



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
14/383,341	09/05/2014	Siu Wa Lee	3145.067US1	3561
21186	7590	09/06/2018	EXAMINER	
SCHWEGMAN LUNDBERG & WOESSNER, P.A. P.O. BOX 2938 MINNEAPOLIS, MN 55402 UNITED STATES OF AMERICA			KIM, JONATHAN C	
			ART UNIT	PAPER NUMBER
			2659	
			NOTIFICATION DATE	DELIVERY MODE
			09/06/2018	ELECTRONIC

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

uspto@slwip.com  
SLW@blackhillsip.com

UNITED STATES PATENT AND TRADEMARK OFFICE

---

BEFORE THE PATENT TRIAL AND APPEAL BOARD

---

*Ex parte* SIU WA LEE, LING CEN, HAIZHOU LI,  
YAOZHU PAUL CHAN, and MINGHUI DONG

---

Appeal 2018-000374  
Application 14/383,341<sup>1</sup>  
Technology Center 2600

---

Before LARRY J. HUME, JAMES W. DEJMEK, and  
JOHN D. HAMANN, *Administrative Patent Judges*.

DEJMEK, *Administrative Patent Judge*.

DECISION ON APPEAL

Appellants appeal under 35 U.S.C. § 134(a) from a Final Rejection of claims 1–19. We have jurisdiction over the pending claims under 35 U.S.C. § 6(b).

We affirm.

---

<sup>1</sup> Appellants identify the Agency for Science, Technology and Research as the real party in interest. Br. 3.

## STATEMENT OF THE CASE

### *Introduction*

Appellants' disclosed and claimed invention generally relates to "template-based personalized singing synthesis." Spec. ¶ 2. In particular, the disclosed approach uses speech-to-singing (STS) synthesis. Spec. ¶ 4. According to the Specification, in STS synthesis, the synthesized singing of a speaker preserves the timbre of the speaker such that the synthesized singing will sound like it is being sung by the speaker. Spec. ¶ 6. In a disclosed embodiment, characteristics of a singing voice for a first individual (e.g., a professional singer) are derived and the vocal characteristics of a voice of a second individual (i.e., a speaker/unprofessional singer) are modified in response to the characteristics of the first individual's singing voice to generate a synthesized singing voice for the second individual. Spec. ¶ 9.

Claim 1 is representative of the subject matter on appeal and is reproduced below with the disputed limitation emphasized in *italics*:

1. A method for speech-to-singing synthesis comprising:  
deriving characteristics of a singing voice for a first individual; and

modifying vocal characteristics of a voice for a second individual in response to the characteristics of the singing voice of the first individual to generate a synthesized singing voice for the second individual,

*wherein modifying the vocal characteristics of the voice of the second individual comprises aligning the voice of the second individual to match the singing voice of the first individual.*

*The Examiner's Rejections*

1. Claims 1–19 stand rejected under 35 U.S.C. § 101 as being directed to patent-ineligible subject matter. Final Act. 8.
2. Claims 1, 3–6, and 8 stand rejected under 35 U.S.C. § 102(a) as being anticipated by Kawashima et al. (US 6,836,761 B1; Dec. 28, 2004) (“Kawashima”). Final Act. 9–12.
3. Claim 19 stands rejected under 35 U.S.C. § 102(a) as being anticipated by Haupt et al. (US 2013/0019738 A1; Jan. 24, 2013 (filed July 22, 2011)) (“Haupt”). Final Act. 12–13.
4. Claim 2 stands rejected under 35 U.S.C. § 103(a) as unpatentable over Kawashima and Haupt. Final Act. 13–14.
5. Claim 7 stands rejected under 35 U.S.C. § 103(a) as unpatentable over Kawashima and Oytun Türk and Marc Schröder, Evaluation of Expressive Speech Synthesis With Voice Conversion and Copy Resynthesis Techniques, IEEE Transactions on Audio, Speech, and Language Processing, Vol. 18, No. 5, 965–73 (July 2010) (“Türk”). Final Act. 14–15.
6. Claims 9, 10, and 14–17 stand rejected under 35 U.S.C. § 103(a) as unpatentable over Haupt and Ling Cen et al., Segmentation of Speech Signals in Template-based Speech to Singing Conversion, APSIPA ASC 2011 Xi'an (2011) (“Cen”). Final Act. 15–18.
7. Claims 11–13 stand rejected under 35 U.S.C. § 103(a) as unpatentable over Haupt, Cen, and Türk. Final Act. 19–21.
8. Claim 18 stands rejected under 35 U.S.C. § 103(a) as unpatentable over Haupt, Cen, and Nakano et al. (US 2009/0306987 A1; Dec. 10, 2009) (“Nakano”). Final Act. 21–22.

## ANALYSIS<sup>2</sup>

### *Rejection under 35 U.S.C. § 101*

Appellants dispute the Examiner’s conclusion that the pending claims are directed to patent-ineligible subject matter. Br. 7–9. In particular, Appellants argue the claims involve converting digital signals from one form (i.e., speech) to another (i.e., synthesized singing) to achieve a technical result—“to create a professional-quality singing voice for poor singers using the timbre of the speaker’s voice while applying the melody of a professional voice.” Br. 8. Appellants assert the claims provide a major advancement in speech to singing technology. Br. 8. Additionally, Appellants analogize the pending claims to the claims at issue in *McRO, Inc. v. Bandai Namco Games America Inc.*, 837 F.3d 1299 (Fed. Cir. 2016). Br. 9. More specifically, Appellants argue the court in *McRO* concluded the claims in that case were not directed to an abstract idea but rather recited “‘perform[ing] a distinct process to automate a task previously performed by humans’ and ‘us[ing] a combined order of specific rules that renders information into a specific format that is then used and applied to create desired results.’” Br. 9. Appellants contend the pending claims similarly apply rules for speech-to-singing synthesis. Br. 9.

The Supreme Court’s two-step framework guides our analysis of patent eligibility under 35 U.S.C. § 101. *Alice Corp. Pty. v. CLS Bank Int’l*, 134 S. Ct. 2347, 2355 (2014). If a claim falls within one of the statutory categories of patent eligibility (i.e., a process, machine, manufacture, or

---

<sup>2</sup> Throughout this Decision, we have considered the Appeal Brief, filed June 5, 2017 (“Br.”); the Examiner’s Answer, mailed August 25, 2017 (“Ans.”); and the Final Office Action, mailed May 6, 2016 (“Final Act.”), from which this Appeal is taken. Appellants did not file a Reply Brief.

composition of matter) then the first inquiry is whether the claim is directed to one of the judicially recognized exceptions (i.e., a law of nature, a natural phenomenon, or an abstract idea). *Alice*, 134 S. Ct. at 2355. If so, the second step is to determine whether any element, or combination of elements, amounts to significantly more than the judicial exception. *Alice*, 134 S. Ct. at 2355.

Although the independent claims each broadly fall within the statutory categories of patentability, the Examiner concludes the claims are directed to a judicially recognized exception—i.e., an abstract idea. Final Act. 8. In particular, the Examiner concludes the claims are directed to “modifying vocal characteristics of a voice for a second individual.” Ans. 2; *see also* Final Act. 8. The Examiner explains such modification consists of a series of mathematical operations. Final Act. 8. Further, the Examiner finds the claims only require generic computer elements performing “generic computer functions that are well-understood, routine and conventional activities previously known to the industry” and concludes, therefore, the claims do not recite significantly more to transform the claims into patent-eligible subject matter. Final Act. 8.

Instead of using a definition of an abstract idea, “the decisional mechanism courts now apply is to examine earlier cases in which a similar or parallel descriptive nature can be seen—what prior cases were about, and which way they were decided.” *Amdocs (Isr.) Ltd. v. Openet Telecom, Inc.*, 841 F.3d 1288, 1294 (Fed. Cir. 2016) (citing *Elec. Power Grp., LLC v. Alstom S.A.*, 830 F.3d 1350, 1353–54 (Fed. Cir. 2016)); *accord* United States Patent and Trademark Office, July 2015 Update: Subject Matter Eligibility 3 (July 30, 2015), <https://www.uspto.gov/sites/default/files/>

documents/ieg-july-2015-update.pdf (instructing Examiners that “a claimed concept is not identified as an abstract idea unless it is similar to at least one concept that the courts have identified as an abstract idea.”). As part of this inquiry, we must “look at the ‘focus of the claimed advance over the prior art’ to determine if the claim’s ‘character as a whole’ is directed to excluded subject matter.” *Affinity Labs of Tex., LLC v. DirecTV, LLC*, 838 F.3d 1253, 1257 (Fed. Cir. 2016).

Our reviewing court has concluded that abstract ideas include the concepts such as the collection and analysis of information. *Elec. Power*, 830 F.3d at 1353; *see also Elec. Power*, 830 F.3d at 1354 (“analyzing information by steps people go through in their minds, or by mathematical algorithms, without more, as essentially mental processes within the abstract-idea category”). Similarly, “collecting, displaying, and manipulating data” is an abstract idea. *Intellectual Ventures I LLC v. Capital One Fin. Corp.*, 850 F.3d 1332, 1340 (Fed. Cir. 2017) (additionally concluding that “customizing information and presenting it to users based on particular characteristics” is an abstract idea (citing *Intellectual Ventures I LLC v. Capital One Bank (USA)*, 792 F.3d 1363, 1370 (Fed. Cir. 2015))). Further, our reviewing court has concluded that a process that employs mathematical algorithms to manipulate existing information to generate additional information is not patent eligible. *Digitech Image Techs., LLC v. Elec. for Imaging, Inc.*, 758 F.3d 1344, 1351 (Fed. Cir. 2014); *see also Parker v. Flook*, 437 U.S. 584, 595 (1978) (“If a claim is directed essentially to a method of calculating, using a mathematical formula, even if the solution is for a specific purpose, the claimed method is [patent ineligible].”) (internal quotations omitted).

Further, merely combining several abstract ideas does not render the combination any less abstract. *RecogniCorp, LLC v. Nintendo Co.*, 855 F.3d 1322, 1327 (Fed. Cir. 2017) (“Adding one abstract idea (math) to another abstract idea . . . does not render the claim non-abstract.”); *see also FairWarning IP, LLC v. Iatric Sys., Inc.*, 839 F.3d 1089, 1093–94 (Fed. Cir. 2016) (determining the pending claims were directed to a combination of abstract ideas).

Here, Appellants’ claims are directed to generating a synthesized singing voice by modifying vocal characteristics of one individual based on the characteristics derived from the singing voice of another individual. Deriving characteristics of a signal (e.g., a singing voice) is similar to the abstract idea of analyzing data. *See Elec. Power*, 830 F.3d at 1353–54; *Intellectual Ventures*, 850 F.3d at 1340. Similarly, modifying a signal (e.g., the vocal characteristics of another’s voice) based on the results of analyzing a first signal is also similar to the abstract idea of manipulating data using a mathematical relationship. *See Intellectual Ventures*, 850 F.3d at 1340; *Digitech*, 758 F.3d at 1351. Neither these steps, nor those recited in the dependent claims alters the character of the claims, which is directed to an abstract idea.

To the extent Appellants assert the pending claims are not directed to an abstract idea, but are similar to those held patent eligible in *McRO* (*see* Br. 9), we disagree. As our reviewing court noted in *McRO*, it is important to determine “whether the claims . . . focus on a specific means or method that improves the relevant technology or are instead directed to a result or effect that itself is the abstract idea and merely invoke generic processes and machinery.” 837 F.3d at 1314. In other words, a claim that defines the way

by which a computer-related result is achieved is distinguishable from a patent-ineligible claim that simply describes a result. Here, we agree with the Examiner that the claims recite at a high level deriving characteristics and modifying vocal characteristics in response to the derived characteristics, but do not otherwise recite how such derivation or modification is done. Ans. 5–6. Accordingly, we are unpersuaded of Examiner error.

Because we determine the claims are directed to an abstract idea or combination of abstract ideas, we analyze the claims under step two of *Alice* to determine if there are additional limitations that individually, or as an ordered combination, ensure the claims amount to “significantly more” than the abstract idea. *Alice*, 134 S. Ct. at 2355 (citing *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 566 U.S. 66, 72, 78–79 (2012)). The implementation of the abstract idea involved must be “more than [the] performance of ‘well-understood, routine, [and] conventional activities previously known to the industry.’” *Content Extraction and Transmission LLC v. Wells Fargo Bank N.A.*, 776 F.3d 1343, 1347–48 (Fed. Cir. 2014) (quoting *Alice*, 134 S. Ct. at 2359) (alteration in original). “Whether something is well-understood, routine, and conventional to a skilled artisan at the time of the patent is a factual determination.” *Berkheimer v. HP Inc.*, 881 F.3d 1360, 1369 (Fed. Cir. 2018).

Appellants do not dispute the Examiner’s findings that the claims recite “generic computer functions that are well-understood, routine and conventional activities.” See Final Act. 8. Appellants do not advance any arguments related to the Examiner’s findings and conclusions related to whether the claims recite significantly more to transform the abstract idea

into a patent-eligible application. Accordingly, we are not apprised of Examiner error.

For the reasons discussed *supra*, we sustain the Examiner’s rejection under 35 U.S.C. § 101 of independent claim 1. For similar reasons, we also sustain the Examiner’s rejection under 35 U.S.C. § 101 of independent claims 9 and 19, which recite similar limitations and were not argued separately. *See* Br. 9; *see also* 37 C.F.R. § 41.37(c)(1)(iv) (2017). Additionally, we sustain the Examiner’s rejection under 35 U.S.C. § 101 of claims 2–8 and 10–18, which depend directly or indirectly therefrom and were not argued separately. *See* 37 C.F.R. § 41.37(c)(1)(iv).

*Rejections under 35 U.S.C. § 102*

a. Kawashima (claims 1, 3–6, and 8)

Appellants contend Kawashima fails to disclose “modifying the vocal characteristics of the voice of the second individual comprises aligning the voice of the second individual to match the singing voice of the first individual,” as recited in claim 1. Br. 9–11. Instead, Appellants argue Kawashima discloses allocating vocal characteristics of a first individual to the vocal characteristics of a second individual, but does not align the voice of the second individual to match the singing voice of the first individual. Br. 10. Specifically, Appellants contend the alignment processing section in Kawashima derives alignment information from the second individual’s vocal characteristics, but does not align the voice of the second individual to match the singing voice of the first individual. Br. 10 (citing Kawashima, col. 18, ll. 24–31). Rather, Appellants assert, Kawashima is generally directed to “a voice converter for assimilating a user voice to be processed to

a different target voice.” Kawashima, col. 1, ll. 9–11. Figure 1 of Kawashima, as relied on by the Examiner (*see, e.g.*, Ans. 7–8) is illustrative and is reproduced below:

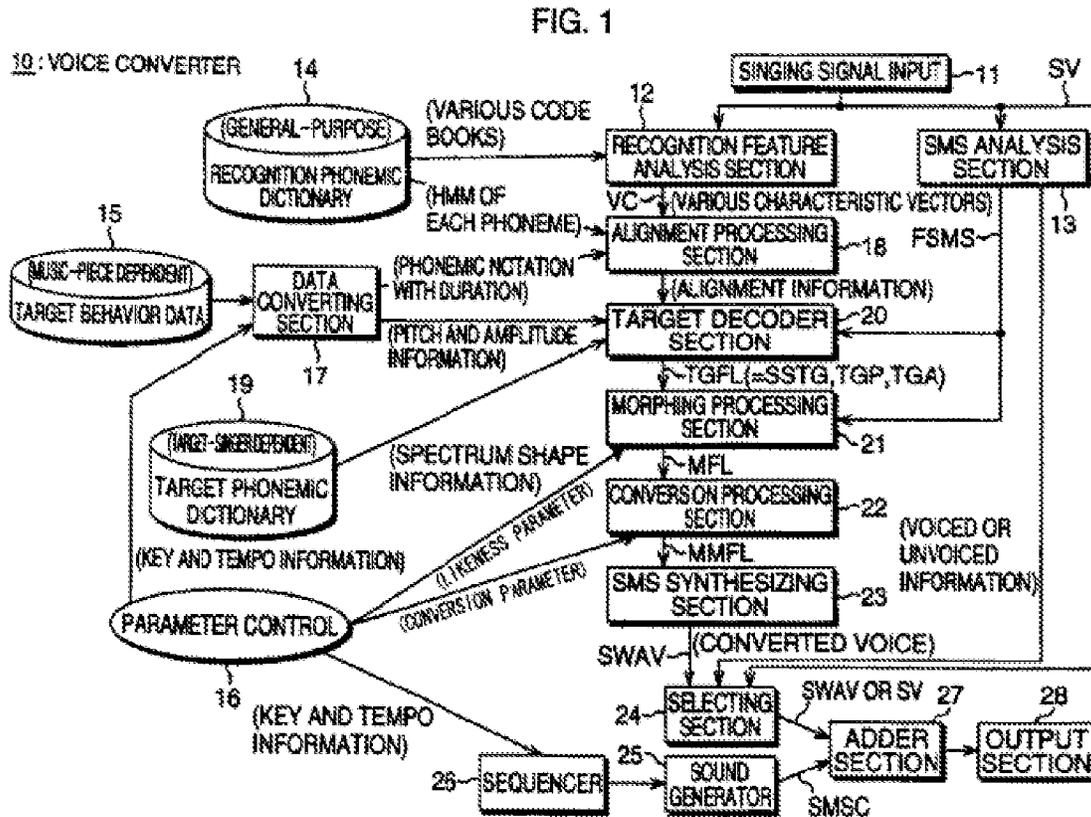


Figure 1 of Kawashima is a high-level block diagram of Kawashima’s disclosed voice converter (10). Kawashima, col. 6, ll. 59–61. Kawashima discloses the voice converter as applied to a karaoke apparatus is “capable of performing imitation of a target singer.” Kawashima, col. 7, ll. 49–52. In the disclosed embodiment, Kawashima describes that data, such as pitch and amplitude temporal changes, are extracted from the singing voice of a target singer and such information is stored in Target Behavior Data storage (15). Kawashima, col. 9, ll. 22–38. Additionally, “spectrum information corresponding to respective phonemes of the target singer to be imitated” are

stored in Target Phonemic Dictionary (19). Kawashima, col. 9, ll. 39–46. Kawashima further describe the singing signal input section (11) inputs a voice signal generated by a karaoke singer (i.e., the voice to be modified by the target characteristic data). Kawashima, col. 17, ll. 55–58. The input voice signal is analyzed in the Recognition Feature Analysis Section (12) that calculates characteristic vectors of the signal. Kawashima, col. 17, ll. 59–67. Additionally, the SMS Analysis Section (13) also analyzes the input singing signal to generate, *inter alia*, frame data. Kawashima, col. 18, ll. 10–14. Kawashima discloses the Alignment Processing Section (18):

finds respective parts of the music piece sung by the karaoke singer using Viterbi algorithm on the basis of various characteristic vectors VC outputted from the recognition feature analysis section 12, HMM [(Hidden Markov Model)] of respective phonemes from the recognition phonemic dictionary 14, and the phonemic notation information with duration included in the target behavior data.

Kawashima, col. 18, ll. 24–31. On the basis of the alignment information (i.e., the output from the Alignment Processing Section (18)), Target Decoder Section (20) calculates frame information (a pitch, an amplitude, and a spectrum shape) of the target singer. Kawashima, col. 18, ll. 42–49. Morphing Processing Section (21) morphs the singing signal (based on target frame data, frame data, and likeness data; Conversion Processing Section (22) transforms morphing frame data into conversion frame data; and SMS Synthesizing Section (23) outputs a synthesized output signal. Kawashima, col. 18, l. 50–col. 19, l. 5.

Contrary to Appellants' arguments, we agree with the Examiner that Kawashima discloses aligning the voice of the second individual (i.e., the singing signal input) to match the singing voice of the first individual (i.e.,

the target singer). *See* Ans. 8. As described, the characteristics of a singing voice of a first individual are derived and stored at least in Target Behavior Data (15) and Target Phonemic Dictionary (19). The voice of a second individual (i.e., the singing signal input) is analyzed (12) to obtain vocal characteristics, which are aligned with the singing voice data of the first individual (i.e., target singer) as part of the vocal characteristic modification process of the singing signal (i.e., the voice of the second individual).

For the reasons discussed *supra*, we are unpersuaded of Examiner error. Accordingly, we sustain the Examiner's rejection of independent claim 1. Additionally, we sustain the Examiner's rejection of claims 3–6 and 8, which depend directly or indirectly therefrom and were not argued separately. *See* Br. 11; *see also* 37 C.F.R. § 41.37(c)(1)(iv).

b. Haupt (claim 19)

Appellants contend Haupt fails to disclose “synthesizing a singing voice for the second individual in response to the converted spectral parameter sequence and the pitch contour information of the singing voice of the first individual,” as recited in independent claim 19. Br. 11–12. Rather than synthesizing a singing voice for the second individual, Appellants assert Haupt discloses adapting a pitch of the first individual. Br. 11–12 (citing Haupt ¶¶ 42–43). Additionally, Appellants argue such pitch adaptation is not in response to the converted spectral parameter sequence and pitch contour information. Br. 12.

Haupt is generally directed to converting a spoken voice to a singing voice sung in the manner of a target singer. Haupt, Title. In a disclosed embodiment, a user speaks the lyrics of a song and the captured voice is

sampled and processed to “cause[] the user’s voice to sound as if the user were singing the song with the same pitch and timing as the original artist.” Haupt ¶ 9. Haupt describes major blocks of the voice synthesis process: (i) voice analysis of the spoken fragments/words of the source speaker; (ii) voice analysis of the target singer; (iii) re-estimation of parameters of the source speaker to match the target singer; and (iv) re-synthesis of the source model to make a singing voice. Haupt ¶ 30. Haupt describes the voice analysis system is a “low-order parameterization of speech” that uses a linear prediction coefficient (LPC) analysis to break the voice into two separate components—a pitch signal and a filter signal. Haupt ¶ 32. Haupt further describes obtaining pitch estimates for the speaker’s and singer’s voice data using, for example, autocorrelation methods. Haupt ¶ 33. Haupt also discloses using timing information from (i.e., the length of a triphone) to scaled the spoken utterance to be the same length as the sung triphone. Haupt ¶ 35. Haupt further discloses the pitch of the singer “is transposed to the speaker’s voice”, using a pitch multiplication constant. Haupt ¶ 36. “Deterministic and stochastic sections of speech are transposed from the singer to the speaker.” Haupt ¶ 36. Haupt also describe that a singing voice model may be added to change a characteristic of speech into singing. Haupt ¶ 37. In addition to matching transitions between boundaries between triphones, Haupt describes an energy-shaping component to match the amplitude shape. Haupt ¶¶ 39–40. Using different variable and filter coefficients, Haupt describes a singing voice sample may be synthesized. Haupt ¶¶ 41–42. Additionally, using the timing information of the triphones, “the resultant speech sample after voice synthesis **86** is then placed in a signal timed in such a manner that the sung voice and the newly

formed sample occur at the exact same point in the song.” Haupt. ¶ 42. Haupt describes that the resulting track “will be singing in a speaker’s voice in the manner of a target singer.” Haupt ¶ 42.

Contrary to Appellants’ arguments, we agree with the Examiner that Haupt discloses synthesizing a singing voice for a second individual (i.e., the speaker) in response to the converted spectral parameter sequence and the pitch contour information of the singing voice of the first individual.

Ans. 9–10; *see also* Haupt ¶¶ 30–43. Specifically, the Examiner finds, and we agree, Haupt discloses that filter coefficients for the singer and source are transformed to the frequency domain and a scaling factor is calculated to match the amplitude shape of the source speaker to the target. Ans. 9 (citing Haupt ¶¶ 40–42). The filter coefficients and scaling factor correspond to the claims converted spectral parameters used to transform the source signal to a target singing voice. Ans. 9. Further, as the Examiner explains, Haupt discloses transposing the pitch of the singing voice to a speaker voice.

Ans. 9–10 (citing Haupt ¶¶ 36, 43). The Examiner correctly identifies that Haupt describes, as part of the synthesis of a singing voice for a second individual, the pitch of the singing voice is transposed to the second individual. *See* Haupt ¶ 36. Thus, we agree with the Examiner that Haupt discloses “synthesizing a singing voice for the second individual in response to the converted spectral parameter sequence and the pitch contour information of the singing voice of the first individual,” as recited in independent claim 19.

For the reasons discussed *supra*, we are unpersuaded of Examiner error. Accordingly, we sustain the Examiner’s rejection of independent claim 19.

*Rejections under 35 U.S.C. § 103*

a. Claim 2

Claim 2 depends from claim 1 and recites, *inter alia*, “the voice of the second individual is speech.” Appellants argue Kawashima discloses aligning vocal characteristics of the first individual’s singing voice to the second individual’s (i.e., the karaoke singer’s) *singing* voice—i.e., not to the spoken voice (speech) of the second individual. Br. 12–13. Further, Appellants argue Haupt discloses aligning a spoken utterance of a second individual with a first individual’s singing utterance. Br. 13. Appellants assert Haupt discloses performing the alignment using timing information and the pitch-synchronous overlap and add (PSOLA) method, whereas Appellants’ Specification discloses characteristics may include pitch contour information, alignment information and spectral parameter sequence information. Br. 13.

We are not persuaded of Examiner error. As the Examiner explains, and as discussed above, Haupt describes aligning the spoken voice (i.e., utterances or speech) with the singing voice of a target (i.e., first individual). Ans. 10–11 (citing Haupt ¶¶ 31–40, 48). Appellants’ discussion that the Specification discloses certain characteristics as opposed to Haupt’s use of timing information and PSOLA to perform alignment is also unavailing at least because Appellants’ characteristics of pitch contour information, alignment information and spectral parameter sequence information are not recited in either claim 1 or claim 2. *See In re Van Geuns*, 988 F.2d 1181, 1184 (Fed. Cir. 1993) (limitations are not to be read into the claims from the Specification). Additionally, as the Examiner explains (Ans. 10) and as

discussed above, Kawashima discloses aligning a singing voice of a second individual to match the singing voice of a first individual. In the rejection, the Examiner explains it would have been obvious to an ordinarily-skilled artisan to modify the singing voice of the second individual (as described by Kawashima) with the spoken voice of a second individual (as described by Haupt) to synthesize a singing voice using less effort. Final Act. 13–14; Ans. 11–12. We agree. Accordingly, we sustain the Examiner’s rejection under 35 U.S.C. § 103(a) of claim 2.

b. Claims 9, 10, and 14–17

Appellants also argue the Examiner erred in finding Cen teaches or reasonably suggests “deriving a template of first speech characteristics and first singing characteristics in response to a first individual’s speaking voice and singing voice,” as recited in independent claim 9. Br. 14–15. Rather, Appellants assert Cen teaches using “a pair of template speaking and singing voices,” not speech and singing characteristics. Br. 15.

In response, the Examiner finds Cen teaches template-based Speech-to-Singing (STS) synthesis “trains the transformation models by analyzing a pair of template speaking and singing voices that are usually [derived] from a professional singer.” Cen, Introduction; Ans. 14 (emphasis omitted). In other words, “the template-based models are first obtained from a first individual.” Ans. 14. Further, Cen describes that in the learning stage of the STS synthesis system template singing and speaking voices are further analyzed to derive transformation models based on acoustic features such as pitch contour, phoneme duration, and spectrum. Cen, Section II; Ans. 14. As the Examiner explains, “the acoustic features (characteristics) are

extracted [(from the speaking and singing voices)] to form a *template[]-based transform function* (obtained from a first individual), and a new utterance (obtained from a second individual) is converted into a singing voice by modifying the acoustic features using the transformation model.” Ans. 14 (emphasis added). In other words, the Examiner finds Cen starts with a pair of template singing and speaking voices that are then analyzed to extract acoustic features (i.e., speech and singing characteristics). Thus, the speech and singing characteristics are template based and derived from a first individual’s speaking voice and singing voice. Appellants do not rebut persuasively the Examiner’s findings.

Additionally, Appellants contend the Examiner erred in finding the combination of Haupt and Cen teaches or reasonably suggests “aligning acoustic features of the second individual’s approximated singing voice to the template of the first speech characteristics, the first singing characteristics and the second speech characteristics to generate the second individual’s synthesized singing voice.” Br. 15. In particular, Appellants argue the utterances disclosed in Haupt are not the acoustic features and speech/singing characteristics recited in claim 9. Br. 15. Further, Appellants again assert Cen does not disclose speech and singing characteristics. Br. 15.

The Examiner provides a comprehensive response detailing how each of Haupt and Cen teaches or reasonably suggests the alignment of the claimed elements. *See* Ans. 15–19. Regarding Haupt, the Examiner finds, and we agree (*see* discussion of Haupt *supra*), Haupt describes extracting acoustic features and characteristics from a speech signal (i.e., utterance) to be converted to a synthesized singing signal. Ans. 19; *see also* Haupt ¶ 33

(e.g., extracting pitch estimates). Additional characteristics extracted or determined include alignment information (e.g., triphone boundaries) and spectral parameter sequence information (e.g., filter coefficients). *See* Haupt ¶¶ 38–40; Ans. 19. Regarding Appellants’ argument the teaching of Cen, for similar reasons discussed earlier in this section, we are unpersuaded of Examiner error.

For the reasons discussed *supra*, we are unpersuaded of Examiner error. Accordingly, we sustain the Examiner’s rejection of independent claim 9. For similar reasons, we also sustain the Examiner’s rejection of claims 10 and 14–17, which depend directly or indirectly therefrom and were not argued separately. *See* Br. 16; *see also* 37 C.F.R. § 41.37(c)(1)(iv).

c. Claims 7, 11–13, and 18

Regarding dependent claims 7, 11–13, and 18, Appellants assert the additionally cited references fail to overcome the alleged deficiencies related to the independent claims. Br. 13–14, 16. Because we do not find the Examiner’s rejection of the independent claims to be in error, we similarly do not find the rejection of these claims to be in error. Accordingly, we sustain the Examiner’s rejections of claims 7, 11–13, and 18.

## DECISION

We affirm the Examiner’s decision rejecting claims 1–19 under 35 U.S.C. § 101.

We affirm the Examiner’s decision rejecting claims 1, 3–6, 8, and 19 under 35 U.S.C. § 102(a).

Appeal 2018-000374  
Application 14/383,341

We affirm the Examiner's decision rejecting claims 2, 7, and 9–18 under 35 U.S.C. § 103(a).

No time period for taking any subsequent action in connection with this appeal may be extended under 37 C.F.R. § 1.136(a)(1)(iv). *See* 37 C.F.R. § 41.50(f).

AFFIRMED