

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opinion in support of the decision being entered today (1) was not written for publication in a law journal and (2) is not binding precedent of the Board.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. 59

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte YOICHI TAKEBAYASHI,
HIROSHI KANAZAWA, AND
HIROYUKI CHIMOTO

Appeal No. 1997-4121
Application No. 08/427,272

HEARD: March 7, 2000

This is a decision on appeal from the final rejection of claims 1 through 15, all of the claims pending in the application.

The invention pertains to time series signal recognition in general and to speech recognition in particular. Highly accurate recognition tests in quiet environments have been known but the recognition rate is drastically reduced in noisy environments due to false pattern matching or failure of appropriate word boundary detection. While word spotting, or continuous pattern matching, has been used to alleviate the problem, other problems have arisen. As depicted in prior art Figure 1, the word boundary for a particular signal gets narrower as the noise level increases. But, the speech recognition dictionary

employed in such speech recognition systems is typically prepared by using the word feature vectors obtained by using specific word boundaries and a specific noise level. Therefore, when the conventional dictionary is used with the word spotting method, matching with the word feature vector obtained from an unfixed word boundary for a speech mixed with noise having a low signal/noise ratio, as in a practical environment, causes some recognition errors. The instant invention is directed to obtaining a high recognition rate even in noisy environments. It also includes an effective learning system for word spotting methods of speech recognition.

Representative independent claim 1 is reproduced as follows:

1. An apparatus for time series signal recognition, comprising:
 - means for inputting signal patterns for time series signals to be recognized;
 - means for recognizing the time series signals according to a word spotting scheme using continuous pattern matching, including;
 - means for extracting a plurality of candidate feature vectors for characterizing an individual time series signal from the signal patterns;
 - recognition dictionary means for storing reference patterns with which the individual time series signals are matched;
 - means for calculating similarity values for each of the extracted candidate feature vectors and the reference patterns;
 - means for determining a recognition result by selecting one of said stored reference patterns that matches with one of the candidate feature vectors by the continuous pattern matching for which the similarity value calculated by the calculating means is greater than a prescribed threshold value; and
 - means for learning new reference patterns to be stored in the recognition dictionary means, including:
 - means for mixing speech patterns with noise database patterns representing background noises, to form signal patterns for learning, and supplying the signal patterns for learning to the recognizing means;
 - means for extracting feature vectors for learning from the recognition results and the

Appeal No. 1997-4121
Application No. 08/427,272

similarity values obtained by the recognizing means using the signal patterns for learning; and means for obtaining new reference patterns from the feature vectors for learning extracted by the extracting means and storing the obtained new reference patterns in the recognition dictionary means.

The examiner relies on the following references:

Takebayashi et al (Takebayashi)	4,783,802	November 8, 1988
Bahler	4,481,593	November 6, 1984

Claims 1 through 15 stand rejected under 35 U.S.C. 103 as unpatentable over Takebayashi in view of Bahler.

Reference is made to the brief and answer for the respective details of the positions of appellants and the examiner.

OPINION

We reverse.

Each of the independent claims 1, 6 and 11 contains limitations requiring, *inter alia*, that the time series signals are recognized according to “a word spotting scheme” and that the means for learning new reference patterns to be stored in the recognition dictionary means includes a “means for extracting feature vectors for learning from the recognition results...”[emphasis ours]. Thus, every claim requires a combination of using word spotting, which uses a continuous pattern matching, as the recognizing scheme, and a learning scheme which employs the results from the recognition unit (note the connection between

Appeal No. 1997-4121
Application No. 08/427,272

element 8 and element 14 in Figure 2) in extracting word feature vectors in the learning means 3.

Word spotting schemes were well known and such is admitted by appellants (see top of page 2 of the specification). Thus, appellants do not rely on “word spotting,” per se, for patentability.

Rather, it is the combination of the word spotting scheme and the specific learning scheme which distinguishes over the applied prior art.

While word spotting was well known, the references to Takebayashi and Bahler are alleged by appellants to be deficient in showing any word spotting schemes. Since word spotting is a continuous pattern matching method in which the word boundary is taken to be not fixed but flexible (see the definition at lines 3-4 of page 2 of the specification), it does appear that neither Takebayashi nor Bahler discloses such a method. Takebayashi uses a recognition means which detects starting and ending points, thus fixing the word boundary. While Bahler does disclose the recognition of keywords in a “continuous” audio signal, and would thus appear to be indicative of a word boundary which is not fixed, further review of the reference would appear to indicate otherwise. For example, the reference discloses a plurality of templates representing silence, or nonspeech, signals, which would appear to indicate that these templates would set some kind of boundary around a spoken word. In any event, appellants submit, as evidence, a declaration under 37 CFR 1.132, from Mr. Hidenori Shinoda, a clearly skilled artisan. Mr. Shinoda explains that the concatenation technique of Bahler “is a technique for connecting two or more specific word templates in order to detect word sequences..., which is part of the syntax control used in determining beginning and ending boundaries of unknown keywords in Bahler.” Mr. Shinoda explains, further, that this “technique is totally unrelated to whether the recognition is ‘word spotting’ or not, because the difference between ‘word spotting’ and ‘non-word spotting’ is not a matter of how word templates are

connected. Rather, it is a matter of whether continuous matching using a plurality of candidate feature vectors is made or not.” Mr. Shinoda also explains the “optional dwell time” of Bahler which is only related to how long each pattern in Bahler is permitted to be. “This has nothing to do with whether the recognition is ‘word spotting’ or not because the difference between ‘word spotting’ and ‘non-word spotting’ is not a matter of how long each pattern is permitted to be, rather it is a matter of whether or not continuous matching using a plurality of candidate feature vectors is made.”

The examiner’s response is merely to contend that Mr. Shinoda’s declaration “clearly contradicts the plain teachings of patent 4,783,802...” (Answer-page 9). When we weigh the evidence provided by the references and by Mr. Shinoda’s declaration against the examiner’s opinion that the cited references clearly teach a word spotting scheme, we find for appellants.

Even if, arguendo, Bahler is somehow construed to disclose a word spotting scheme, it is not at all apparent to us, how or why the artisan would have combined such a teaching with Takebayashi to arrive at the instant claimed subject matter.

Moreover, even if a word spotting scheme was incorporated into Takebayashi, it is still not clear how the limitation requiring the learning means to have a means for extracting feature vectors for learning from the recognition results...” is met. Clearly, this limitation refers to the connection, in Figure 2, of the similarity decision unit, 8, to word feature vector extraction unit, 14. Yet, Takebayashi, in Figure 1, shows no such connection from the pattern matching section 18 to the learning section 22. Accordingly, we find

Appeal No. 1997-4121
Application No. 08/427,272

that the examiner has not established a prima facie case of obviousness with regard to the claimed subject matter.

Appeal No. 1997-4121
Application No. 08/427,272

The examiner's decision rejecting claims 1 through 15 under 35 U.S.C. 103 is reversed.

REVERSED

Errol A. Krass)	
Administrative Patent Judge)	
)	
)	
)	BOARD OF PATENT
Anita Pellman Gross)	APPEALS AND
Administrative Patent Judge)	INTERFERENCES
)	
)	
)	
Lee E. Barrett)	
Administrative Patent Judge)	

EAK/wgb

Appeal No. 1997-4121
Application No. 08/427,272

Foley & Lardner
3000 K Street, NW
Suite 500
Washington, DC 20007-5109