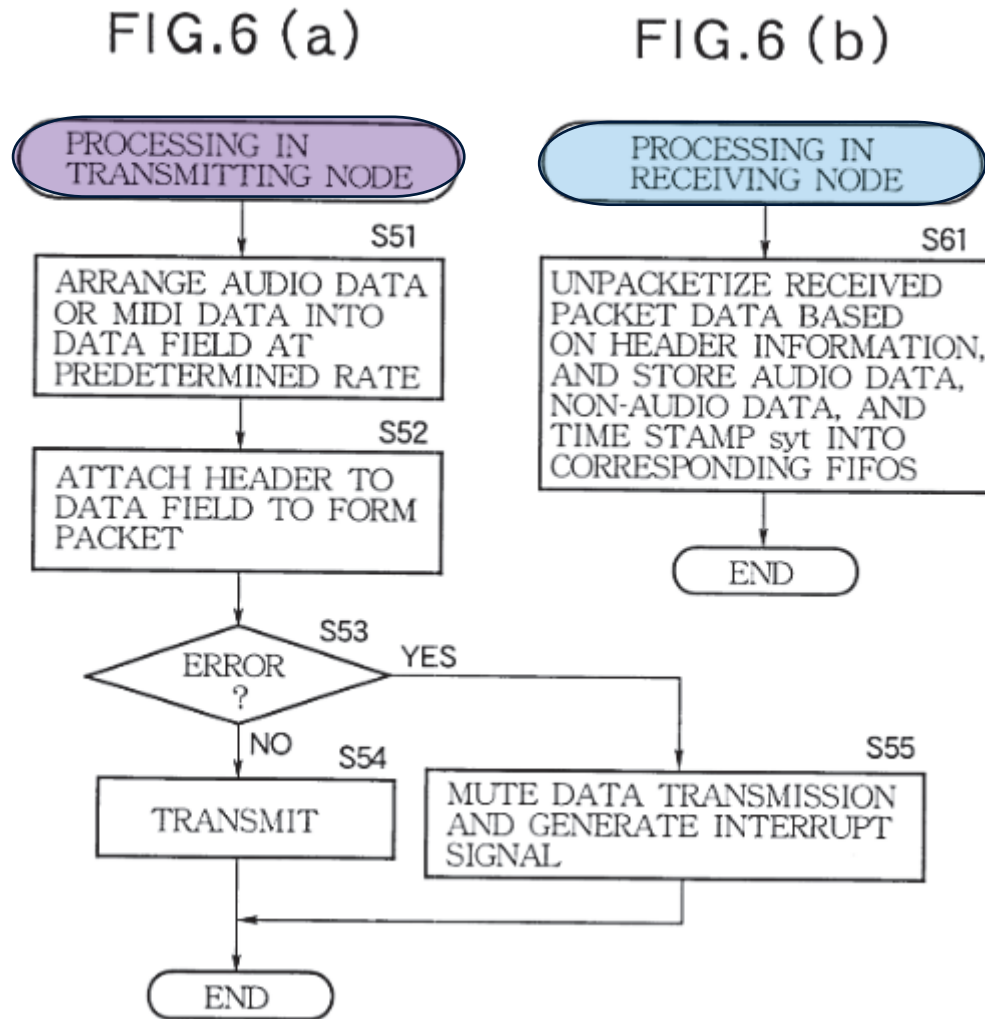


Kawamura, FIG. 2.

68. In my opinion, like the '001 patent, Kawamura is in the technical field of controlling audio playback. Compare Ex. 1001, 9:5-36, claims 1, 7, 13, with Kawamura, [0088], [0206].

C. Okamura (Ex. 1007)

69. Okamura relates to synchronizing audio data reproduction between a transmitting node and a receiving node. Okamura, Abstract. In Okamura, the **transmitting node** functions as the master, generating a timestamp “on the basis of the time provided by the cycle timer,” which “specifies the reproduction time at the receiving side.” Okamura, 1:42–56. The **receiving node**, acting as the slave, uses “an offset setting section that sets an offset time ... and adds the offset time to a time indicated by the timestamp,” and initiates playback “when the time of the timestamp added with the offset time coincides with a current time indicated by an internal cycle timer.” Okamura, 5:53–67. Okamura also discloses “estimating propagation delay” on the transmitting side to set the timestamp as a future reproduction time, and “adjusting the offset on the receiving side” to shift playback timing. Okamura, 24:7–12.



Okamura, FIGS. 6(a) and 6(b).

70. In my opinion, like the '001 patent, Okamura is in the technical field of adjusting playback timing. *Compare* Ex. 1001, 24:45-25:57 with Okamura, 1:42-56, 2:49-67, 3:1-19, 5:53-67, 24:7-25:57.

D. Independent Claims

71. In my opinion, Janevski teaches most of claim 1, including multiple PVRs and a PVR performing the initiator role to control synchronous playback based

on commands from a remote. Janevski, Abstract, 6:4-26, 15:43-47. Okamura teaches alternative timing information that identifies a future time for playback to maintain synchronous playback. Okamura, 1:42–56; 5:53–67. Kawamura discloses that a user can send commands from a remote using a network interface, such as Bluetooth, rather than infrared. Thus, in my opinion, a POSITA would have found claim 1 obvious over the combination of Janevski, Kawamura, and Okamura.

1. Claim 1

a. [1preamble]: “A method performed by a first zone player, the method comprising:”

72. In my opinion, Janevski teaches [1preamble] because it discloses a sending PVR (first zone player) that performs a “method for synchronizing presentation from bit streams based on their content.” Janevski, 1:1-3; *see also id.*, 1:8-11, 13:22-24, 17:7-16. In my opinion, Janevski’s PVRs “play[] back respective, identical copies of a video,” which included audio, and meet the agreed upon construction of a “data network device configured to process and output audio.” Janevski, 2:6–15; Ex. 1015.

b. [1a]: receiving, via a network interface at the first zone player, a request to engage in synchronous playback of audio content as part of a synchrony group that includes at least a second zone player that is communicatively coupled to the first zone player via at least one data network;

73. In my opinion, Janevski in combination with Kawamura teaches [1a].

74. Janevski discloses that “PVR 114a” (first zone player) and “PVR 114b” (second zone player) form a “synchronized viewing session” (synchrony group), where the “PVR 114a” (first zone player) performs an “initiator” (master) role. Janevski, 1:13, 6:13-27, 7:57-63, 8:42-47, FIG. 1.

75. Janevski’s “[p]ersonal video recorder (PVR)” “114a” discloses *a first zone player*. *Id.* Janevski’s “[p]ersonal video recorder (PVR)” “114b” discloses *a second zone player*. *Id.* Each “PVR” describes a zone player because they are “playback” devices that “playback ... renderable content,” including “visual” and/or “aural” information (audio content). Janevski, Abstract, 1:8-11, 2:23-45, 5:29-32, 10:30-35, 15:22-26, 16:33-43.

76. In my opinion, Janevski’s PVRs also contain at least one *network interface*. First, as explained below, PVRs use an Internet connection that allows the first zone player to communicate with other zone players, satisfying *a network interface at the first zone player* requirement of the element. In my opinion, as combined with Kawamura, the PVR would also include a second *network interface* to enable Bluetooth communication. The ’001 patent similarly describes its zone players being implemented with multiple network interface devices, including those that connect zone players to a local network and those that connect the local network to external networks such as the Internet or PSTN. ’001 patent, 4:14-21; 24:45–

25:57. Therefore, the network interface referenced in the claim may encompass multiple interfaces. Moreover, in my opinion, Janevski's disclosure of communication over the Internet "to synchronize their playbacks" (Janevski, Abstract, 6:44–49) suggests the presence of a network interface, as Internet-based communication necessarily requires such an interface. *See* Kessler, *Overview of TCP/IP and the Internet*, 1994. In my opinion, the network interface used for the Internet communication is the network interface used to send the messages for the remaining elements of claim 1.

77. The "PVR" "114a" and "PVR" "114b" playback audio content in "a **synchronized viewing session**," which discloses *a synchrony group* comprising the two zone players "PVR" "114a" and "114b." Janevski, 6:4-38, FIG. 1; *see also id.*, Abstract, 5:50-51, 7:4-16, 7:25-50, 11:4-11, 12:41-46. Figure 1 illustrates the "synchronized viewing session," where "[t]he Internet 118 [the claimed data network] supplies the means 119 for communicating information between the PVRs 114a, b such that synchronization may be achieved." Janevski, 6:48-51. Accordingly, In my opinion, the devices are *communicatively coupled*.

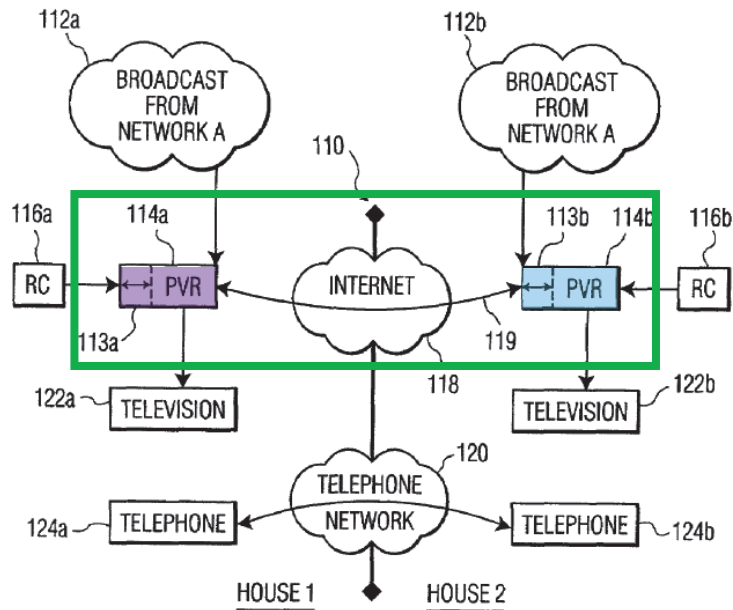


FIG. 1

Janevski, FIG. 1.

78. In my opinion, Janevski further teaches *a request to engage in synchronous playback* by disclosing that a control command from a remote control or keyboard connected to the initiator PVR 114a activates a menu for synchronized playback and sends the menu for synchronized playback to the other participants. Janevski, 7:4-16. Specifically, Janevski discloses that “user 1 takes the lead as the ‘initiator’, user 1, via a remote control 116, via controls on the PVR 114a itself, ... [which] would activate a menu for synchronized viewing on the PVR 114a.” Janevski, 7:4-9. More specifically, Janevski discloses that “User 1 would respond to and send the menu. Resulting other menus would be sent to user 2. The users would

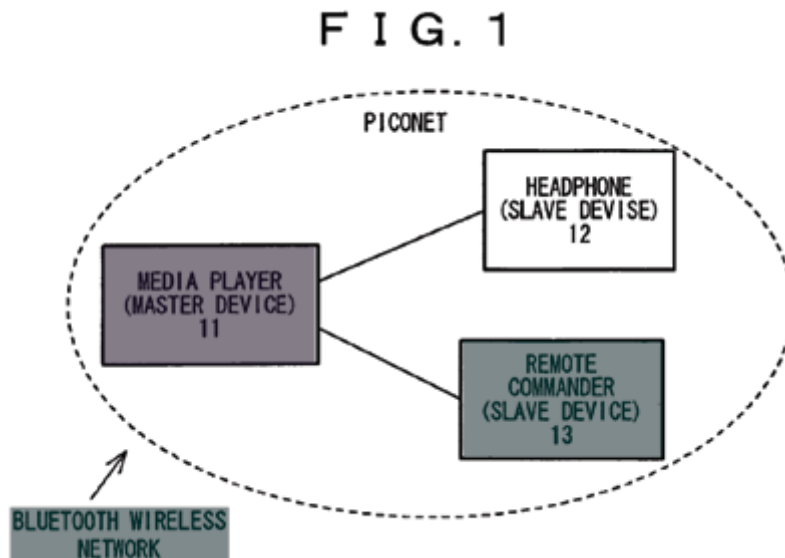
each indicate by their respective response to the menus they received whether or not they will be participating in the session. Based on the responses, a synchronized viewing session is established to begin at an agreed upon time.” Janevski, 7:9-16. By taking the lead “via a remote control 116” and “indicat[ing] by ... response to the menus” to participate in the session, Janevski teaches a request to engage in synchronous playback. Janevski, 7:9-16. Janevski also discloses that the “PVR” “114a” performs the “initiator” role for the “synchronized viewing session.” Janevski, Abstract, 6:16-22, 8:39-52; *see also id.*, 6:4-38, 7:25-50, 11:12-42.

79. Kawamura teaches that it was also known to receive commands from a “remote commander” via a “Bluetooth interface” (via the network interface) instead of using infrared. Kawamura, [0088]-[0090].

80. In Janevski’s exemplary embodiment, the “PVR 114a includes ... a remote control sensor 204,” which “receives a signal from the remote control 116.” Janevski, 7:51-8:1. But Janevski is not limited to this configuration. In my opinion, while Janevski’s “remote control” is “typically” “configured to transmit an infrared signal,” a POSITA would have understood other wireless communication techniques were compatible with and advantageous for Janevski. Janevski, 6:40-44.

81. In my opinion, Kawamura discloses one known suitable option to replace infrared—Bluetooth. Kawamura discloses a Bluetooth “piconet 10”

including “media player 11 ... and a remote commander 13.” Kawamura, [0088], [0090]-[0091], FIGs. 1, 2. Within the piconet, “[t]he remote commander 13 remotely controls the media player 11 for executing play control such as play, stop, pause, fast forward and rewind,” which discloses a controller device transmitting playback commands to the media player. Kawamura, [0088].

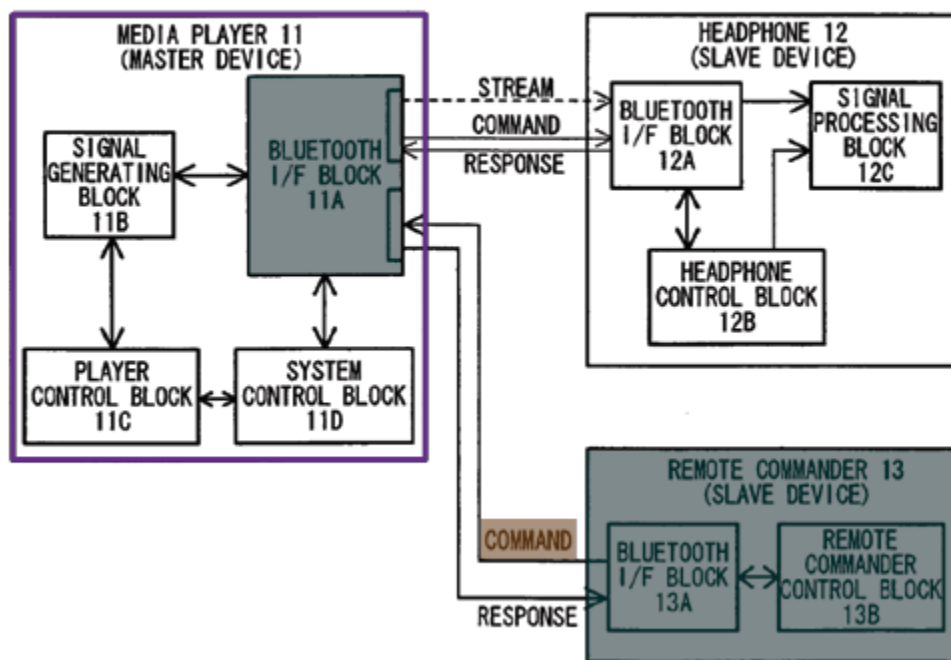


Kawamura, FIG. 1.

82. Kawamura’s “media player” includes “a Bluetooth interface,” which is a network interface that connects to the “Bluetooth wireless network.” Kawamura, [0090]-[0091], FIGs. 1-, 2. The “Bluetooth interface” of “media player” implements “Bluetooth wireless connection in piconet 10 and transfer[] commands and responses with the ... [remote commander] 13.” *Id.* Kawamura’s “remote commander” similarly includes “a Bluetooth interface” which “transfers commands

and responses with the [media player] 11.” Kawamura, [0100]-[0104]. Thus, Kawamura’s “media player” (zone player) receives “commands” from the “remote commander” (network device) via the “Bluetooth interface.” (network interface). Kawamura, [0088], [0090]-[0091].

F I G . 2



Kawamura, FIG. 2.

83. In my opinion, as combined, Kawamura’s “remote commander” would have served as Janevski’s “remote control” and Janevski’s “PVR 114a” would have received its “control function” via Kawamura’s “Bluetooth interface.”

84. In my opinion, a POSITA would have sought Bluetooth solutions because Kawamura teaches that “short-distance wireless communication,” including

“Bluetooth,” is “expected a great deal” for use “as the transmission media for home networks.” Kawamura, [0003]-[0005]. Kawamura also teaches that “obviously, [Bluetooth] is also designed for the application to stationary equipment” including “a stereo unit (or other media players).” Kawamura, [0010].

85. In my opinion, a POSITA would have recognized that infrared communications (as in Janevski’s exemplary embodiment) are limited to line-of-sight transmission. In my opinion, for example, the prior art explains that “[m]any video sources, especially VCRs and cable converters, are designed to cooperate with hand-held controllers that send out infrared control signals upon command of the user,” but “[u]nfortunately signals from these devices do not travel between rooms unless there is a line-of-sight path between transmitter and source.” Goodman, 4:43-49; Braithwaite, 1:54-60. So, in my opinion, a POSITA would have been motivated to implement Kawamura’s Bluetooth remote to improve communications and successfully transmit commands to Janevski’s PVR even when walls, furniture, people, or other household obstacles obstructed the line-of-sight.

86. In my opinion, a POSITA would have reasonably expected success because Janevski and Kawamura use similar components in similar ways to achieve similar goals, e.g., AV media players, control signals, and remote controllers. *See*, Kawamura, [0088]-[0091], [0206]; Janevski, 6:4-44. Moreover, “because Bluetooth

is a worldwide standard ... the globalization of communication environments may be promoted with ease.” Kawamura, [0011]. Kawamura’s system also uses “a conventional Bluetooth control profile such as AV Remote Control Profile.” Kawamura, [0023]. Janevski’s “PVR 114a” is also already configured with an Internet network interface (“receiver[] 113a”) as explained above. Janevski, 7:57-63; 6:6-16. Thus, in my opinion, implementing Kawamura’s Bluetooth “remote and interface in Janevski (to the extent not expressly disclosed) would have used known features according to their known functions (remote controls and Bluetooth), to achieve predictable results (Janevski’s PVR receives the control function via a Bluetooth interface).

87. In my opinion, a POSITA would further appreciate that implementing Bluetooth as taught by Kawamura in Janevski’s synchronized playback system would not require any undue experimentation, as both references describe standard network interfaces and remote control protocols that are widely adopted in consumer electronics. In my opinion, the technical compatibility between Bluetooth-enabled media players and networked playback devices is well-established, and the prior art demonstrates that AV control signals, including playback and synchronization commands, can be reliably transmitted over Bluetooth or Internet interfaces. Indeed,

a specific Bluetooth profile, Audio/Video Remote Control Profile (AVRCP 1.0), was standardized by the Bluetooth Special Interest Group (SIG) on May 22, 2003.

88. Moreover, in my opinion, a POSITA would recognize that integrating Bluetooth remote control into Janevski's system would enhance user experience by enabling seamless control of synchronized playback across multiple devices, regardless of physical barriers or room layout. Because Bluetooth is a global standard designed for interoperability and multi-device environments, a POSITA, in my opinion, would have a reasonable expectation of success in combining these features to achieve predictable and robust results in home network audio systems.

89. In my opinion, Janevski's "receiver" and Kawamura's "Bluetooth interface" meet the construction of "physical component of a device that provides an interconnection with a data network" because they are physical components that connect to the "Internet" and "Bluetooth" networks, respectively. Janevski, 7:57-63; 6:6-16; Kawamura, [0088]-[0091]. The Bluetooth network interface conveys to the first zone player the user's request to engage in synchronous playback of audio content as part of a synchrony group and Janevski's receiver conveys (from the Internet) the remote user's request to engage in synchronous playback of audio content as part of a synchrony group.

90. Therefore, in my opinion, Janevski when combined with Kawamura teaches receiving, via a network interface at the first zone player, a request to engage in synchronous playback of audio content as part of a synchrony group that includes at least a second zone player communicatively coupled to the first zone player via at least one data network.

c. [1b] “after receiving the request to enter into the synchrony group:”

91. In my opinion, Janevski in combination with Kawamura teaches [1b] because Janevski discloses that its synchronization process occurs after the request to participate in a synchronized viewing session and the corresponding responses are received. Janevski, 7:25-50. Specifically, Janevski discloses that “[b]ased on the responses, synchronized viewing session is established to begin at an agreed upon time.” Janevski, 7:14-16. Alternatively, a user may initiate a request to engage in synchronous playback of audio via the Bluetooth interface by the user of Janevski’s PVRa as modified by Kawamura’s teachings as explained above in [1a].

d. [1ci] “detecting an indication that the first zone player is to operate in (a) one of a control-master mode or a control-slave mode for the synchrony group”

92. In my opinion, Janevski teaches [1ci] by disclosing that Janevski’s PVR synchronized viewing system dynamically assigns roles based on user actions after the synchronized viewing session is started. Janevski, 8:48-52. Specifically,

Janevski discloses that “[e]ach time that a participant changes the watching mode— i.e. executes a control function (rewind, fast forward, etc.)—that participant broadcasts a command to all other participants (including the ‘current’ initiator) and thereby becomes the new initiator.” Janevski, 8:42–52. In my opinion, this broadcasted command is the *indication* that is detected by the system to determine role assignment.

93. In my opinion, the first zone player detects that it is to operate in control-master mode when it “executes a control function ... [and] broadcasts a command” to all participants “and thereby becomes the new initiator.” Janevski, 8:48-52. In my opinion, this confirms that Janevski’s PVR synchronized viewing system detects the control function from the user’s remote control as an indication that the device should assume the initiator role (control-master mode).

94. In my opinion, conversely, the first zone player detects that it is to operate in control-slave mode when it receives a command broadcasted by another PVR that has executed a control function and assumed the initiator role. Janevski, 8:48-52. Janevski discloses that “the initiator performs the same process to synchronize each participant, whereby all participants become synchronized.” Janevski, 8:42–48. In my opinion, this shows that the receiving PVRs detect their

role as participants (i.e., control-slaves) based on the broadcasted command and synchronize playback accordingly.

e. [1cii] “and (b) one of an audio-master mode or an audio-slave mode for the synchrony group;”

95. In my opinion, Janevski teaches [1cii] by describing synchronized playback of aural content across multiple PVRs (synchrony group). Janevski, 5:10-12; 16:36-37. Janevski discloses that “renderable content, as used herein, refers to content that is presentable in a form that a user can sense, e.g. visually or aurally.” Janevski, 5:10-12. Janevski discloses synchronizing playback of “two or more digital bit streams,” and “the presentation may be merely visual or merely aural or both visual and aural.” Janevski, 16:36-37. In my opinion, this confirms that Janevski’s PVR synchronized viewing system supports synchronized audio playback, and that each PVR may be involved in presenting aural content.

96. As explained in [1ci], in my opinion, Janevski discloses that initiator PVR (e.g., 114a) controls synchronization and presenting content in aural form, thereby operating in an audio-master mode. Janevski, 5:10-12; 6:1-22; 16:36-37. Janevski discloses that “the initiator performs the same process to synchronize each participant, whereby all participants become synchronized.” Janevski, 8:42–48. The participant PVRs (e.g., 114b) receive synchronization commands and adjust accordingly, operating in an audio-slave mode through aural presentation. Janevski,

8:48-52. Janevski further discloses that “the characteristic information may be characteristic of viewable images or of audible sounds,” confirming that synchronization applies to audio content. Janevski, 16:38–41.

97. In my opinion, to enable synchronized playback, Janevski describes a time synchronization between the initiator and the participants, disclosing that “[t]ime synchronization is performed by the initiator PVR 114a individually with each participant PVR 114b, and involves sending an originating synchronization message ... and sending a reply synchronization message Differences between the respective transmission times are resolved by... ‘fine tune’ aligning based on content of respective digital bit streams.” Janevski, 8:65-9:14. In my opinion, this shows that the initiator PVR detects its role as the **audio-master** by initiating synchronization and distributing audio and timing information. In my opinion, the participant PVRs detect their role as **audio-slaves** by responding to synchronization messages and aligning playback and timing accordingly.

98. Thus, in my opinion, Janevski teaches that each PVR detects whether it is operating in an audio-master or audio-slave mode based on its role in synchronizing playback of aural content.

f. [1d] “beginning to operate in the synchrony group in accordance with the indication;”

99. In my opinion, Janevski teaches [1d] by disclosing that “[e]ach time that a participant changes the watching mode i.e. executes a control function (rewind, fast forward, etc.), that participant broadcasts a command to all other participants (including the ‘current’ initiator) and thereby becomes the new initiator.” Janevski, 8:48–52. Following this role change, Janevski discloses that “the initiator performs the same process to synchronize each participant, whereby all participants become synchronized” (synchrony group). Janevski, 8:42–47. This includes PVRs broadcasting status messages and commands to the other PVRs to maintain synchronization in their respective roles. Janevski, 7:25-50. As explained in [1ci], in my opinion, this broadcasted command is the *indication* that is detected by the system to determine role assignment. Therefore, in my opinion, Janevski confirms that the zone player begins operating in the synchrony group in accordance with the detected indication—whether as initiator or participant—by initiating or responding to synchronization commands.

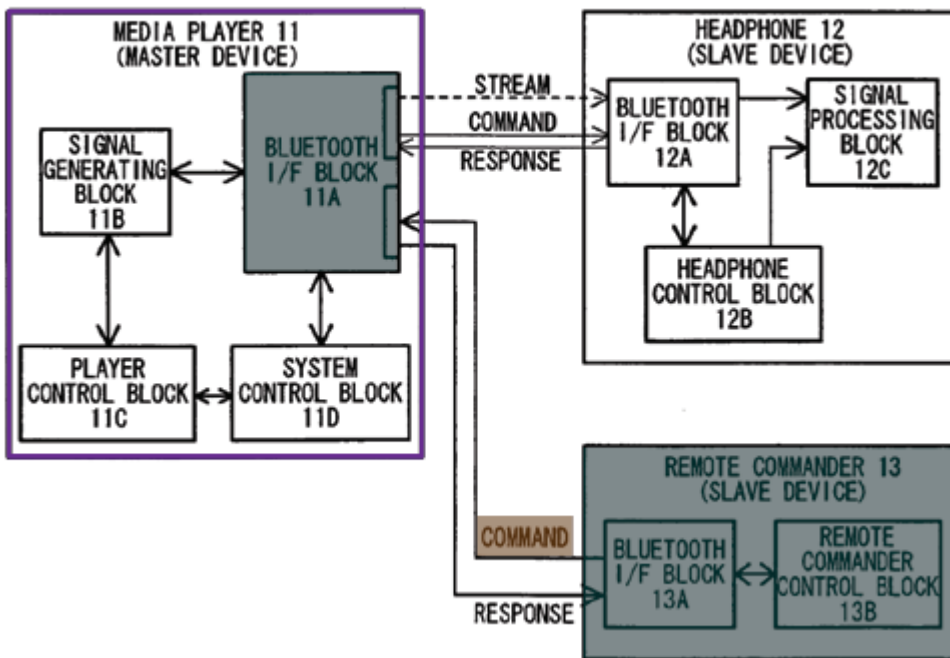
- g. [1e] “wherein, while operating in the control-master mode for the synchrony group, the first zone player is configured to: receive, via the network interface, first control information for the synchrony group from a network device that is communicatively coupled to the first zone player; and”**

100. In my opinion, a POSITA would have found claim element [1e] obvious over the combination of Janevski and Kawamura. First Janevski teaches most of [1e] by disclosing that a “remote control 116a, b is commonly associated with the personal video recorder 114a, b to allow the user to operate the personal video recorder 114a, b remotely.” Janevski, 6:39–41. While operating in control-master mode, the initiator receives this control input and uses it to direct synchronization: “Each time that a participant changes the watching mode i.e. executes a control function (rewind, fast forward, etc.), that participant broadcasts a command to all other participants ... and thereby becomes the new initiator.” Janevski, 8:42–52. A control input for executing a control function is the claimed *first control information*. Janevski’s PVR synchronized viewing system includes “an Internet network 118 that interconnects the PVRs 114a, b ... such that synchronization may be achieved,” (synchrony group) confirming that the control information is received and transmitted via a network interface. Janevski, 6:44–49.

101. In my opinion, as discussed above in element 1a, Kawamura teaches that it was also known to receive commands from a “remote commander” (e.g., a

network device) via a “Bluetooth interface” (via the network interface). Kawamura, [0088]-[0090].

F I G . 2



Kawamura, FIG 2.

102. In my opinion, Janevski at least suggests receiving the playback commands from a network device via a network interface because Janevski discloses a “remote control 116a, b to allow the user to operate the personal video recorder 114a, b remotely.” Janevski, 6:39-44. The “remote control” is a network device *communicatively coupled* with the PVR, as shown in Figure 1. *Id.*

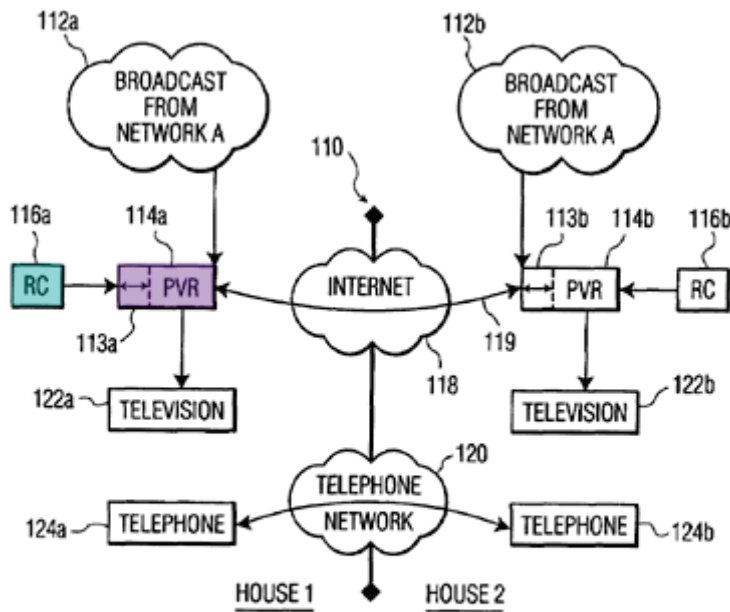


FIG. 1

Janevski, FIG. 1.

103. In my opinion, in Janevski's exemplary embodiment, the "exemplary PVR 114a includes ... a remote control sensor 204," which "receives a signal from the remote control 116." Janevski, 7:51-8:1. But Janevski is expressly not limited to this configuration in my opinion, and while Janevski contemplates that its "remote control" is "typically" "configured to transmit an infrared signal to the television 122a, b," a POSITA would have understood other wireless communication techniques were compatible with and advantageous for Janevski's system. Janevski, 6:40-44.

104. In my opinion, as discussed above in element [1a], Kawamura discloses one known suitable option to replace infrared—Bluetooth. And Kawamura’s “media player” includes “a Bluetooth interface,” which is a network interface that connects to the “Bluetooth wireless network.” Kawamura, [0090]-[0091], FIGs. 1-, 2. The “Bluetooth interface” of “media player” implements “Bluetooth wireless connection in piconet 10 and transfer[] commands and responses with the ... [remote commander] 13.” Kawamura, [0091]. Kawamura’s “remote commander” similarly includes “a Bluetooth interface” which “transfers commands and responses with the [media player] 11.” Kawamura, [0100]-[0104]. Thus, Kawamura’s “media player” (zone player) receives “commands” from the “remote commander” (network device) via the “Bluetooth interface.” (network interface). Kawamura, [0088], [0090]-[0091].

105. In my opinion, as combined above and in element [1a], Kawamura’s “remote commander” would have served as Janevski’s “remote control” and Janevski’s “PVR 114a” would have received its “control function” via Kawamura’s “Bluetooth interface.” In my opinion, this combination would have been obvious to a POSITA for the reasons discussed in element 1a.

106. In my opinion, Janevski’s “receiver” and Kawamura’s “Bluetooth interface” meet the construction of “physical component of a device that provides

an interconnection with a data network” because they are physical components that connect to the “Internet” and “Bluetooth” networks, respectively. Janevski, 7:57-63; 6:6-16; Kawamura, [0088]-[0091].

h. [1f] “based on the first control information, cause, via the network interface, at least one playback action to be applied in the synchrony group;”

107. In my opinion, Janevski teaches [1f] by disclosing “control function[s],” such as “stop, pause, fast forward, [and] reverse.” Janevski, 6:18-22; Abstract. Each of these “control function[s]” is a playback action because they control playback. *Id.* For example, “fast forwards” means “moves forward in the video 308 faster than the video moves in normal play.” 14:34-43.

108. In my opinion, the “initiator” controls synchronous playback in response to the “control function,” which is the claimed *first control information* as explained in element 1e, because “[t]he initiator directs all participants ... **upon execution of each control function** ... to synchronize their playbacks to that of the initiator, whereby all playbacks are synchronized and maintain[ed] in synchronization” (synchrony group). Janevski, Abstract (emphasis added).

109. In my opinion, these playback actions are applied across the synchrony group via the network interface, as the initiator “broadcasts a command to all other participants.” Janevski, 8:42–52. Janevski also discloses *via the network interface*,

by disclosing that “an Internet network 118 interconnects 114a, b, located at the two different locations....[and] supplies the means 119 for communicating information between the PVRs 114a, b such that synchronization may be achieved.” Janevski, 6:45-51. In my opinion, a POSITA would have understood that the term “network interface” is not limited to a single physical or logical interface. As described in the ’001 patent, the system may include multiple network interface devices, including those that connect zone players to a local network and those that connect the local network to external networks such as the Internet or PSTN. ’001 patent, 4:14-21; 24:45–25:57. In my opinion, therefore, the network interface referenced in the claim may encompass multiple interfaces.

110. Moreover, in my opinion, Janevski’s disclosure of communication over the Internet “to synchronize their playbacks” (Janevski, Abstract, 6:44–49) suggests the presence of a network interface, as Internet-based communication necessarily requires such an interface. *See Kessler, Overview of TCP/IP and the Internet, 1994.* In my opinion, the network interface used for the Internet communication is the network interface used to send the messages for the remaining elements of claim 1.

- i. **[1g] “wherein, while operating in the control-slave mode for the synchrony group, the first zone player is configured to: receive, via the network interface, second control information from another zone player; and”**

111. In my opinion, Janevski teaches [1g] by disclosing that “[e]ach time that a participant changes the watching mode i.e. executes a control function (rewind, fast forward, etc.), that participant broadcasts a command to all other participants ... and thereby becomes the new initiator.” Janevski, 8:42–52. As discussed above in element 1e, in my opinion, these control functions of Janevski teach the claimed control information. In my opinion, in the elements above, the first zone player is the initiator (claimed *control-master*) that receives control functions from its remote and broadcasts them to the other PVRs and, those control functions are the claimed *first control information*. However, once a PVR other than the first zone player becomes the initiator, then the first zone player operates as a participant (i.e., in control-slave mode) and receives these commands from the current initiator.

112. In my opinion, Janevski also discloses sending and receiving those commands *via the network interface*, by disclosing that “an Internet network 118 interconnects 114a, b, located at the two different locations ... [and] supplies the means 119 for communicating information between the PVRs 114a, b such that synchronization may be achieved.” Janevski, 6:45-51.

j. [1h] “perform one or more playback actions in accordance with the second control information;”

113. In my opinion, Janevski teaches [1h] by disclosing that “each time that a participant changes the watching mode i.e. executes a control function (rewind, fast forward, etc.) [(i.e., *second control information*)], that participant broadcasts a command to all other participants.” Janevski, 8:42–52. The broadcast command is the second control information as explained in [1g], Janevski also discloses that “that participant’s PVR 114b broadcasts a command for that function that is immediately communicated and effected [(i.e., performed)] in the PVR 114b of each participant, to keep the presentation synchronized.” Janevski, 11:31–36. Therefore, in my opinion, Janevski discloses that the receiving PVRs apply playback actions, such as rewind or fast forward, in accordance with control information received from another zone player.

k. [1i] “wherein, while operating in the audio-master mode for the synchrony group, the first zone player is configured to: obtain audio information that is representative of the audio content;”

114. In my opinion, Janevski teaches [1i] by disclosing that “PVR 114a” performs the “initiator” role for the “synchronized viewing session.” Janevski, Abstract, 6:16–27, 8:39–52; *see also id.*, 6:4–38, 7:25–50, 11:12–42. The “initiator” role reflects a master device that controls synchronous playback because “[t]he

initiator directs all participants ... to synchronize their playbacks to that of the initiator, whereby all playbacks are synchronized and maintain[ed] in synchronization.” Janevski, Abstract. Specifically, “time synchronization is performed by the initiator PVR 114a individually with each participant 114b.” Janevski, 6:16–27, 8:65–9:4, 15:32–33, 16:34–43.

115. In my opinion, Janevski discloses synchronizing playback of audio information because it “applies generally to synchronizing presentation of renderable content of two or more digital bit streams” which “may be ... aural.” Janevski, 6:16–27, 16:34–43; see also *id.*, claim 3 (“said streams include audio data and said presentation is at least partially aural.”). Thus, in my opinion, the initiator PVR that initiates synchronized playback of the audio is operating in audio-master mode.

116. In my opinion, to control synchronization, “a status message is sent out periodically by the ‘initiator’,” i.e., the device operating in audio-master mode, “[t]o ensure that the PVRs 114a, b participating in a session remain synchronous.” Janevski, 7:39–50; see also *id.*, 1:53–2:5. “The status message includes” information to ensure synchronized playback like “the current mode of watching (e.g., normal play, fast forward, pause), an indication of the time into the program, and *information characteristic of content of a digital bit stream.*” *Id.* (emphasis added).

In my opinion, for example, Janevski discloses comparing “corresponding content or ‘landmarks’ of pairs of video playbacks to be synchronized.” Janevski, 3:53-57. This “characteristic information may be characteristic of ... audible sounds” and Janevski discloses that “audio transform coefficients” “can be used to characterize the content of the bit streams.” Janevski, 16:38-43.

117. In my opinion, a person of ordinary skill in the art would understand that, in order to synchronize playback, the system must identify which portion of the audio content is being played, and this identification can be accomplished using landmarks or any information that identifies the location of the content – which, for example, can be a piece of the content itself. In my opinion, a POSITA would have understood that “audio transform coefficients” are computed from frames of audio samples and are part of the compressed audio stream *i.e.*, they are *obtained*. *Id.* The prior art explains that an “audio signal” is “transformed from the time domain ... to another domain that facilitates analysis” and “[t]he transformation produces a set of amplitude coefficients of a variable other than time, typically frequency.” Lim, [0006]-[0008]. “The transform coefficients [are] derived from a frame.” Lim, claim 3, [0006]-[0012]. Therefore, in my opinion, audio transform coefficients are audio information. In my opinion, this is consistent with the ’001 patent specification which explains that audio information can be less than the full content for playback

because, for example, “each frame compris[es] digital audio information for a predetermined period of time.” Ex. 1001, 19:49-53. Therefore, in my opinion, Janevski discloses obtaining audio information that is representative of the audio content by disclosing obtaining “information characteristic of content of a digital bit stream,” which can include audio information to identify the location of the content for synchronization. Janevski, 7:39–50.

- I. **[1j] “generate playback timing information associated with the obtained audio information that is indicative of at least one future time relative to a reference clock time that denotes a time at which at least the first and second zone players are to engage in synchronous playback of a corresponding portion of the obtained audio information; and”**

118. In my opinion, a POSITA would have found claim element [1j] obvious over the combination of Janevski, Kawamura, and Okamura. As discussed above, Janevski discloses that “to achieve precise synchronization, the present invention compares corresponding content or ‘landmarks’ of pairs of video playbacks to be synchronized, determines video replay ‘distance’ between the landmark pairs, and slows down or speeds up selected playbacks in accordance with these distances.” Janevski, 3:52–57. This includes providing synchronization messages with “obtained audio information” in the form of an “information characteristic of content of a digital bit stream” identifying those landmarks. Janevski, 7:40-50. In my

opinion, based on the disclosures of Okamura it also would have been obvious to a person of skill in the art to also include the claimed playback timing information associated with those landmarks (i.e., the obtained audio information).

119. In my opinion, Janevski already disclosed that the synchronization messages should include playback timing information. For example, Janevski discloses that “time synchronization is performed by the initiator PVR 114a individually with each participant PVR 114b, and involves sending an originating synchronization message 402 from the initiator PVR 114a to a participant PVR 114b and sending a reply synchronization message 404 from the participant PVR 114b to the initiator PVR 114a.” Janevski, FIG. 4; 8:39-10:2. Janevski discloses that “each message is timestamped based on the internal video timer of the respective PVRs at the moment of sending or receiving.” Janevski, 8:65–9:4. Janevski’s PVR synchronized viewing system collects four time points: “the message 402 is sent at a time A ... arrives at the participant ... at a time B ... the participant sends back a reply ... at a time C ... received by initiator ... at a time D.” Janevski, 9:15-35. The time misregistration (TM) is calculated using the formula: $TM = \frac{1}{2}[(A + D) - (C + B)]$.” Janevski, 9:30-35. Janevski explains, “once the TM value is calculated, the participant PVR adjusts its timer forward or backward by that amount to match the initiator's playback time.” Janevski, 9:11–13. The “differences between the

respective transmission times are resolved by the present innovative ‘fine tune’ aligning based on content of respective digital bit streams of the initiator and the participant.” Janevski, 9:10-14. In my opinion, this method allows Janevski’s PVR synchronized viewing system to determine whether the initiator or participant is ahead or behind, and by how much, based on the sign and magnitude of TM. For example, in CASE 2, “the initiator’s video timer 212 lags the participant’s timer 212,” and in CASE 3, “the participant lags the initiator by ½ second.” Janevski, 9:40-10:3. After determining time misregistration, Janevski discloses determining frame misregistration in which media characteristics are compared between devices to ensure that the media time stamp information at the different devices is accurate. If it is not, then further compensation is employed to synchronize the playback. For convenience in referring to Janevski’s process, I call this *reactive* synchronization because it generates timing compensation in response to media and transmission characteristics rather than preventing frame misregistration as will be discussed below with respect to Okamura.

120. But, in my opinion, Janevski also discloses that “[t]ime synchronization can be implemented in many different known ways.” Janevski, 8:53-54. So, in my opinion, while Janevski does not disclose *future time relative to a reference clock*

time as its exemplary timing solution, a POSITA would have been motivated to look to other known solutions.

121. In my opinion, Okamura discloses that using playback timing information associated with the obtained audio information that is indicative of at least one future time relative to a reference clock time that denotes a time at which at least the first and second zone players are to engage in synchronous playback of a corresponding portion of the obtained audio information was another known solution that would have been obvious to a POSITA to implement. In my opinion, Okamura offers a known improvement over reactive solutions like Janevski by disclosing that the system “generates a timestamp (abbreviated as ‘syt’), one in every 8 sampling clocks (or 8 data blocks), on the basis of the time provided by the cycle timer... The timestamp specifies the reproduction time at the receiving side of an event sequence.” Okamura, 1:42–56. Okamura also discloses that “each transmitting node also arranges audio data of one or more channels into a data field and arranges the associated timestamp into a syt field so as to form a packet composed of the data field and the syt field, and sends the packet.” Okamura, 1:42–56. Therefore, in my opinion, Okamura shows that the playback timing information is packaged (i.e., included in the same packet) with the audio data and is associated with the specific portions of the audio stream to enable synchronized playback.

122. In my opinion, a POSITA would have recognized that Janevski's approach of broadcasting content and sending synchronization messages separately, and Okamura's approach of packaging playback timing information with audio data in the same packet, are both known and suitable options for achieving synchronized playback. In my opinion, it would have been obvious to implement either method, or a combination thereof, to coordinate playback timing across multiple devices in a synchrony group.

123. In my opinion, the receiving node (slave) uses its own internal cycle timer as a local clock and initiates playback when the adjusted timestamp matches its current time. Okamura discloses that the system includes "an offset setting section that sets an offset time for the receiving node relative to the transmitting node and adds the offset time to a time indicated by the timestamp..." and "a reproduction time control section that operates when the time of the timestamp added with the offset time coincides with a current time indicated by an internal cycle timer for controlling the data output section to effect synchronous reproduction." Okamura, 5:53–67.

124. In my opinion, Okamura also accounts for transmission delay, disclosing that "on the transmitting side, the time of timestamp is set as the value of a reproduction time on the receiving side by estimating propagation delay." Okamura, 24:7–9. Additionally, Okamura teaches that "[b]y adjusting the offset

value on the receiving side, the time of reproducing the audio data supplied from each transmitting node can be shifted from the time of timestamp.” Okamura, 24:9–12.

125. In my opinion, these teachings align with the ’001 patent’s requirement to generate playback timing information indicative of a future time for synchronized playback. See ’001, 2:49–67, 3:1–19, 24:45–25:57. Okamura teaches that the transmitting node (e.g., the first zone player) generates a timestamp “on the basis of the time provided by the cycle timer” and associates it with specific portions of the audio stream. Okamura, 1:42–56. The timestamp is transmitted along with the audio data and “specifies the reproduction time at the receiving side of an event sequence,” enabling the receiving node (e.g., the second zone player) to initiate playback of the corresponding portion of the audio information at the designated future time. *Id.* Thus, in my opinion, Okamura discloses generating playback timing information that is both associated with the obtained audio information and indicative of a future time at which at least the first and second zone players engage in synchronous playback of that audio content.

126. In my opinion, a POSITA would have been motivated to combine Janevski and Kawamura with Okamura to improve synchronization precision and scalability across networked playback devices. Janevski relies on reactive

synchronization using internal video timers and message exchanges to correct misalignment between devices. Janevski, 8:65–10:3. However, in my opinion, this approach assumes symmetrical transmission delays and lacks a predictive mechanism for future playback timing based on a shared reference clock. Janevski also discloses that if playback is obtained from two different sources, then playback between devices may become misaligned due to slight variations in the program of interest 2:59-3:39. A POSITA seeking to address this problem would also find a solution in the teachings of Okamura which ensure synchronized playback by sending audio information along with the synchronization information, thus avoiding variations in the audio source data.

127. In my opinion, Okamura addresses these limitations by teaching a synchronization method that uses a cycle timer as a reference clock to generate timestamps representing future playback times, and adjusts for propagation delay and clock offset. Okamura, 1:42–56, 5:53–67, 24:7–12. In my opinion, specifically in the context of the Janevski, Kawamura, and Okamura combination, Okamura's cycle timer corresponds to the reference clock maintained by the initiator zone player in Janevski's system. In my opinion, A POSITA would have recognized that Okamura's predictive synchronization mechanism could be advantageously

implemented in Janevski's multi-device synchronization framework system to improve timing accuracy and reduce reliance on reactive correction.

128. In my opinion, A POSITA in the field of networked audio-visual systems would have a strong foundational understanding of both reactive and predictive synchronization techniques, as well as the practical challenges of achieving precise timing across distributed devices. In my opinion, given the prevalence of timestamping, clock synchronization, and packet-based transmission in standards like IEEE 1394 and Bluetooth AV profiles, a POSITA would be well-versed in methods for associating playback timing information with audio data to enable coordinated playback. In my opinion, the skill level in the art includes familiarity with both hardware and software implementations for time alignment, buffer management, and compensation for network delays, making the integration of Okamura's predictive timestamping approach into Janevski's synchronization framework a routine and straightforward task.

129. Furthermore, in my opinion, a POSITA would recognize the advantages of moving from purely reactive synchronization (as in Janevski) to predictive, reference-clock-based synchronization (as in Okamura). The ability to generate and transmit playback timing information that specifies a future time for synchronized playback allows for more robust and scalable multi-zone systems,

reduces the risk of drift or misalignment, and improves user experience by ensuring seamless audio across devices. In my opinion, the combination leverages well-known engineering principles and industry standards, and a POSITA would expect that implementing Okamura's timestamping and offset adjustment techniques in Janevski's system would yield predictable, reliable, and commercially valuable results.

130. Moreover, in my opinion, both references are directed to synchronizing playback across AV devices in networked environments and use similar components to achieve similar goals. See Janevski, 6:4–44; Okamura, 1:42–56, 5:53–67. In my opinion, a POSITA would have reasonably expected success in combining Okamura's reference-clock-based timestamping with Janevski's multi-device synchronization framework to achieve more robust and scalable synchronous playback.

131. In my opinion, implementing Okamura's timestamp generation in Janevski's multi-device synchronization framework would have used known techniques (reference clocks and offset adjustments) according to their known functions (predictive synchronization), to achieve predictable results (accurate future playback timing across devices).

- m. **[1k] “transmit, via the network interface, the obtained audio information and the generated playback timing information to the second zone player; and”**

132. In my opinion, a POSITA would have found claim element [1k] obvious over the combination of Janevski, Kawamura, and Okamura. First, Janevski discloses that “the Internet 118 supplies the means 119 for communicating information between the PVRs 114a, b such that synchronization may be achieved.” Janevski, FIG. 1, 6:44–49. Accordingly, messages sent in Janevski are sent “via the network interface” used to communicate over the Internet, as discussed above. Janevski, 6:45-51. To control synchronization, “a status message is sent out periodically by the ‘initiator’, i.e., the PVR 114a” “[t]o ensure that the PVRs 114a, b participating in a session remain synchronous.” Janevski, 7:39-50; *see also id.*, 1:53-2:5. “The status message includes” information to ensure synchronized playback like “the current mode of watching (e.g., normal play, fast forward, pause), an indication of the time into the program, and information characteristic of content of a digital bit stream [(obtained audio information)].” Janevski, 7:39-50.

133. To the extent audio information means the full content for playback, it would have been obvious to modify Janevski’s “initiator” “PVR” “114a” to transmit the “broadcast” to the “participant” “PVR” “114b” if it did not already have the broadcast recorded. Janevski, 6:13-38; 6:45-51, 6:59-7:3.

134. However, in my opinion, while Janevski teaches transmission of synchronization messages and audio content with timing information, Janevski does not disclose transmitting playback timing information indicative of a future time relative to a reference clock as taught in the '001 patent. But that specific type of timing information is disclosed by Okamura and would have been obvious to replace the timing information of Janevski.

135. In my opinion, Okamura offers a known improvement over reactive solutions by teaching transmission of *the generated* playback timing information based on a reference clock as taught in the '001 patent. Okamura explains that the transmitting node “generates a timestamp (abbreviated as ‘syt’) ... on the basis of the time provided by the cycle timer,” and that “the timestamp specifies the reproduction time at the receiving side of an event sequence.” Okamura, 1:42–56. In my opinion, this timestamp is transmitted along with the audio data previously obtained by the transmitting node (e.g., the first zone player), as explained for 1[i], to the receiving node, which uses it to determine when playback should occur.

136. Okamura further discloses that “on the transmitting side, the time of timestamp is set as the value of a reproduction time on the receiving side by estimating propagation delay.” Okamura, 24:7–12. In my opinion, this confirms that the playback timing information is generated and transmitted in advance, allowing

the receiving node to synchronize playback at a future time. Additionally, Okamura teaches that “by adjusting the offset value on the receiving side, the time of reproducing the audio data supplied from each transmitting node can be shifted from the time of timestamp.” Okamura, 24:7–12. This demonstrates that the transmitted timing information is used to control playback timing at the receiving node.

137. Thus, in my opinion, Okamura teaches transmitting both the obtained audio information as explained for 1[i] and the generated playback timing information via a network interface to a second zone player, satisfying the requirements of [1k].

- n. **[11] “wherein, while operating in the audio-slave mode for the synchrony group, the first zone player is configured to: receive, via the network interface, audio information and playback timing information associated with the received audio information from another zone player; and”**

138. In my opinion, a POSITA would have found claim element [11] obvious over the combination of Janevski, Kawamura, and Okamura. First, Janevski teaches element [11] by disclosing that “the Internet 118 supplies the means 119 for communicating information between the PVRs 114a, b such that synchronization may be achieved.” Janevski, 6:44–51. Accordingly, messages sent in Janevski are sent “via the network interface” used to communicate over the Internet, as discussed above. *Id.* The initiator transmits a **status message** to each participant to maintain

synchronization. Janevski discloses that “the synchronization process includes the transmission of a status message ... to ensure that the PVRs 114a, b participating in a session remain synchronous” (synchrony group). Janevski, 7:39–50. In my opinion, this confirms that the participant PVRs (e.g., the first zone player in audio-slave mode) receives synchronization-related information via a network interface from another zone player (e.g., the initiator).

139. Janevski further discloses that its status message contains “the current mode of watching (e.g., normal play, fast forward, pause), an indication of the time into the program, and information characteristic of content of a digital bit stream” and discloses that the content “may be aural” (audio information). Janevski, 5:10–12, 7:39-50, 16:36–37.

140. In my opinion, while Janevski discloses “an indication of the time into the program,” Janevski does not teach playback timing information as taught by the ’001 patent. So, while Janevski does not expressly disclose *playback timing information* as its exemplary timing solution, in my opinion, a POSITA would have been motivated to look to other known solutions like Okamura.

141. In my opinion, Okamura discloses that using playback timing information associated with the obtained audio information that is indicative of at least one future time relative to a reference clock time that denotes a time at which

at least the first and second zone players are to engage in synchronous playback of a corresponding portion of the obtained audio information was another known solution that would have been obvious to a POSITA to implement.

142. In my opinion, Okamura offers a known improvement over reactive solutions by disclosing that “the timestamp specifies the reproduction time at the receiving side of an event sequence,” and that it is generated “on the basis of the time provided by the cycle timer.” Okamura, 1:42–56. This timestamp is transmitted along with the audio data and is used by the receiving node to determine when playback should occur. Okamura discloses that “by adjusting the offset value on the receiving side, the time of reproducing the audio data supplied from each transmitting node can be shifted from the time of timestamp.” Okamura, 24:7–12. In my opinion, this demonstrates the playback timing information as taught in the ’001 patent.

143. In my opinion, a POSITA in the field of networked audio-visual systems would be well-acquainted with the use of network interfaces—such as Ethernet, Wi-Fi, or Bluetooth—for transmitting synchronization and playback control messages between devices. In my opinion, the skill level in the art includes not only familiarity with the hardware and protocols for network communication, but also an understanding of how to structure and transmit control information

(including playback timing) to ensure reliable and coordinated operation of multiple playback devices. In my opinion, given the widespread adoption of packet-based communication and timestamping in consumer and professional AV systems, a POSITA would readily recognize that transmitting playback timing information via a network interface is a routine and expected design choice for achieving synchronized playback in a multi-zone environment.

144. Moreover, in my opinion, a POSITA would appreciate the technical advantages of combining Janevski's approach to network-based synchronization with Okamura's predictive timestamping and offset adjustment techniques. By leveraging Okamura's method of packaging playback timing information with audio data and transmitting it over the network, in my opinion, a POSITA would expect to achieve more precise and scalable synchronization across devices, overcoming the limitations of reactive synchronization and manual time alignment. In my opinion, the knowledge and experience in the art would enable a POSITA to implement these improvements using standard components and protocols, with a reasonable expectation of success and predictable results in multi-zone audio playback systems.

145. Thus, in my opinion, Janevski's "indication of the time into the program" would be modified to include Okamura's *playback timing information indicative of a future time relative to a reference clock*, as required by the '001

patent. In short, Janevski discloses receiving information, related to playback timing; Okamura defines the specific content of the playback timing information received.

- o. [1m] “engage in synchronous playback of the received audio information with at least the second zone player based on the received playback timing information associated with the received audio information while a local clock time of the first zone player differs from a local clock time of the second zone player.”**

146. In my opinion, a POSITA would have found claim element [1m] obvious over the combination of Janevski, Kawamura, and Okamura. Janevski teaches engaging in synchronous playback of audio information between zone players. Specifically, Janevski discloses that “the initiator directs all participants ... to synchronize their playbacks to that of the initiator, whereby all playbacks are synchronized and maintain in synchronization.” Janevski, 7:39–50. The synchronization process includes transmission of a status message containing “an indication of the time into the program, and information characteristic of content of a digital bit stream.” *Id.* Because Janevski defines “renderable content” as “content that is presentable in a form that a user can sense, e.g. visually or aurally,” and confirms that “the presentation may be merely visual or merely aural or both visual and aural,” this necessarily includes audio information. Janevski, 5:10–12, 16:36–37. In my opinion, Janevski does not disclose that the zone players operate with different local clock times or how synchronization is achieved in such a scenario.

So, while Janevski does not disclose *local clock time* as its exemplary timing solution, in my opinion, a POSITA would have been motivated to look to other known solutions because, a POSITA would have been motivated to look to other known solutions because persistent clock drift, hardware differences, and asynchronous timers are well-recognized challenges in distributed playback systems, and industry practice as well as technical literature reflect the need for more robust synchronization mechanisms that can account for differing local clocks to ensure reliable and accurate synchronous playback across devices.

147. In my opinion, Okamura discloses that local clocks for the transmitting and receiving node was another known solution that would have been obvious to a POSITA to implement. In my opinion, Okamura offers a known improvement by explicitly teaching synchronization across devices with differing local clocks. In Okamura's architecture, the transmitting and receiving nodes operate on different clocks. The transmitting node functions as the master, generating timestamps based on a cycle timer, which serves as the system's reference clock. Specifically, Okamura discloses that "each of the transmitting nodes... generates a timestamp (abbreviated as 'syt'), one in every 8 sampling clocks (or 8 data blocks), on the basis of the time provided by the cycle timer." Okamura, 1:42–56. Okamura's receiving node acts as the slave, using its own internal cycle timer as a local clock to determine

when to play back the audio data. Okamura describes “a reproduction time control section that operates when the time of the timestamp added with the offset time coincides with a current time indicated by an internal cycle timer for controlling the data output section to effect synchronous reproduction.” Okamura, 5:53–67. In my opinion, these disclosures confirm that Okamura’s transmitting and receiving nodes operate on separate clocks with one serving as the reference and the other as the local playback clock. Playback is initiated when “the time of the timestamp added with the offset time coincides with a current time indicated by an internal cycle timer.” *Id.* Okamura further discloses that “by adjusting the offset value on the receiving side, the time of reproducing the audio data supplied from each transmitting node can be shifted from the time of timestamp,” allowing synchronization even when local clocks differ. Okamura, 24:7–12.

148. In my opinion, as described in element [11], the first zone player (in audio-slave mode) receives, via the network interface, both the audio information and the playback timing information from another zone player. *See* element [11]. The playback timing information is associated with the received audio information, as Okamura teaches “each transmitting node... arranges audio data of one or more channels into a data field and arranges the associated timestamp into a syt field so as to form a packet composed of the data field and the syt field.” Okamura, 1:42–56.

In my opinion, this confirms that the audio and timing information are received together and used to initiate playback at a future time.

149. Thus, in my opinion, Okamura teaches engaging in synchronous playback based on received timing information while local clock times differ, satisfying the remainder of [1m].

150. In my opinion, a POSITA would have been motivated to combine Janevski and Okamura to enable more robust and accurate synchronization across zone players with differing internal clocks. Janevski provides a framework for synchronized playback between devices but assumes that timing differences can be detected and corrected reactively using symmetrical message exchanges and internal video timers. In my opinion, this approach does not account for persistent clock drift, asynchronous local timers, network delays, or playback glitches which can degrade synchronization quality over time.

151. In my opinion, Okamura addresses this limitation by introducing a reference clock (cycle timer) and offset-based synchronization, allowing devices with different local clocks to align playback precisely. In my opinion, a POSITA would have recognized that Okamura's predictive synchronization mechanism could be advantageously implemented in Janevski to improve timing accuracy and reduce reliance on reactive correction.

152. In my opinion, a POSITA familiar with networked audio/video playback systems and synchronization protocols would have recognized the limitations of reactive synchronization methods like those in Janevski, which rely on symmetrical message exchanges and internal timers but do not account for persistent clock drift or asynchronous local clocks. Okamura addresses these limitations by introducing a predictive synchronization mechanism using a reference clock (cycle timer) and offset-based timing, allowing devices with different local clocks to align playback precisely. In my opinion, a POSITA would have been motivated to combine Janevski's multi-device playback framework with Okamura's reference-clock-based synchronization because doing so would enable more robust and accurate synchronous playback across zone players, even when their internal clocks differ. In my opinion, this combination would use known techniques, reference clocks, offset adjustments, and timestamp scheduling, according to their established functions, to achieve predictable and scalable results in networked environments, thereby improving user experience and system reliability.

153. Both references are directed to synchronizing playback across AV devices in networked environments and use similar components, e.g., media players, control messages, and timing mechanisms, to achieve similar goals. See Janevski, 6:4–44; Okamura, 1:42–56, 5:53–67. In my opinion, a POSITA would have

reasonably expected success in combining Okamura’s reference-clock-based synchronization with Janevski’s multi-device playback framework to achieve predictable and scalable results.

2. Claims 12 and 23

154. In my opinion, independent claims 12 and 23 are unpatentable for the same reasons as claim 1 because they simply repeat the elements of claim 1 and one additional unpatentable element. The table below highlights the language in claims 12 and 23 that is copied from or substantively identical to claim 1.

Elements of claim 1	Elements of claims 12 and 23
[1a]	a network interface that is configured to communicatively couple the first zone player to at least one data network; receiving, via a network interface at the first zone player, a request to engage in synchronous playback of audio content as part of a synchrony group that includes at least a second zone player that is communicatively coupled to the first zone player via at least one data network;
[1b]	after receiving the request to enter into the synchrony group:
[1ci]-[1cii]	detecting an indication that the first zone player is to operate in (a) one of a control-master mode or a control-slave mode for the synchrony group and (b) one of an audio-master mode or an audio-slave mode for the synchrony group; and
[1d]	beginning to operate in the synchrony group in accordance with the indication;
[1e]	wherein, while operating in the control-master mode for the synchrony group, the first zone player is configured to: receive, via the network interface, first control information for the

	synchrony group from a network device that is communicatively coupled to the first zone player; and
[1f]	based on the first control information, cause, via the network interface, at least one playback action to be applied in the synchrony group;
[1g]	wherein, while operating in the control-slave mode for the synchrony group, the first zone player is configured to: receive, via the network interface, second control information from another zone player; and
[1h]	perform one or more playback actions in accordance with the second control information;
[1i]	wherein, while operating in the audio-master mode for the synchrony group, the first zone player is configured to: obtain audio information that is representative of the audio content;
[1j]	generate playback timing information associated with the obtained audio information that is indicative of at least one future time relative to a reference clock time that denotes a time at which at least the first and second zone players are to engage in synchronous playback of a corresponding portion of the obtained audio information; and
[1k]	transmit, via the network interface, the obtained audio information and the generated playback timing information to the second zone player; and
[1l]	wherein, while operating in the audio-slave mode for the synchrony group, the first zone player is configured to: receive, via the network interface, audio information and playback timing information associated with the received audio information from another zone player; and
[1m]	engage in synchronous playback of the received audio information with at least the second zone player based on the received playback timing information associated with the received audio

	information while a local clock time of the first zone player differs from a local clock time of the second zone player.
--	--------------------------------------------------------------------------------------------------------------------------

155. In my opinion, Claims 12 and 23 only differ from claim 1 in their preambles.

156. Claim 12 further recites “at least one processor; a tangible, non-transitory computer-readable medium; and program instructions stored on the tangible, non-transitory computer-readable medium that are executable by the at least one processor such that the first zone player is configured to perform functions comprising.” Janevski discloses “a processor ... in either the PVRs 114a, b or one of the other devices associated with system 110.” Janevski, 7:20-24. Janevski also discloses these limitations because “PVR 114a includes ... a digital memory 206,” which is a non-transitory computer-readable media, as shown in Figure 2. Janevski, 7:57-8:7. Janevski further discloses that its memory includes instructions that are executed because “[a] memory device and a processor preferably reside in either the PVRs 114a, b” and “programming code associated with the system 110 preferably resides in the memory device and is processed by the processor.” Janevski, 7:20-24;

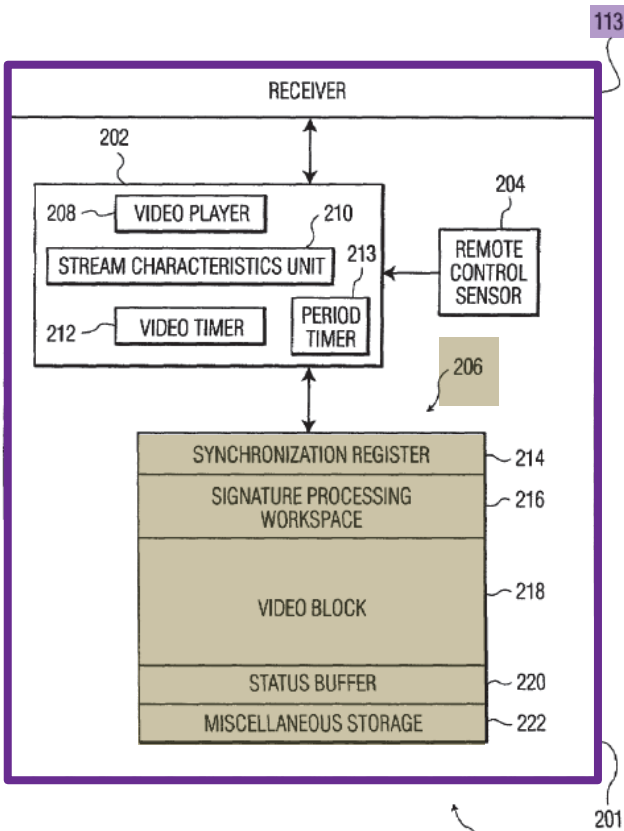


FIG. 2

Janevski, FIG. 2.

157. Claim 23 only differs from claim 1 and includes an element from claim 12 in its preamble because the “[t]angible, non-transitory computer-readable media having instructions stored therein, wherein the instructions, when executed, cause a first zone player to perform functions comprising:” are recited in claim 12’s elements as discussed above.

E. Dependent Claims

158. In my opinion, a POSITA would have found dependent claims 13, 17-22, 24, and 28-33 obvious over the combination of Janevski, Kawamura, and Okamura.

1. **Claims 2, 13, and 24: “wherein detecting an indication that the first zone player is to operate in (a) one of a control-master mode or a control-slave mode for the synchrony group and (b) one of an audio-master mode or an audio-slave mode for the synchrony group comprises detecting an indication that the first zone player is to operate in (a) the control-master mode for the synchrony group and (b) the audio-master mode for the synchrony group.”**

159. Janevski teaches the additional elements of claims 2, 13, and 24 because it discloses that when a PVR “performs a control function (e.g., start, pause, rewind, fast forward),” that PVR “broadcasts a command to all participants ... and ... becomes the new initiator.” Janevski, 7:31–50; 8:48–52. In my opinion, this execution of a control function serves as the indication that the zone player is to operate in the control-master mode, as it triggers the role transition and initiates synchronization across the group.

160. In my opinion, Janevski also teaches that the same indication reflects operation in the audio-master mode. Specifically, Janevski discloses that “the status message [also] includes ... an indication of the time into the program, and information characteristic of content of a digital bit stream from which playback to

the message sender is being generated.” Janevski, 7:36–41. Janevski also discloses “synchronizing presentation of renderable content of two or more digital bit streams” which “may be ... aural.” Janevski, 6:16–27, 16:34–43; see also *id.*, claim 3 (“said streams include audio data and said presentation is at least partially aural.”). Because Janevski defines “renderable content” as “content that is presentable in a form that a user can sense, e.g. visually or aurally,” and confirms that “the presentation may be merely visual or merely aural or both visual and aural,” this necessarily includes audio information. Janevski, 5:10–12, 16:36–37. In my opinion, the fact that the initiator is generating playback from a digital bit stream and transmitting timing and content information to other zone players confirms that it is functioning as the audio-master. Thus, in my opinion, Janevski teaches that the execution of a control function, and the resulting transmission of a status message, together serve as the detected indication that the zone player is to operate in both control-master and audio-master modes.

2. **Claims 6, 17, and 28: “wherein beginning to operate in the synchrony group in accordance with the indication comprises either (a) transitioning from operating in the audio-master mode to operating in the audio-slave mode or (b) transitioning from operating in the audio-slave mode to operating in the audio-master mode.”**

161. In my opinion, Janevski teaches the additional elements of claims 6, 17, and 28 because it discloses that that “each time that a participant changes the

watching mode i.e. executes a control function (rewind, fast forward, etc.), that participant broadcasts a command to all other participants (including the ‘current’ initiator) and thereby becomes the new initiator.” Janevski, 8:42–52. In my opinion, this transition reflects a change from audio-slave mode to audio-master mode, as the participant assumes control over synchronized playback.

162. Conversely, Janevski discloses that “the initiator directs all participants ... to synchronize their playbacks to that of the initiator,” and that “the synchronization process includes the transmission of a status message which includes “information characteristic of content of a digital bit stream from which playback to the message sender is being generated.” Janevski, 7:39–50. Because the initiator is the source of playback timing and content, it operates in audio-master mode, and participants operate in audio-slave mode. Thus, in my opinion, Janevski teaches that beginning to operate in the synchrony group includes transitioning between audio-master and audio-slave modes, satisfying the requirements of Claims 6, 17, and 28.

3. **Claims 7, 18 and 29: “wherein the first control information identifies particular audio content to be played back by the synchrony group that is available at an audio source outside of the at least one data network, and wherein causing the at least one playback action to be applied in the synchrony group comprises causing a zone player operating in the audio-master mode to obtain audio information that is representative of the particular audio content.”**

163. In my opinion, a POSITA would have found claims 7, 18, and 29 obvious over Janevski.

164. Figure 1 illustrates the “synchronized viewing session,” where “[t]he Internet 118 [the claimed *data network*] supplies the means 119 for communicating information between the PVRs 114a, b such that synchronization may be achieved.” Janevski, 6:48-51.

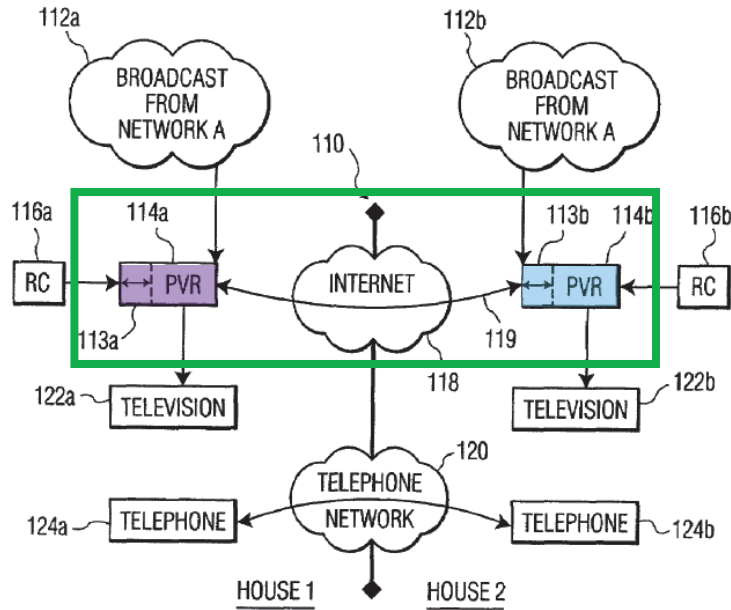


FIG. 1

Janevski, FIG. 1.

165. FIG. 1 also discloses that “broadcasts **112a, b** of a television program from network A is made pursuant to any communication means known to one having ordinary skill in the art, such as cable, digital cable, satellite, antenna, over the Internet or combinations thereof.” Janevski, FIG. 1, 6:6-10. In my opinion, since network A (112a) is outside the at least one data network (the Internet 118), Janevski teaches that particular audio content to be played back by the synchrony group can be available at an audio source outside of the data network used for synchronization. *Id.* Specifically, in my opinion, broadcasts from network A – including cable, digital cable, satellite, antenna, or other sources – are external to the Internet-based data

network that supplies synchronization information between the PVRs. Janevski, 6:6-39, 6:48-51; FIG. 1.

166. Thus, in my opinion, the first control information in Janevski's system can identify particular audio content, e.g., a television broadcast or audio stream, available from network A, which is outside the Internet data network used for device-to-device communication and synchronization. Janevski, 6:6-39, FIG. 1.

167. Thus, in my opinion, Janevski teaches that the zone player can obtain and play back audio information from an external source, network A (112a), satisfying the requirements of claims 7, 18, and 29.

4. Claims 8, 19, and 30: “wherein the at least one future time relative to the reference clock time comprise at least one first future time that is determined based on a local clock of a zone player other than the first zone player.”

168. In my opinion, the combination of Janevski and Okamura teaches the additional elements of claims 8, 19, and 30 because it discloses that the transmitting node “generates a timestamp (abbreviated as ‘syt’), one in every 8 sampling clocks (or 8 data blocks), on the basis of the time provided by the cycle timer... The timestamp specifies the reproduction time at the receiving side of an event sequence.” Okamura, 1:42–56’ The “cycle timer” serves as the system's reference clock. See also '001, 2:49–67.

169. In my opinion, the receiving node (the second zone player) uses its own internal cycle timer, i.e., a local clock, to determine when to initiate playback. Okamura discloses “a reproduction time control section that operates when the time of the timestamp added with the offset time coincides with a current time indicated by an internal cycle timer.” Okamura, 5:53–67. In my opinion, this confirms that the playback time is determined based on the local clock of the receiving node, which is a zone player other than the first.

170. Okamura further explains that “by adjusting the offset value on the receiving side, the time of reproducing the audio data supplied from each transmitting node can be shifted from the time of timestamp.” Okamura, 24:7–12. In my opinion, this adjustment is based on the receiving node’s local clock and ensures synchronized playback. Thus, in my opinion, Okamura teaches that the future time for playback, relative to the reference clock, is determined based on the local clock of another zone player, satisfying Claims 8, 19, and 30.

5. Claims 9, 20 and 31: “wherein the at least one future time relative to the reference clock time comprise at least one first future time that is determined based on a local clock of the first zone player.”

171. In my opinion, Okamura teaches the additional elements of claims 9, 20, and 31 because it discloses that the transmitting node “generates a timestamp (abbreviated as ‘syt’) ... on the basis of the time provided by the cycle timer.”

Okamura, 1:42–56. The “cycle timer” serves as the system’s reference clock. *Id.* The timestamp “specifies the reproduction time at the receiving side of an event sequence.” *Id.*

172. Okamura further discloses “an offset setting section that sets an offset time for the receiving node relative to the transmitting node and adds the offset time to a time indicated by the timestamp.” Okamura, 5:53–67. The receiving node uses its own internal cycle timer, i.e., its local clock, to determine when playback should occur, because in this context, the first zone player is operating in audio-slave mode, and the future playback time is determined based on its own local clock. Okamura describes “a reproduction time control section that operates when the time of the timestamp added with the offset time coincides with a current time indicated by an internal cycle timer.” Okamura, 5:53–67.

173. Thus, opposite to claims 8, 19, and 30 where the first zone player is the transmitted node, here in claims 9, 20, and 31, the future time for playback is determined based on the local clock of the receiving node, which in this case corresponds to the first zone player operating in audio-slave mode. Okamura confirms that “by adjusting the offset value on the receiving side, the time of reproducing the audio data supplied from each transmitting node can be shifted from the time of timestamp.” Okamura, 24:7–12. This adjustment is based on the local

clock of the receiving zone player, satisfying the requirements of Claims 9, 20, and 31.

- 6. Claims 10, 21, and 32: “wherein the second control information comprises information indicative of a volume adjustment, and wherein performing one or more playback actions in accordance with the second control information comprises adjusting a playback volume of the first zone player.”**

174. In my opinion, a POSITA would have found claims 10, 21, and 32 obvious over the combination of Janevski, Kawamura, and Okamura. Janevski teaches control functions such as “start, pause, rewind, fast forward,” and synchronization of playback across devices, but does not disclose volume adjustment. *See* Janevski, 6:18–22; 8:42–52.

175. Okamura mentions volume only in the context of preset playback conditions by disclosing that “in step S81, the audio data are reproduced and outputted to the reproduction channels (audio0 through audio7) under conditions such as volume, effect, etc. preset to these channels.” Okamura, 21:1-13. However, Okamura also does not disclose *information indicative of a volume adjustment* nor *adjusting a playback volume of the first zone player*.

176. In my opinion, Kawamura teaches volume adjustment as another playback function. Kawamura discloses that “the remote commander 13 remotely controls the media player 11 for executing play control such as play, stop, pause, fast

forward and rewind and remotely controls the headphone 12 for executing sound volume adjustment (volume up/down) for example.” Kawamura, [0088]. The headphone control block 12B realizes sound output functions such as volume up, volume down, and mute. Kawamura, [0097]. The remote commander issues volume adjustment commands, which are transferred by the media player to the headphone for execution. Kawamura, [0118].

177. In my opinion, a POSITA would have recognized that volume adjustment is a standard playback function, just like play, pause, or fast forward, and that Bluetooth protocols as discussed in Kawamura support transmission of all such playback commands, including volume adjustment, between devices. In my opinion, a POSITA would have been motivated to combine Kawamura’s volume adjustment playback function in addition to Janevski and Okamura’s playback functions to implement a standard set of playback commands (including volume) in a synchronized playback system, using Bluetooth or similar protocols to transmit all relevant commands as needed.

178. In my opinion, as Bluetooth protocols (as discussed in Kawamura) are designed to transmit a standard set of playback commands, including volume adjustment, between devices, a POSITA would have known that integrating volume control into a synchronized multi-zone playback system is both technically

straightforward and highly desirable for user experience. The ability to remotely adjust playback volume in real time allows users to tailor audio output to their preferences and environment, which is especially important in multi-zone or distributed audio systems. Furthermore, in my opinion, a POSITA would have known that the mechanisms for transmitting control information, such as volume commands, are well-established and compatible with the synchronization frameworks described in Janevski and Okamura. In my opinion, the combination of these references would result in a robust system where playback actions, including volume adjustment, can be coordinated across multiple devices, providing users with granular control and predictable, reliable performance. Thus, in my opinion, the combination of Janevski, Kawamura, and Okamura would have used known techniques (volume control via remote input) in known systems (multi-device synchronized playback), to achieve predictable results (volume adjustment during synchronized playback). In my opinion, a POSITA would have reasonably expected success in combining Kawamura's volume control mechanism with Janevski and Okamura's synchronization framework to improve user experience and provide more granular playback control across devices.

7. Claims 11, 22 and 33

- a. 11, 22, and 33 (pre): “wherein the playback timing information that is received while operating in the audio-slave mode comprises at least one future time at which at least the first and second zone players are to engage in synchronous playback of a corresponding portion of the received audio information, and wherein operating in the audio-slave mode to engage in synchronous playback of the received audio information with at least the second zone player comprises:”**

179. In my opinion, a POSITA would have found claim elements 11(pre), 22 (pre), and 33 (pre) obvious over the combination of Janevski, Kawamura, and Okamura as explained in [1j], [11], [1m], and as further specified in dependent claims 8, 19, and 30 and 9, 20, and 31.

- b. 11, 22, and 33 (a): “updating the at least one future time to account for a differential between the local clock time of the first zone player and a local clock time of another zone player; and”**

180. In my opinion, Janevski and Okamura teach claim elements 11(a), 22 (a), and 33 (a), because Okamura discloses that there is “an offset setting section that sets an offset time for the receiving node relative to the transmitting node and adds the offset time to a time indicated by the timestamp.” Okamura, 5:53–67. This offset time is used to adjust the received timestamp, which represents the future playback time, so that playback can be synchronized even when the local clocks of the zone players differ. In my opinion, Okamura teaches that the receiving node (first zone

player) maintains its own local clock (“internal cycle timer”) and receives timestamps generated by the transmitting node (another zone player) based on its local clock. In my opinion, the receiving node updates the future playback time by applying an offset to the received timestamp, thereby accounting for the differential between the local clock time of the first zone player and the local clock time of another zone player.

181. Okamura further explains, “by adjusting the offset value on the receiving side, the time of reproducing the audio data supplied from each transmitting node can be shifted from the time of timestamp.” Okamura, 24:7–12. In my opinion, this adjustment ensures that the future time for synchronous playback is updated to account for the differential between the local clock time of the first zone player and the local clock time of another zone player, satisfying the requirements of claims 11(a), 22(a), and 33(a).

- c. **11, 22, and 33 (b): “when the local clock time of the first zone player reaches the updated at least one future time, engaging in synchronous playback of the corresponding portion of the received audio information with at least the second zone player.”**

182. In my opinion, Janevski and Okamura teach claim elements 11(b), 22(b) and 33 (b), because Okamura discloses “a reproduction time control section that operates when the time of the timestamp added with the offset time coincides with a

current time indicated by an internal cycle timer for controlling the data output section to effect synchronous reproduction.” Okamura, 5:53–67.

183. Okamura discloses that, “the timestamp specifies the reproduction time at the receiving side of an event sequence.” Okamura, 1:42–56. Playback is initiated when “the time of the timestamp added with the offset time coincides with a current time indicated by an internal cycle timer,” ensuring that synchronous playback occurs at the correct time. Okamura, 5:53–67.

XI. Ground 2: Claims 3, 14, and 25 Are Rendered Obvious by Janevski, Kawamura, Okamura, and Kono.

A. Kono (Ex. 1008)

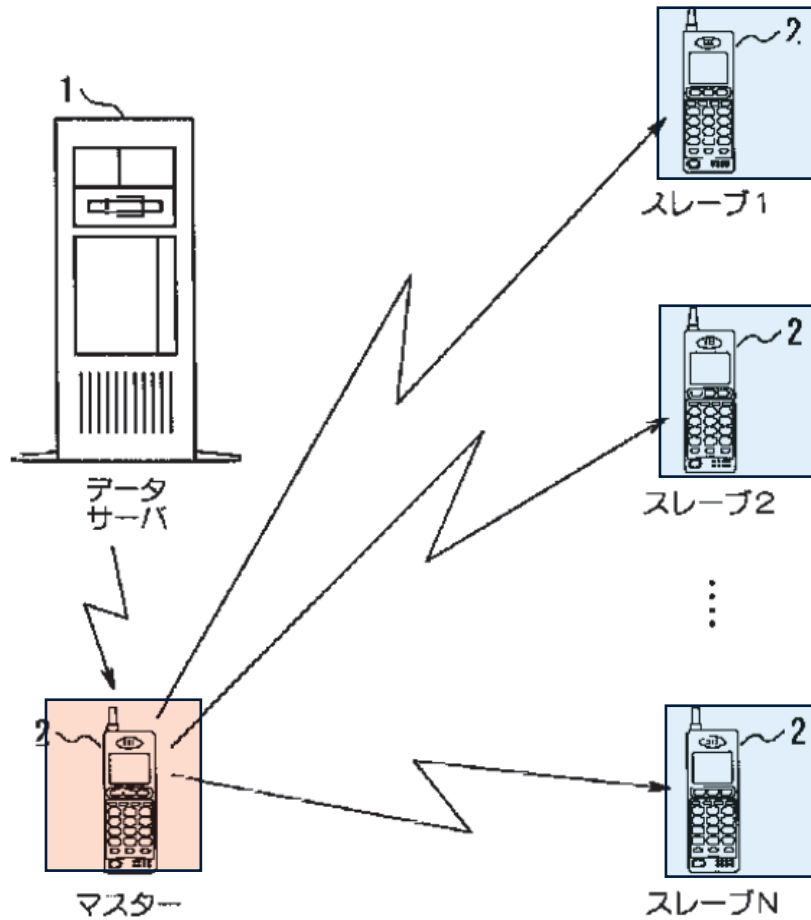
184. Kono “relates to a karaoke system that employs a plurality of mobile communication terminals, such as mobile phone devices.” Kono, [0001]. The system enables “a plurality of mobile communication terminals ...to play karaoke” using their mobile phones overcoming limitations of prior systems where only one device would participate. Kono, [0001], [0003], [0004].

185. Kono’s system includes a “master mobile terminal” and one or more “slave mobile phone devices,” each equipped with a wireless communication module such as Bluetooth. Kono, [0011]. The master “accesses a site on the server 1 via the Internet, selects song data, and downloads the song data by a prescribed

procedure.” Kono, [0012]. The downloaded song data includes “timing information, which indicates a playback timing for a song...is embedded in the song data.” *Id.*

186. After downloading, “the song data downloaded by the master is transferred and distributed to each of the slaves 1–N via the wireless communication modules.” Kono, [0015]. Playback is initiated when “[o]n the **master** side,...the playback start button is pressed,” which causes “a start notification packet [to be] sent to the **slave**,” and “playback of audio data, subtitle data, and image data is started.” Kono, [0021]. On the **slave** side, “when the aforementioned start notification packet is received ... playback of audio data, subtitle data, and the image data is started.” Kono, [0022].

【図1】



Kono, FIG. 1.

187. During playback, “the timing information for the song data on the master side is sent to each of the slaves via the wireless communication modules, and each of the slaves controls its own playback speed on the basis of that timing information.” Kono, [0013]. The slave checks “whether the timing at which the above-mentioned next block must be played back matches the timing indicated in

the above-mentioned timing information that was received ... [and] then playback of that block is started.” Kono, [0022]. This ensures synchronized playback across devices.

188. In my opinion, like the '001 patent, Kono is in the technical field of controlling synchronous audio playback across multiple devices. Compare Ex. 1001, 1:25–45, 2:35–48, 5:33–39, with Kono, [0001], [0005], [0011], [0012].

B. Claims 3, 14, and 25: “wherein the obtained audio information comprises a beginning of the obtained audio information, and wherein the playback timing further comprises a future time relative to the reference clock time that denotes a time at which at least the first and second zone players are to initiate synchronous playback at the beginning of the obtained audio information.”

189. In my opinion, a POSITA would have found claims 3, 14, and 25 obvious over the combination of Janevski, Kawamura, and Okamura with Kono. Janevski teaches claims 3, 14, and 25 because it discloses that “each message is timestamped based on the internal video timer of the respective PVRs at the moment of sending or receiving.” Janevski, 8:65–9:4. Janevski’s PVR synchronized viewing system collects four time points: “the message 402 is sent at a time A ... arrives at the participant ... at a time B ... the participant sends back a reply ... at a time C ... received by initiator ... at a time D.” Janevski, 9:15-35. The time misregistration (TM) is calculated using the formula: $TM = \frac{1}{2}[(A + D) - (C + B)]$.” Janevski, 9:30-35. Janevski explains, “once the TM value is calculated, the participant PVR adjusts

its timer forward or backward by that amount to match the initiator's playback time.” Janevski, 9:11–13. In my opinion, this confirms that playback synchronization is achieved across devices.

190. Janevski further teaches that the initiator controls playback timing for the group by disclosing that “the initiator directs all participants ... to synchronize their playbacks to that of the initiator, whereby all playbacks are synchronized and maintain in synchronization.” Janevski, 7:39–50. In my opinion, this synchronization is based on control functions and timing alignment, but Janevski does not disclose that playback begins at the start of the audio content or that a reference clock is used to schedule a future playback time.

191. In my opinion, Okamura explicitly teaches *a future time relative to the reference clock time*. Okamura discloses that “the transmitting node generates a timestamp (abbreviated as ‘syt’), one in every 8 sampling clocks (or 8 data blocks), on the basis of the time provided by the cycle timer... The timestamp specifies the reproduction time at the receiving side of an event sequence.” Okamura, 1:42–56. The cycle timer serves as the reference clock, and the timestamp represents a future time for playback.

192. Okamura further explains that “an offset setting section ... sets an offset time for the receiving node relative to the transmitting node and adds the offset time

to a time indicated by the timestamp.” Okamura, 5:53–67. Playback is initiated when “the time of the timestamp added with the offset time coincides with a current time indicated by an internal cycle timer for controlling the data output section to effect synchronous reproduction.” Okamura, 5:53–67. In my opinion, this confirms that playback timing is scheduled for a future time relative to a reference clock.

193. However, neither Janevski nor Okamura explicitly disclose *initiat[ing] synchronous playback at the beginning of the obtained audio information*. In my opinion, Kono offers a known improvement by teaching that the master mobile terminal downloads song data from a server and initiates playback from the beginning. Kono states, “first, the master accesses a site on the server 1 via the Internet, selects song data, and downloads the song data by a prescribed procedure.” Kono, [0012].

194. Kono further explains that when “[o]n the master side, ... the playback start button is pressed,” which causes “a start notification packet [to be] sent to the slaves,” and “playback of audio data, subtitle data, and image data is started.” Kono, [0021]. On the slave side, “when the aforementioned start notification packet is received ... playback of the audio data, subtitle data, and the image data is started.” Kono, [0022]. In my opinion, this confirms that playback is initiated at the beginning of the audio content.

195. Kono also teaches synchronization based on timing information embedded in the audio data. It discloses that “each of these data is partitioned into N blocks, each having a prescribed data size, and timing information is written per block for all the blocks.” Kono, [0014]. During playback, “the timing information for the song data on the master side is sent to each of the slaves via the wireless communication modules, and each of the slaves controls its own playback speed on the basis of that timing information.” Kono, [0013]. The slave checks “whether the timing at which the above-mentioned next block must be played back matches the timing indicated in the above-mentioned timing information that was received ... [and] then playback of that block is started.” Kono, [0022]. In my opinion, this ensures that playback across devices is synchronized from the beginning of the audio stream.

196. In my opinion, a POSITA would have been motivated to combine Kono with the Janevski–Okamura combination to implement a synchronized playback system that initiates playback at the beginning of the audio content. Janevski teaches reactive synchronization using internal video timers and message exchanges to align playback timing across devices. Janevski, FIG. 4; 8:65–9:35. In my opinion, Okamura improves upon this by introducing predictive synchronization using a reference clock and timestamp-based scheduling. Okamura, 1:42–56, 5:53–67.

However, in my opinion, neither reference discloses initiating playback specifically at the beginning of the audio stream.

197. In my opinion, Kono offers a known improvement by teaching that the master mobile terminal downloads song data from a server and initiates playback from the beginning of the audio content. Kono discloses that “[f]irst, the master accesses a site on the server 1 via the Internet, selects song data, and downloads the song data by a prescribed procedure.” Kono, [0012]. Playback is initiated when “[o]n the master side, ... the playback start button is pressed,” which causes “a start notification packet [to be] sent to the slaves,” and “playback of audio data, subtitle data, and image data is started.” Kono, [0021]. On the slave side, “when the aforementioned start notification packet is received ... playback of the audio data, subtitle data, and the image data is started.” Kono, [0022]. Kono confirms that playback begins at the start of the obtained audio information and is synchronized across devices using timing information embedded in each block. Kono, [0021].

198. In my opinion, a POSITA would have recognized that combining Kono’s initiation mechanism with the Janevski–Okamura synchronization framework would yield a more complete and predictable multi-device playback system. In my opinion, the combination uses known techniques, such as start-of-stream playback, timestamp-based scheduling, and inter-device synchronization,

according to their known functions to achieve predictable results: synchronized playback beginning at the start of the audio content. By combining Kono's mechanism for starting playback from the beginning of downloaded content with Janevski's and Okamura's synchronization frameworks, in my opinion, a POSITA would have understood that users could reliably experience simultaneous playback of the same audio program across multiple devices, regardless of the source or network conditions. This would minimize user confusion, prevent playback drift, and ensure a consistent experience for all participants in a synchrony group. Furthermore, in my opinion, a POSITA would have known that the use of timing information embedded in the audio data, as taught by Kono and Okamura, is compatible with the control and synchronization methods described in Janevski. In my opinion, the combination leverages well-established techniques, such as start-of-stream playback, timestamp-based scheduling, and inter-device coordination, to achieve predictable, scalable, and user-friendly results.

199. Moreover, in my opinion, a POSITA would have reasonably expected success in combining Kono with the Janevski, Kawamura, and Okamura combination, because all three references operate in the domain of networked or distributed media playback and use compatible components such as timing controllers, synchronization messages, and playback coordination logic.

XII. CONCLUSION

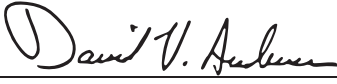
200. For the reasons set forth in Section X and XI, it is my opinion that one skilled in the art would have found claims 1-3, 6-14, 17-25, 28-33 of the '001 patent to be obvious.

201. In signing this declaration, I understand that the declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I acknowledge that I may be subject to cross-examination in this case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

202. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Date: October 14, 2025

Respectfully submitted,



Dr. David Anderson