

Please type a plus sign (+) inside this box →

PTO/SB/16 (2-98)
Approved for use through 01/31/2001. OMB 0651-0037
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
e required to respond to a collection of information unless it displays a

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (c).

INVENTOR(S)					
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)			
D. Lee Mark	Walker Mackelprang	Tucson, ARIZONA Tucson, ARIZONA			
☒ Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto					
TITLE OF THE INVENTION (280 characters max)					
A method for enhancing quality of text-to-speech synthesis based on contextual analysis of data					
Direct all correspondence to:		CORRESPONDENCE ADDRESS			
☐ Customer Number		Place Customer Number Bar Code Label here			
OR Type Customer Number here					
☒ Firm or Individual Name	Authentix Network, Inc.				
Address	Attn: D. E. WINN				
Address	4400 E. Broadway, Ste.200				
City	Tucson	State	ARIZONA	ZIP	85711
Country	USA	Telephone	5203233280	Fax	5203204177
ENCLOSED APPLICATION PARTS (check all that apply)					
☐ Specification Number of Pages		☒ Small Entity Statement			
☐ Drawing(s) Number of Sheets		☒ Other (specify) Claim, Assignment			
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)					
☒ A check or money order is enclosed to cover the filing fees					FILING FEE AMOUNT (\$)
☐ The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 					\$150
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
☒ No.					
☐ Yes, the name of the U.S. Government agency and the Government contract number are: _____					

Respectfully submitted,

Date | 5,17,00

SIGNATURE 
DELBERT E. WINN

REGISTRATION NO.

N/A

TYPED or PRINTED NAME

(if appropriate)

Docket Number:

ATX-TTSS

TELEPHONE 770-587-5699

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the complete provisional application to the PTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, Washington, D.C., 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C., 20231.

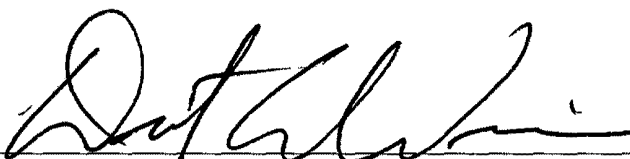
Samsung v. Cerence
IPR2025-00458
Exhibit 1015

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

I certify that this document and all enclosures are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" under 37 CFR § 1.10 on the date indicated below and is addressed to the Assistant Commissioner for Patents, Washington, DC 20231

"Express Mail" Mailing Label Number: ET130473727 US

Date of Deposit: 5/17/00



Signature of Person Mailing

Type or Print Name of Person Mailing:

Delbert E. Winn
Authentix Network, Inc.
4400 E. Broadway, Suite 200
Tucson, AZ 85711
(520) 323-3280

Documents enclosed:

1. Return Postcard
2. Statement Claiming Small Entity Status – Small Business Concern
3. Certificate of Mailing by Express Mail
4. Claim
5. Recordation Form Cover Sheet
6. Assignment
7. Check (\$150 – for Provisional Patent Application)
8. Check (\$ 40 – for Assignment)

A method for enhancing quality of text-to-speech synthesis based on contextual analysis of data

Summary

This invention provides a sophisticated, unique filtering system to modify text based on its context, prior to it being synthesized by a text-to-speech tool. The result is a higher quality, smoother, more naturally flowing speech pattern produced by a text-to-speech (TTS) engine, which is critical for the wider acceptance of voice-computer interfaces.

The direct translation of the written word to the spoken word is not usually a smooth process. Given text from an email message, news story, web page, or any text source, TTS engines do their best to synthesize the words into understandable speech. However, the result is often an unnatural speech delivery because of the diversity and context of the written word. Even small differences in what people are accustomed to in normal conversation can cause large divides in how people perceive the quality and naturalness of computer generated speech.

Because TTS engines are general-purpose tools they deal with text pre-processing in a general way and don't perform satisfactorily when presented with out of the ordinary text. For instance, if a document containing the text "10-5" is processed by a TTS engine, it must make a decision as to how to say it. Because it doesn't know the context of the document it will likely select something that has the highest general chance of being correct. So, perhaps it would be pronounced "ten minus five." This might be correct, but it might not. If the document had to do with athletics, the correct pronunciation would likely have been "ten to five."

Additionally, textual documents are often not of a single theme or topic. There may be a dominant theme and perhaps a number of sub-themes in some sort of priority order. For example, in a news story about a baseball player recovering from a serious injury, the dominant theme may be baseball, but a sub-theme would be medicine. In this example, the system will process the text first for baseball context, then for medicine, and lastly a pass to pick up any general TTS issues.

This invention provides three methods for processing text prior to it being sent to a TTS engine for synthesis.

- 1) The system automatically detects the context of a document by comparing the data to an internal database of known words and phrases. The system can be configured to select a maximum number of contexts. If there are multiple contexts detected, it prioritizes based on the number of hits for each one.
- 2) A user indicates specifically the context of the information. This works where the user has control of the text prior to its feeding to the TTS engine, and knows content of the text.

3) Using mark-up language tags, a user controls within a document where selected contexts can be used. This allows contexts to be applied to specific parts of a document.

This invention consists of the following software components:

1) Context Dictionary Rules Manager – Used to define the rules that make up a Context Dictionary. These are the rules that are applied at run-time to replace, modify or otherwise change the text before it is synthesized. A dictionary can be thematic, cultural, regional, or industry specific. Additionally, TTS engine-specific dictionaries can be added to handle the differences in the TTS engines on the market. Also, for each dictionary, key words and phrases can be attached which the Context Detector uses to try to determine what context dictionaries to apply to text. These are used only in the automatic detection mode when the user does not specify a context. Additionally, global rules can be defined that apply to all text processed.

2) Context Detector – Used to analyze a document based on the key words and phrases associated with the dictionary. This is used to determine which Context Dictionary to use only if not specified by the user.

3) Text Cleaning Server – Used to process text in real-time by applying rules of the selected Context Dictionary or by parsing mark-up tags in the text itself.

While this invention is based on the processing of English language text, it is designed in apply to any language.

Procedure

1. The user creates and names a Context Dictionary and adds rules to the dictionary on how to modify text. The user optionally identifies key words and phrases that when found in a document tell the system
2. When an application wants to convert text to speech, it sends the text to the Text Cleaning Server and specifies 1) a specific context dictionary or dictionaries to apply, or 2) to process the text based on tags in the text itself, or 3) to automatically try to detect a matching context. Each successive method requires increasingly more time and processing power to complete.
3. The Text Cleaning Server, based on the information passed to it, applies all appropriate rules, modifies the text, and passes it to the proprietary TTS Server.
4. The TTS Server queues the TTS task for a specific TTS engine to process.
5. The synthesized text is then streamed back to the original application where it is delivered to the user.

