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.
11/277,793	03/29/2006	Changxue C. Ma	CML02876HI	5379

20280 7590 02/06/2013
MOTOROLA MOBILITY LLC
600 NORTH US HIGHWAY 45
W2-55BB
LIBERTYVILLE, IL 60048-5343

EXAMINER
HE, JIALONG

ART UNIT	PAPER NUMBER
2659	

NOTIFICATION DATE	DELIVERY MODE
02/06/2013	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):

DOCKETING.MOBILITY@MOTOROLA.COM

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Ex parte CHANGXUE C. MA

Appeal 2010-009737
Application 11/277,793
Technology Center 2600

Before JOHN A. JEFFERY, MARC S. HOFF, and DANIEL N. FISHMAN, *Administrative Patent Judges*.

FISHMAN, *Administrative Patent Judge*.

DECISION ON APPEAL

Appellant appeals under 35 U.S.C. § 134(a) from the Examiner's rejection of claim 20. Br. 7.[1] We have jurisdiction under 35 U.S.C. § 6(b)(1). We affirm.

---

[1] Throughout this opinion, we refer to: (1) the Revised Appeal Brief ("Br.") filed February 23, 2010; and (2) the Examiner's Answer ("Ans.") mailed March 31, 2010.

STATEMENT OF THE CASE

Appellant's invention relates to a method for producing phonetic tag variants in voice-to-phoneme conversion (such as is useful in voice recognition systems). The invention generates a feature vector from a first spoken utterance, and generates a first phonetic voice tag from that feature vector. The invention then applies perturbations to the feature vector to produce perturbed feature vectors. The perturbed feature vectors are converted into phonetic voice tag variants. A second utterance is then recognized from the phonetic voice tag variants. *See generally* Abstract.

Claim 20 is reproduced below with the key disputed limitation emphasized:

> 20. A method for producing phonetic voice tag variants in voice-to-phoneme conversion comprising:
> generating a feature vector from a first spoken utterance;
> generating a first phonetic voice tag from said feature vector;
> applying one or more perturbations to said feature vector for producing one or more perturbed feature vectors;
> converting said perturbed feature vectors into one or more phonetic voice tag variants; and
> *recognizing a second spoken utterance from said one or more phonetic voice tag variants and said first phonetic voice tag,*
> wherein a phonetic voice tag is a string of symbolic characters representing phonemes of speech.

The Examiner relies on the following as evidence of unpatentability:

Steve Young et al., THE HTK BOOK (3d rev. 2000).

Yongwon Jeong & Hyung Soon Kim, *Recognition Confidence Scoring Using Recognition Results from Perturbed Input Feature Vectors*, 37 ELECS. LETTERS 1143 (2001).

Yan Ming Cheng et al., *Voice-to-Phoneme Conversion Algorithms for Speaker-Independent Voice-Tag Applications in Embedded Platforms*, 2005 IEEE WORKSHOP ON AUTOMATIC SPEECH RECOGNITION & UNDERSTANDING 403 (2005).

The Rejections

1. The Examiner rejected claim 20 under 35 U.S.C. § 103(a) as unpatentable over Cheng in view of Jeong. Ans. 4-7.
2. The Examiner rejected claim 20 under 35 U.S.C. § 103(a) as unpatentable over Young in view of Jeong. Ans. 7-9.

The First Rejection (Cheng in view of Jeong)

The Examiner finds that Cheng in the left column of page 407 shows the step of generating a feature vector from a first utterance, and finds that Cheng in the right column of page 404 teaches generating a first phonetic voice tag from the feature vector. Ans. 5. The Examiner further finds that Cheng's Table 2 on page 408 shows that a phonetic voice tag is a string of symbolic characters representing phonemes of speech. *Id.* The Examiner acknowledges that Cheng fails to show the steps of applying perturbations, converting the perturbed feature vectors to generate phonetic voice tag variants, and recognizing a second utterance from one or more voice tag variants. *Id.* The Examiner then finds that Jeong shows these steps, citing Jeong in the left column of page 1144 as showing three different methods for applying perturbations to feature vectors. Ans. 5-6. The Examiner cites Jeong in the right column of page 1144 as showing converting the perturbed feature vectors into phonetic voice tag variants, and as showing recognizing a second utterance from the phonetic voice tag variants. Ans. 6. The Examiner asserts that there is motivation to combine Cheng and Jeong because they are in similar fields (voice recognition) and Jeong's Abstract suggests its utility for improving out-of-vocabulary rejection. Ans. 7.

The Second Rejection (Young in view of Jeong)

The Examiner finds that Young shows generating a feature vector at page 28, section 3.1.5, and shows generating a first voice tag at page 29, section 3.2. Ans. 7. The Examiner further points to Young at page 9, section 1.5, as teaching recognizing an utterance from one or more voice tags. Ans. 8. The Examiner acknowledges that Young fails to show applying perturbations or converting the perturbed feature vectors into phonetic tag variants. *Id.* The Examiner then finds that Jeong shows these steps, citing portions of Jeong as above in the first rejection. Ans. 7-9.

Appellant argues, with respect to both rejections, that Jeong fails to show the recited step of recognizing a second utterance from phonetic voice tag variants and the voice tag generated from the first utterance. Appellant appears to emphasize that Jeong fails to show anything comparable to recognizing a second utterance from the combination of a voice tag of a first utterance and voice tag variants generated from perturbations of the feature vector of the first utterance. *See* Br. 8. More specifically, Appellant suggests that: "Jeong uses the voice-tag variants of the ***first*** spoken utterance in order to improve the recognition rate of the ***first*** spoken utterance. In Jeong, there is nothing comparable to the ***second*** spoken utterance of the presently pending claim 20." Br. 8.

ISSUE

Under § 103, has the Examiner erred by finding that the cited prior art collectively teaches or suggests recognizing a second spoken utterance from

one or more phonetic voice tag variants and a first phonetic voice tag in rejecting claim 20 over Cheng and Jeong or over Young and Jeong?

ANALYSIS

First Rejection (Cheng in view of Jeong)

Based on the record before us, we find no error in the Examiner's rejection of claim 20. As regards the disputed limitation (recognizing a second utterance), the Examiner cites Jeong on the right column of page 1144 as showing the recognition of 800 utterances (test data) based on perturbed voice tag variants (perturbations of feature vectors of training data). Ans. 6. In response to Appellant's arguments, the Examiner further clarifies that Jeong shows using the training data (Jeong, p. 1144, left column) as the recited "first utterance" and perturbations thereof to recognize the test data (Jeong, p. 1144, right col., including Figure 1 test results) as the second utterance. Ans. 9-10. The Examiner further clarifies that although Jeong speaks of "feature vector" data, Cheng makes clear that conversion between feature vector representations of an utterance and a voice tag or phoneme string representation of an utterance is well known to those of ordinary skill in the art. Ans. 10.

We agree with the Examiner's finding that the training data of Jeong corresponds to the recited first utterance. Specifically, in the right column of page 1144, Jeong states: "For training, we used the Korean phonetically optimized words (POW) speech database [4]. The size of the vocabulary is 3848 words spoken by 40 male speakers" (brackets in original). Jeong further states: "Twelve mel-frequency cepstral coefficients (MFCC) and their derivatives were used as feature vectors." *Id.* Thus, the training

database of Jeong comprises a first utterance, and Jeong generates a feature vector from such a first utterance. Jeong further shows that the feature vector is perturbed when in the left column of page 1144 the reference states: "Therefore, the recognition result from perturbed input feature vectors can be employed to improve the robustness of the confidence score. . . . [T]hree different methods were tried for perturbing input feature vectors." Thus, Jeong teaches applying perturbations to produce perturbed feature vectors. Returning to the right column of page 1144 of Jeong, it is further clear that using the training data and one or more perturbations of a first utterance in that training data improves recognition of words in test data (second utterances). Thus, Jeong teaches the recited features except for conversion from a feature vector to a first phonetic voice tag and conversion from a perturbed feature vector to a phonetic voice tag variant. Appellant has not disputed that these features are taught by the art. Regardless, we find that Cheng teaches that such conversions are well known to those of ordinary skill in the art of voice recognition. For example, Cheng's Abstract on page 403 recites (emphasis added):

> In the first approach, a voice-to-phoneme conversion in batch mode manages this task by preserving the commonality of input feature vectors of multiple voice-tag example utterances. Given multiple example utterances, a developed feature combination strategy *produces an "average" utterance, which is converted to phonetic strings as a voice-tag representation* via a speaker-independent phonetic decoder.

Cheng therefore teaches the recited steps of conversion from a feature vector to a voice tag. We therefore find that Appellant has not persuasively rebutted the Examiner's finding that Cheng and Jeong collectively teach all the elements of claim 20, and we therefore affirm the Examiner's rejection.

For the foregoing reasons, Appellant has not persuaded us of error in the rejection of claim 20 over Cheng and Jeong.

Second Rejection (Young in view of Jeong)

The Examiner rejected claim 20 over Young in view of Jeong in a manner similar to the rejection over Cheng in view of Jeong. We note that in this second rejection, the Examiner does not specifically cite in the prior art recognition of a *second* utterance from the voice tag and voice tag variants generated from a *first* utterance. Rather, the Examiner generally notes Young at page 9, section 1.5, as showing recognizing *an* utterance from voice tag information. We deem this error harmless for the Examiner made clear in the first rejection (Cheng in view of Jeong) that Jeong teaches recognizing a *second* utterance (test data) from the feature vector and the perturbed feature vectors generated from a *first* utterance (training data). Further, as noted above in the Examiner's response to arguments, it was clarified that Jeong shows essentially every element of claim 20. Ans. 9. The Examiner further clarifies reliance on either Cheng or Young for showing that it is known to convert from a feature vector (or perturbed feature vector) into a corresponding voice tag.

We therefore find that Appellant has not persuasively rebutted the Examiner's finding that Young and Jeong collectively teach or suggest all the elements of claim 20. We therefore sustain the Examiner's rejection.

For the foregoing reasons, Appellant has not persuaded us of error in the rejection of claim 20 over Young and Jeong.

CONCLUSION

The Examiner has not erred in finding that the cited prior art collectively teaches or suggests recognizing a second spoken utterance from one or more phonetic voice tag variants and a first phonetic voice tag. Thus, the Examiner has not erred in rejecting claim 20 over Cheng and Jeong or over Young and Jeong.

DECISION

The Examiner's decision rejecting claim 20 is affirmed.

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).

AFFIRMED

babc