THIS OPINION WAS NOT WRITTEN FOR PUBLICATION

The opinion in support of the decision being entered today
(1) was not written for publication in a law journal and
(2) is not binding precedent of the Board.

Paper No. 49

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte SHINJI KOGA,
TAKAO WATANABE and
KAZIMAGA YOSHIDA

Appeal No. 1996-2198
Application 08/077,506[1]

ON BRIEF

Before THOMAS, MARTIN, and GROSS, Administrative Patent Judges.

MARTIN, Administrative Patent Judge.

---

[1] Application for patent filed June 17, 1993. The front of the file wrapper identifies this application as a continuation of Serial No. 07/704,717, filed May 20, 1991, as a continuation of Serial No. 07/308,277, filed February 9, 1989.

DECISION ON APPEAL

This is an appeal from the final rejection of claims 1-5, all of the pending claims, under the first and second paragraphs of 35 U.S.C. § 112 and under § 103. We reverse.

A. The invention

The invention at issue is described by appellants as an improvement in reference pattern generators using the hidden Markov model (HMM) and more particularly as an improvement in Rabiner-Levinson-type speech recognition systems (Brief at 3).

By way of background, appellant's brief describes HMMs having a discrete number of states N and a discrete number of symbols M (Brief at 9). A particular HMM is defined by:

(1) A, the state transition probability distribution $\{a_{ij}\}$ which defines the probability that the next state would be j, given that the current state is i;

(2) B, the occurrence probability distribution $\{b_j(k)\}$ which defines the probability of observing (k) given that the state is j; and

(3) C, which defines that the probability that the sequence of events will begin at state i (Brief at 10).

Page 7 of the brief shows that the state transition probability distribution A can be represented as a matrix of discrete probability values. As will appear, it was known in the art to represent the occurrence probability distribution B as either (1) a set of discrete probability values obtained from quantized feature vectors or (2) as a continuous probability function derived from non-quantized feature vectors. The § 103 question before us is whether it would have been obvious to represent the occurrence probability distribution as an approximate continuous probability function Bc derived from a set of discrete probability values B obtained from quantized feature vectors.[2]

---

[2] The Answer was accompanied by a copy of Rabiner & Juang, "An Introduction to Hidden Markov Models," IEEE ASSP Magazine, January 1986, pp. 12-15, which the examiner cites as "teach[ing] that it was obvious to extend any discrete model by substituting [sic, replacing] discrete probability functions with continuous density functions" (Answer at 5). This publication will not be considered, because it is not mentioned in the statement of the § 103 rejection and was cited for the first time in the answer. See Ex parte Movva,

(continued...)

---

[2](...continued)
31 USPQ2d 1027, 1028 n.1 (Bd. Pat. App. & Int. 1993):

> The examiner has cited and relied upon four new references in the Examiner's Answer but did not make a new ground of rejection. As set forth in In re Hoch, 57 CCPA 1292, 428 F.2d 1341, 166 USPQ 406 (1970), "[W]hen a reference is relied on to support a rejection, whether or not in a 'minor capacity,' there would appear to be no excuse for not positively including the reference in the statement of rejection." The failure of the examiner to do so here appears to be for the purpose of avoiding a new ground of rejection. Since a new ground of rejection was not made, appellants were not entitled as a matter of right to respond to this new evidence of obviousness by way of amendment and/or evidence. Rather, appellants were limited to presenting argument by way of a Reply Brief. The procedural disadvantage in which appellants were placed by the examiner's action is apparent. Accordingly, we have not considered the four references in determining the correctness of the rejection before us in this appeal. If in further prosecution of this subject matter, the examiner continues to find these references to be relevant evidence of obviousness (see n. 6, infra), a proper rejection should be made.

(continued...)

The prior art Rabiner-Levinson speech recognition system described at pages 2 and 3 of appellants' specification ("Rabiner-Levinson system"), employs an occurrence probability distribution B that is a set of discrete probability values obtained from quantized feature vectors. As is apparent from the drawing identified as Appendix D to the brief (wherein the reference numbers have the suffix "A"), the Rabiner-Levinson system uses many of the same components as appellants' invention, shown in appellants' Figure 1. Both systems have a feature analyzer 10, 10A for analyzing an input speech pattern to produce a time sequence of vectors V. When the invention is in the training mode, the mode selection switch (13; unnumbered in App. D) applies these vectors to a converting

---

[2](...continued)
Accord, In re Raske, 28 USPQ2d 1304, 1304-05 (Bd. Pat. App. & Int. 1993); Ex parte Hiyamizu, 10 USPQ2d 1393, 1394 (Bd. Pat. App. & Int. 1988). See also MPEP § 706.02(j) (6th ed., rev 3, July 1997)) ("Where a reference is relied on to support a rejection, whether or not in a minor capacity, that reference should be positively included in the statement of the rejection. See In re Hoch, 428 F.2d 1341, 166 USPQ 406, [407] n.3 (CCPA 1970).").

circuit (15, 15A) which, using code vectors (R) stored in a code vector table (14, 14A) converts the time sequence of feature vectors into a time sequence of feature codes (C, Ck), a process referred to by appellants as "quantization of the feature codes into feature codes" (Brief at 13). As noted by appellants, the advantages and disadvantages of this quantization technique are described as follows in Juang et al. (Juang) Patent No. 4,783,804:

> The recognition scheme disclosed in U.S. patent application Ser. No. 434,516, filed Sept. 2, 1982 [now Patent No. 4,587,670], discloses an arrangement that utilizes vector quantization to generate Markov model output symbol probability signals $b(O_t)$. While vector quantization techniques permit the use of permanently stored probability tables, the resulting model probabilities are only an approximation to the actual likelihood. The graph of FIG. 2 illustrates the effect of the vector quantized approximation.
>
> In FIG. 2, curve 201 represents the actual likelihood $b(O_t)$ as a function of acoustic features, and the vertical lines correspond to the discrete probability values of the discrete vector quantized features. An input acoustic feature derived from a speech pattern is first quantized to the nearest prototype feature and the probability of the prototype feature is selected as $b(O_t)$. It is readily apparent that an input feature at x_1 on the feature axis corresponds to a probability y_1 from curve 201 but that the probability for the nearest prototype vector 205 is y_2. The difference between y_1 and y_2 is the error due to quantization and the

> error affects the accuracy of speech pattern recognition. [Col. 5, lines 11-32.]

According to appellants, Juang avoids errors relating to discrete representations by replacing the code vector table and converting circuit of the Rabiner-Levinson system with a function generator which generates sets of continuous probability density functions (Brief at 13-14). Because the § 103 rejection is not based on Juang, we need not determine whether this is an accurate characterization of Juang's invention.

Appellants' invention, in contrast, does not discard the code vector table (14, 14A) and converting circuit (15, 15A) of the Rabiner-Levinson system. Instead, it adds a function generator 17 which generates an approximate continuous occurrence probability function Bc derived from the set of discrete probability values B generated by the first pattern forming circuit (16, 16A). This function can take the form of either a Gaussian probability density function or a Poisson probability density function (Spec. at 11, line 27 to p. 12, line 6). The approximate continuous occurrence probability function Bc and the corresponding state transition probability distribution B are combined in second pattern forming circuit 18 to form a pattern P, which is stored in circuit 18 (Spec. at 13, lines 1-10).

According to appellants, the converting circuit 15 and code vector table 14 permit the first pattern forming circuit 16 to rapidly generate the state transition probability distribution A and the feature code occurrence probability distribution B of the HMM (Brief at 3), while the approximate continuous occurrence probability density function Bc gives a better approximation of the actual probability than would probability distribution B (Brief at 4).

B. The claims

Claim 1, which is representative, reads as follows:

1. A reference pattern generating device including:

a feature analyzer, responsive to a speech signal representative of an input pattern, for producing a time sequence of feature vectors representative of said input pattern; a table storing a plurality of code vectors of known vocabulary and a plurality of feature codes respectively corresponding to said plurality of code vectors; converting means, connected to said feature analyzer and said table, for converting a plurality of time sequences of feature vectors to a plurality of time sequence of feature codes with reference to said table, said plurality of time sequences of feature vectors being produced in response to a plurality of speech signals; and first forming means for forming, in response to said plurality of time sequences of feature codes, a state transition probability in a state transition network and a probability distribution of occurrence of the feature codes in each state in said state transition network;

wherein the improvement comprises:

function generating means, connected to said table and said first forming means, for generating an approximate continuous probability density function said approximate continuous probability density function approximating said probability distribution of occurrence of the feature codes in each state in said state transition network, said code vectors being used as parameters in said approximate continuous probability density functions; and

second forming means, connected to said first forming means and said function generating means, for forming as a reference pattern for said plurality of speech signals a combination of said state transition probability distribution and said approximate continuous probability density function.

Appellants treat claims 3-5 as standing or falling with claim 1 and argues claim 2 separately (Brief at 15).

C. The references

The references relied on the § 103 rejections are:

Baker 4,803,729[3] Feb. 7, 1989

---

[3] See Paper No. 31, at 6. The Answer erroneously (continued...)

(filed Apr. 3, 1987)

Kuroda et al. (Kuroda) 4,829,577 May 9, 1989
(filed Mar. 12, 1987)

Webster's Ninth New Collegiate Dictionary (1986 ed.), p. 806. (Webster).

D. The grounds of rejection

Paper No. 31, the Office action that immediately precedes and is incorporated by reference into the final Office action (Paper No. 34), states (at 6 and 8-10) that the claims are rejected on the following grounds:

Claims 1-5 stand rejected under the first paragraph of § 112 for being based on a specification that "fail[s] to provide an adequate written description of the invention and failing to present the best mode."

Claims 1-5 also stand rejected under the second paragraph of § 112 for being indefinite.

Claims 1 and 3-5 further stand rejected under § 103 for obviousness over Kuroda in view of Baker.

---

[3](...continued)
identifies Baker et al. Patent No. 4,805,219 as the Baker reference relied on in the § 103 rejection.

Claim 2 further stands rejected under § 103 for obviousness over Kuroda in view of Baker and Webster.

E. The merits of the § 112, first paragraph, rejection

Although the statement of the § 112, first paragraph rejection asserts that the claims fail to satisfy the written description and best mode requirements, the examiner's arguments implicate only the enablement requirement.

The examiner contention that "[t]he approximate continuous probability density function Bc is not defined in the specification" (Answer at 7, lines 20-21) is unpersuasive, because the examiner has not explained why the formulas given at page 12 fail to adequately explain how to obtain the Gaussian probability density function.

The examiner also challenges the sufficiency of the disclosure of the second pattern forming circuit 18, which the specification as filed described as follows:

> The second pattern forming circuit 18 receives the state transition probability distribution A from the first pattern forming circuit 16 and the approximate continuous [feature code] probability density function Bc from the function generator 17 and combines them to form a second pattern. The

> second pattern circuit memorizes the second pattern as the reference pattern.
>
> Reference patterns are generated and memorized for different training speeches in the similar manner as described above. [Spec. at 13, lines 1-10.]

The examiner contends that

> The specification . . . fails to indicate what type of apparatus block 18 is or what particular function it performs. The arguments vaguely indicate that the apparatus of the "second pattern forming circuit 18" is capable of "joining or associating" data. This is contrary to the function of pattern forming." [Answer at 6-7.]

The examiner appears to believe that because the first pattern forming circuit 16 performs Baum-Welch calculations and thus is not just a memory device, the same must hold for the second pattern forming circuit 18. We do not agree that the use of the term "pattern forming circuit" to refer to two different claimed elements implies that the elements employ similar structures or perform similar functions. Nor do we agree that the term "pattern forming" necessarily implies that the second pattern forming circuit more be doing more than simply storing a combination of the state transition probability distribution and the approximate continuous probability density function as a reference pattern for later comparison to an unknown incoming speech signal in identifier 12 (Spec. at 13, lines 11-27).

For the foregoing reasons, we cannot sustain the § 112, first paragraph, rejection of claims 1-5 for nonenablement.

F. The merits of the § 112, second paragraph, rejection

The examiner makes separate indefiniteness arguments with respect to claims 1, 3 and 5. Claim 1 is said to be indefinite because "[i]t is unclear how the 'second forming means' and the 'first forming means' operate to produce 'a reference pattern."

This argument, which appears to assert nonenablement rather than indefiniteness, is unconvincing for the reasons given above.

Claims 3 and 5 are allegedly unclear because

> both indicate that the "identifying means" is connected to the "second forming means." However, the "second forming means" is contained in the "reference pattern generating means." The "mode selection switch means" selects between the "training mode" and "recognition mode" which are separately performed by the "reference pattern generating means" and the "identifying means," respectively. Therefore, the "identifying means"

> will be inoperative because the "second forming means" will not operate when the "identifying means" ("recognition mode") is selected by the switch. [Answer at 8.]

As appellant correctly notes (Brief at 26), it is clear from the specification that mode selection switch 13 does not control the power to reference pattern generator 11 and thus to second pattern forming circuit 18, as the examiner appears to believe. Instead, when set for operation in the training mode, the selection switch connects the feature vectors from feature analyzer 10 to the input of reference pattern generator 11, which generates reference patterns and stores them in second pattern forming circuit 18 (Spec. at 8, line 7 to p. 13, line 10). When set for operation in the identifying mode, the selection switch connects the output of feature vectors from feature analyzer 10 to identifier 12 for analysis using the patterns stored in second pattern forming circuit 18 (Spec. at 13, lines 11-27).

Finally, the examiner contends (Answer at 9) that claim 5 is indefinite because the recitation "wherein said approximate continuous probability density function is only calculated while in the training mode" is redundant in view of the fact that the recited function generating means, which makes this calculation, is part of the reference pattern generating means. Appellant argues (Brief at 26) that the "wherein" clause "only clarifies that the density function is being calculated while the system is in the training mode. It does not make the claim indefinite or ambiguous." We agree.

For the foregoing reasons, the rejection of claims 1-5 under § 112, second paragraph, is reversed.

G. The merits of the § 103 rejection of claims 1 and 3-5 over Kuroda in view of Baker

The examiner reads the first three elements of claim 1 onto Kuroda's HMM speech recognition system as follows:

Claim	Kuroda
"feature analyzer"	feature extraction circuit 4
"table"	parameter table 11
"converting means"	labeling circuit 5

We agree with appellants (Brief at 29) that the examiner's reliance on Kuroda's parameter table as the claimed table is incorrect. It appears to us, and appellants do not deny, that this limitation corresponds instead to Kuroda's label prototype dictionary 6, from which labeling block 5 selects the prototype that is closest to the feature extracted by feature extractor 4 (col. 3, lines 27-50).

Next, the examiner reads the claimed first forming means for forming the state transition probability (A) and the probability distribution (B) of occurrence of the feature codes onto Kuroda's training block 8 and adaptation block 9. Appellants concede (Brief at 31 & n.17) that Kuroda's training block 8 generates probability distributions A and B, which Kuroda represents as parameter P(i,j,k). These discrete parameter values, which are stored in parameter table 11 in the manner shown in Figure 6, represent the probability that a transition from the state i to the state j occurs in a Markov model and that a label k is produced at that i6j transition (col. 4, lines 33-40). For the foregoing reasons, the examiner's additional reliance on adaptation block 9, which tailors the stored parameter values to a different speaker or to the same speaker at a different time (col. 1, lines 55-68; col. 7, lines 12-45), is unnecessary.

The examiner next relies on Kuroda's training and adaptation blocks 8 and 9 in combination with Baker to satisfy claim 1's requirement for "function generating means . . . for generating an approximate continuous probability density function." Specifically, the examiner states that such a means is

> suggested by Kuroda's training [block] 8 and adaptation [block] 9, fig. 1[,] for calculating new parameters by computing a weighted sum of the probabilistic frequencies and Baker's figure 2[,] which is a schematic representation of how phonetic

> frame models can be derived from samples of speech. . . .
>
> . . . .
>
> It is noted that Kuroda does not explicitly teach the use of "probability density." However, it would have been obvious to one of ordinary skill in the art to use probability density functions with the Hidden Markov models of Kuroda because Baker teaches that it is well known in the art to develop probability density functions from the analysis parameters used in Markov models (figure 2). [Paper No. 31, at 7.]

We do not agree. As already noted, the parameter values generated by Kuroda's training and adaptation blocks 8 and 9 and stored in parameter table 11 represent discrete probability values rather than a continuous probability function. Turning now to Baker, Figure 2 shows each frame 62 of a spoken sequence 60 having twelve spectral parameters 64, whose energy levels are depicted in bar graph 65 (col. 12, lines 29-42).[4] After the frames are grouped by the operator into segments 66A-E representing different speech sounds (col. 13, lines 3-19), all of the frames of a segment are combined to form a phonetic frame model 70 having a different

---

[4] The actual embodiment tested by the inventor, twenty-four spectral parameters were used (col. 12, lines 42-45).

"dimension" corresponding to each of the twelve parameters 64 (col. 13, lines 20-27). Each dimension is assumed to have the general form shown in Figure 3A, which is a Laplacian probability distribution definable by only two variables, mu and sigma (col. 13, lines 27-35). Baker explains that this simplifies the computation and storage required to represent each dimension (col. 13, lines 35-38). As appellants correctly note (Brief at 32), Baker's Laplacian phonetic frame models are not developed from the analysis parameters used in Markov models, as suggested by the examiner. In fact, the Laplacian phonetic frame models shown in Figures 2, 3A, and 3B are developed without using any HMM analysis. Instead, each

Laplacian phonetic frame model simply represents the features of an incoming segment of speech and thus is more akin to the feature information extracted by Kuroda's feature extraction block 4 than to the HMM parameters stored in Kuroda's parameter table 11. We note in passing that although Baker discloses using the Laplacian phonetic frame models in a "smooth frame labeling" technique that employs HMMs (Figs. 6-

11; col. 15, line 50 to col. 20, line 31), the examiner has not alleged, let alone demonstrated, that any HMM probability distribution is represented as a Laplacian function.

For the foregoing reasons, the § 103 rejection of claims 1 and 3-5 based on Kuroda in view of Baker. The § 103 rejection of claims 2 is reversed because the deficiencies of these two references are not cured by Webster.

REVERSED

JAMES D. THOMAS Administrative Patent Judge	)	
JOHN C. MARTIN Administrative Patent Judge	)	BOARD OF PATENT APPEALS AND INTERFERENCES
ANITA PELLMAN GROSS Administrative Patent Judge	)	

Sughrue, Mion, Zinn, MacPeak & Seas
2100 Pennsylvania Avenue, N.W.
Washington, DC   20037-3202

JCM/cam