(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: METHOD OF OPERATING A VOICE-CONTROLLED NAVIGATION SYSTEM
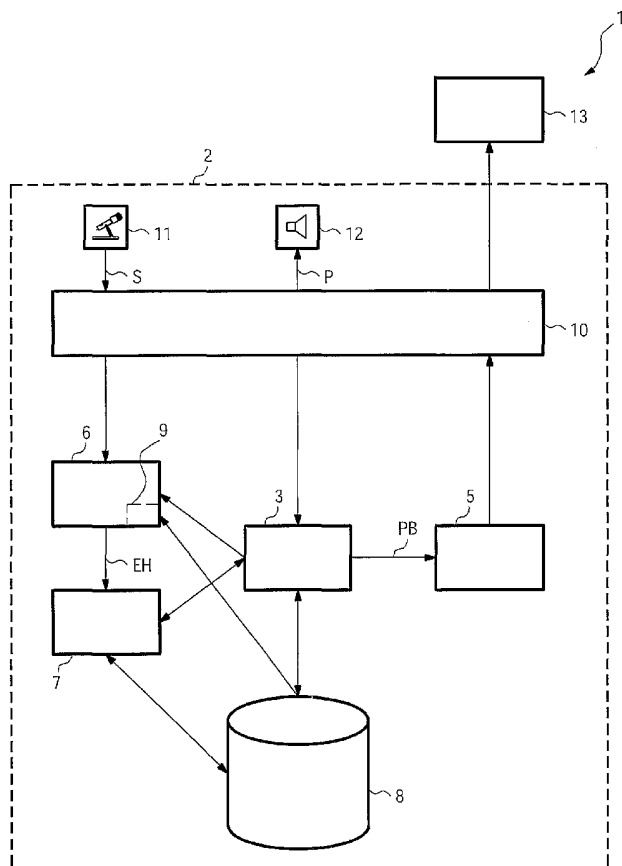
(57) Abstract: A method of operating a voice-con-
trolled navigation system (1) is described, in which,
within an automatically conducted dialog, taking
account of geographical criteria (GK), input requests
(P) are generated and output to a user, and responses
spoken by the user (S) are detected. The spoken
responses (S) are analyzed for recognition of location
data using an automatic voice recognition method,
taking account of the geographical criteria (GK). In
addition, a corresponding voice-data user interface for
a navigation system is described.

1

Method of operating a voice-controlled navigation system

The invention relates to a method of operating a voice-controlled navigation system. In addition, the invention relates to a voice-data user interface for a navigation system, a navigation system with a voice-data user interface of this kind, and a computer program in order to execute the method on a processor of a voice-data interface of a

5   navigation system. Furthermore, the invention relates to a method of generating a geographical database for use in the said method in order to operate a voice-controlled navigation system.

Modern motor vehicles are increasingly using navigation systems. Navigation systems of this kind enable the user to work out routes to particular destinations and to be

10   guided along the route during the journey. To be able to offer these functions, navigation systems use geographical data containing *inter alia* information on geographical areas, towns, locations, buildings, streets, junctions, favorable journey times along specific sections of road, speed limits on roads, etc. Using this geographical data, a navigation system is able to find the optimum route, i.e. the shortest and/or fastest one, from a start point to a particular

15   destination. The start point and/or the destination can be input by the user via a suitable user interface. Alternatively, the start point – to the extent that it is the current start point – can also be determined using automatic position-determining equipment, e.g. GPS, in the case of some navigation systems.

20

The user interface generally comprises a keyboard for inputting the location data. Deemed to be "location data" hereby is geographical data on any locations, areas, buildings, roads etc. The more convenient navigation systems are equipped, alternatively or in addition, with a voice-data user interface, with which the user can communicate in natural

25   language. Since a voice-data user interface enables the hands-free operation of the particular equipment, the controlling of navigation systems in motor vehicles in this manner is to be preferred from the safety aspect. The driver can operate the navigation system during the journey without having to take his hands off the vehicle's steering wheel in order to do so.

2

In the case of use of a voice-data user interface of this kind, a spoken response expressed by the user, e.g. specifying a location or giving a command, is detected as a voice signal by means of a microphone. The spoken response is then sent to a voice recognition device so that the location or command can be recognized and passed on, in machine-
5   readable form, to the control device of the navigation system. Voice recognition systems generally operate in a manner such that the spoken response (also referred to below as the voice signal) is firstly analyzed spectrally or time-wise, and the analyzed voice signal is then compared, section by section, with different models of possible signal trains with known voice information. To this end, the voice recognition system is generally equipped with an
10   entire library of different possible signal trains. Using the comparison of the received voice signal with the available models, the model that best matches a particular section of the voice signal is selected in order to reach a recognition result. The probability with which each model matches the associated section of the voice signal is hereby normally calculated. During the analysis and calculation of the probability of how well the individual models
15   match the relevant sections of a voice signal, grammatical and/or linguistic rules are generally consulted. This avoids the possibility that the individual sections of the longer voice signal simply match well in isolation with the particular models available, and ensures that the context within which the sections of voice signal exist is taken into account in order that a useful recognition result is arrived at overall, thereby reducing the error rate. However,
20   one problem with the use of voice recognition systems in navigation systems consists in the fact that location data, i.e. the designations of towns, federal states, streets, buildings etc., will often involve proper names, the spelling and pronunciation of which may be extremely unusual. The recognition result could be improved by making available to the voice recognition system complete models of all possible location data as a restricted, active
25   vocabulary. Owing to the huge number of possibilities, however, a restriction of the vocabulary of the voice recognition device completely globally to all location data that could possibly arise is not practicable. On the other hand, the problem exists, particularly in the case of voice-data user interfaces for navigation systems, that the inputting of spoken responses generally has to take place under unfavorable conditions, i.e. with a relatively large
30   amount of background noise. To this extent, additional conditions for improvement of recognition quality during voice recognition would be extremely helpful.

3

It is an object of the present invention to specify an improved method of operating a voice-controlled navigation system, and a corresponding, voice-data user interface for a navigation system, which, in a simple manner, increase the quality of voice recognition.

5   This object is achieved by a method of operating a voice-controlled navigation system in which input requests are generated within an automatically conducted dialog, taking account of geographical criteria, and output to a user, responses spoken by the user are detected and the spoken responses are analyzed for recognition of location data using an automatic voice recognition method, taking account of the geographical criteria.

10   In terms of equipment, the object is achieved by a voice-data user interface for a navigation system with an output device for outputting input requests to a user, with a voice-inputting device for detecting spoken responses by the user, a dialog-control device for controlling a dialog with the user, taking account of geographical criteria, a prompt-generation unit for generating input requests, taking account of geographical criteria, a voice

15   recognition device and an analysis unit for analyzing the detected spoken responses for the recognition of location data, taking account of geographical criteria, and a geographical database and/or a data interface for accessing a geographical database, which makes available geographical criteria and/or geographical data for the dialog-control device and/or the prompt-generation device and/or the voice recognition device and/or the analysis unit.

20   Control of the dialog sequence by means of the dialog-control device hereby takes place using a dialog writing system stored in a special dialog-writing language in the system. This may be any dialog-writing language. Examples of usual languages are method-oriented programming languages such as C or C++ or so-called hybrid languages, which are declarative and method-oriented, such as Voice XML or PSP HDDL, which are languages

25   with a structure similar to HTML, the language generally used to write Internet sites. Control hereby takes place essentially through the outputting of corresponding input requests, generally also known as prompts, to the user.

Using the automatically conducted dialog, the desired destination is determined iteratively in multiple query steps. As the individual prompts are output within

30   the automatic dialog to take account of geographical criteria, these geographical criteria can be used accordingly, within the voice recognition method for recognizing a spoken response following the particular prompt, to restrict the active vocabulary or to evaluate language hypotheses, etc. As a result, the recognition results in the individual stages of the dialog are

4

considerably improved, which leads overall to an extremely reliable recognition of the correct location.

A generation of the corresponding input request may, for instance, take place in a manner such that a particular input request is selected from a group of possible input requests already specified. In the same way, the input request may also be completely newly generated as a function of the particular position within the dialog sequence.

A prompt of this kind can, in principle, be generated in any manner by the prompt-generation unit and output to the user, e.g. in written form on a display or similar. Preferably, however, the output takes place in spoken form. This has the advantage that the user can register the prompt while simultaneously continuing to observe the traffic, as a result of which the operation of the navigation device becomes even safer. Both a spoken and a simultaneous visual output are also possible. Used for the spoken output may be, for instance, a voice synthesizer (text-to-speech converter) to convert the prompts output from a text form into a spoken form. To the extent that pre-prepared prompts are used, these may also be stored in an audio database. It is also possible for the prompt-generation unit to compile a prompt section-by-section from prepared audio data, for instance particular parts of a sentence, hereby also generating individual parts by means of a voice synthesizer where applicable.

The dependent claims contain especially advantageous embodiments and further embodiments of the invention. The voice-data user interface in accordance with the invention may hereby be further developed by analogy with the claims relating to the method.

There are various options for using the geographical criteria on the basis of which a prompt has been generated, when a subsequent spoken response is recognized.

In a preferred embodiment, using the geographical criteria taken into account in the generation of the input request, a word list is compiled, which serves as a restricted active vocabulary during the voice recognition of the subsequent spoken response by the user. So, for example, on generation of an input request "In which federal state is your destination located?" (wherein, in this example, it is assumed that this is a navigation system with a geographical database currently in use that covers the whole of Germany), a relatively short word list, comprising only the names of all federal states in the country, would be used in the voice recognition of the subsequent spoken response by the user.

Alternatively or in addition to this, the word list currently active may also be compiled as a function of recognition results of previous spoken responses within the dialog

5

by the user. One example here would be that the user has already input, in a previous stage of the dialog, that the destination is located in the federal state of North Rhine-Westphalia. For a voice recognition of the user's spoken response to the subsequent input request "In what town is your destination?" it is then sufficient for all names of towns within the federal state

5   of North Rhine-Westphalia to be included in the word list.

Similarly, recognition results of the user's spoken responses subsequently within the dialog may also be used to restrict the active vocabulary for a renewed recognition of spoken responses already made by the user, in order to improve the recognition or to enable it in the first place. One example here would be a dialog in which a prompt "In what

10   town is your destination?" is firstly output. To the extent that the voice recognition cannot then provide a reliable recognition result, for instance because none of the recognition hypotheses has an adequate confidence level, a prompt such as "What large conurbation is in the vicinity?" may be output in the next stage of the dialog. A restricted word list with large conurbations can then be used for the spoken response that follows the second prompt. The

15   recognition result from this query can then serve for the compilation of a word list comprising only towns located in the vicinity of the recognized large conurbation for a further attempt at recognition of the spoken response to the first prompt. This kind of repeat recognition of previous spoken responses stored as audio data can – if desired – also take place as a matter of course.

20   This kind of restriction of the active vocabulary, as a result of which only specific recognition results are permitted, is known as a "hard" restriction.

Alternatively or in addition to this, different recognition hypotheses that have been determined during a voice recognition of a spoken response by the user may be evaluated, using the geographical database, by means of the geographical criteria taken into

25   account in the generation of the previous prompt. An evaluation of this kind can also take place as a function of recognition results of spoken responses by the user previously and/or subsequently within the dialog.

This subsequent evaluation may take place in a "soft" form, in which an "n-best list" is compiled, comprising a particular number "n" of the most probable recognition

30   hypotheses in an ordered sequence. In evaluating the hypotheses for compiling the n-best list, it is ensured that, as regards the geographical criteria, the recognition hypotheses are consistent with previous and/or subsequent recognition results and/or the geographical criteria of the input request. An n-best list of this kind is preferably also generated if the active vocabulary has been restricted previously. Otherwise, an evaluation may also take

6

place according to "hard" exclusion criteria so that the active vocabulary is *quasi* restricted subsequently.

Particularly if the data requested from the user within the dialog does not concern clearly defined facts, such as the name of a federal state, a country etc., but relates to
5   "soft" criteria, such as the size of a town, or if it involves the relationship of different geographical locations to one another, such as the proximity to a large conurbation, then considerations of geographical criteria of these kinds in the form of evaluations of the recognition hypotheses are often more useful than hard exclusion criteria, such as restriction of the active vocabulary.

10   A combined utilization method for the geographical voice recognition criteria within one and the same dialog is also perfectly possible. For instance, for the voice recognition of a spoken response to a first prompt, the active vocabulary may be restricted, and, for the voice recognition of a spoken response to the second prompt, the geographical criteria may be used for evaluating the recognition hypotheses. For the voice recognition of a
15   specific spoken response, both a restriction of the active vocabulary as regards certain criteria, and, in addition, an evaluation of recognition hypotheses on the basis of various other criteria may also take place.

There are also various options for the selection of the dialog sequence itself.

For instance, a dialog may, in principle, be strictly hierarchically structured
20   according to geographical criteria, i.e. prompts hierarchically structured according to geographical criteria are generated within the dialog sequence. A typical example here would be, firstly, an enquiry as to the country in which the destination is located, then, for example, as to the federal state if applicable, then a region, and finally the town, followed by the street, wherein the area is further narrowed down at each stage and, accordingly, only the responses
25   that are possibilities in this area are compiled into word lists.

In an alternative procedure, the input requests as to geographical criteria are generated within the dialog as a function of recognition results of previous spoken responses by the user. For example, an enquiry could be made as to the nearest large conurbation if, in a first step, the recognition result of a response to the enquiry as to the destination was
30   unsatisfactory. If, on the other hand, e.g. the town in which the destination is located has been clearly recognized in the first step, the street may be queried immediately in the next step of the dialog.

When structuring a dialog according to hierarchically structured geographical criteria, it is also possible to use, in addition, recognition results of previous spoken responses

7

in order to determine the further steps within the dialog sequence. A typical example is the case where the federal state "Berlin" has been specified in response to the query as to the federal state in which the destination is located. In the subsequent input request, rather than asking for the town within the federal state, it would be more useful to enquire about e.g. the

5   administrative district of the town in which the destination is located.

With both methods, there preferably exists the possibility that – if the user cannot answer the particular question, e.g. the question as to an administrative district of the town of Berlin or about a particular region within the federal state – the dialog step can be skipped by means of an appropriate response, such as "unknown", or can be replaced with a

10   different enquiry that also narrows down the area.

The geographical database used within the dialog, e.g. for compiling a word list and/or for evaluating a recognition hypothesis, is preferably restricted, as far as possible, on the basis of a previous input request and/or a previous recognition result of a spoken response to certain data entries. By restricting the database in previous steps for the

15   subsequent steps, an extraction of the appropriate word list can take place considerably more quickly, since the number of data entries that have to be searched in order to compile the word list is correspondingly smaller.

Further, the use of a geographical database with data entries that each have assigned to them one or more markers representing a type of the data entry concerned is

20   particularly preferred. A geographical type of data entry would be, for example, whether the data entry concerned relates to a country, a federal state, a town or a large conurbation, or also, in which federal state a town is located, etc. The markers may also represent a geographical hierarchy level. Using these markers, a restriction of the database for further steps can be accomplished considerably faster, and/or a word list can be extracted more

25   quickly or post-processed more effectively, since searching is restricted to entries with specific markers, wherein the type of marker, e.g. the current hierarchy level or the currently queried geographical type, is defined for a recognition or evaluation of a specific spoken response by the previous prompt or dialog stage.

The dialog control device, the prompt generation device, the voice recognition

30   device and the analysis unit may each be software components implemented on suitable hardware, for instance a processor of a voice-data user interface of a navigation system. It is hereby not absolutely essential for the voice-data user interface to be equipped with its own processor for this purpose, instead the voice-data user interface may also share the use of a processor that is also used for the remaining functions of the navigation system. It is, in

8

particular, pointed out in this connection that the navigation system does not necessarily have to be a structural unit, but that it is also possible to realize various components of the system on various items of equipment that are connected to one another. This applies, in particular, to the voice-data user interface itself, the components of which may also be implemented on

5   spatially separated processors. Thus, for instance, a voice recognition device could be implemented with an analysis unit on a particularly high-capacity server on the Internet, and be connected via a data connection to the further components of the navigation system, located in, for instance, a motor vehicle of the user.

Since the dialog control device, the prompt generation device, the voice

10 recognition device and the analysis unit may each take the form of software modules, it is possible to retrofit existing navigation systems equipped with an appropriate output device, such as a loudspeaker and/or a display, with the voice-data user interface in accordance with the invention. All that is necessary is for the system to be equipped with a voice input unit, e.g. a simple microphone, and for an appropriately powerful processor to be available or for

15 appropriate connections to a powerful processor to exist. Access to a geographical database exists *per se* in a navigation system, since this requires a database for calculating routes. The database may be stored on a mass storage device located on the navigation system, such as a CD. However, it may also be interrogated via a network, e.g. the Internet.

The database should preferably be modified in advance as part of a pre-

20 processing stage. For example, the markers, which represent the type of the database entry concerned and/or a geographical hierarchy level and/or other geographical features, e.g. the position on a river, etc., may hereby be assigned to the respective individual data entries. In addition, the database may also be sorted hierarchically and/or relationships between the individual database entries may be established, and geographical criteria thereby determined.

25 Geographical criteria of this kind may be stored at a separate location in the database or be contained directly in the database entries. For instance, a database entry "Eilendorf near Aachen" simultaneously also contains the relationship between two towns. With a database of this kind, prepared for use in accordance with the invention, a navigation system in accordance with the invention becomes more fast and efficient.

30

The invention will be further described with reference to examples of embodiments shown in the drawings, to which, however, the invention is not restricted.

9

Fig. 1 shows a schematic diagram of the system architecture of one

embodiment example of a navigation system in accordance with the invention.

Fig. 2 shows a dialog block diagram to explain one possible dialog sequence

5   between a user and the system in accordance with the invention.

In principle, the navigation system 1 shown in Fig. 1 may be a conventional

navigation system equipped with all components that are normally components of a

10   navigation system in order to guarantee the necessary functionalities. For the sake of

simplicity, these components of the navigation system 1 are shown here in just one single

block 13. For communication with the user, the navigation system 1 is equipped with a

voice-data user interface 2 in accordance with the invention, the components of which are

shown in greater detail in Fig. 1.

15   One component of this voice-data user interface 2 is an input/output interface

10, to which a voice inputting device 11, here a microphone 11, and a voice output device 12,

here a loudspeaker 12, are connected. Via the microphone 11, the voice-data user interface 2

can detect spoken responses S by the user. Via the loudspeaker 12, the voice-data user

interface 2 can output prompts P, e.g. in order to induce the user to make a spoken response

20   S.

One further important component of the user interface 2 is a voice recognition

device 6, which pre-processes the incoming spoken responses S, processes them and supplies

recognition hypotheses EH at an output. These recognition hypotheses EH are then further

processed in an analysis unit 7 so that the contents of the spoken responses – for instance

25   commands or location details – can be understood.

The voice-data user interface 2 is further equipped with a prompt generation

unit 5 by means of which the prompts P to be output to the user are generated. Responsible

for controlling the dialog between the voice-data user interface 2 and the user by means of

the prompts P output by the system, and taking account of the spoken responses S by the user

30   input in response, is a dialog control device 3 (also referred to below as the dialog manager

3), which controls the dialog in accordance with a predetermined dialog program. To this

end, the dialog control device 3 is connected to the prompt generation unit 5, the voice

recognition device 6, the analysis unit 7 and the input/output interface 10. As a result, the

dialog control device 3 may transmit, for example, a prompt generation command PB to the

10

prompt generation device 5, thereby inducing it to output a specific prompt P. As soon as a spoken response S is detected by the microphone, the dialog control device 3 is informed via the input/output interface 10, and sends a start command AS to the voice recognition device 6 and the analysis unit 7.

5   A further component of the voice-data user interface 2 that is significant for the invention is a geographical database 8. This database 8 is represented here as a component integral to the voice-data user interface 2. This may, however, also perfectly well be a general geographical database of the navigation system 1, which is used – possibly only in part – by, among others, the voice-data user interface 2 of the navigation system 1.

10   This operating mode of a voice-data user interface 2 structured in this manner is explained below by reference to the dialog block diagram shown in Fig. 2:

          A dialog generally begins – following a normal activation, e.g. by a voice command or by manual operation of the equipment – by the dialog manager 3 outputting a prompt output command PB to the prompt generator 5 in order for a particular prompt P to be

15   output to the user. The generation of this prompt P takes account of specific geographical criteria GK, which are predetermined in the dialog program or which the dialog manager 3 can retrieve from the geographical database 8.

          In this database 8 are located data entries DE, e.g. the names and further geographical data on countries, regions, federal states, towns, streets, significant sights, full

20   addresses, etc. The data entries DE may hereby be entered in different ways into the database 8. For instance, the individual data entries DE may each contain markers M, which indicate the geographical category or the type to which the data entry DE is assigned, such as <country>, <federal state>, <town>, <administrative district of a town>, etc. or <small town>, <large conurbation>, <town of over 1 million inhabitants>, etc. As an alternative or in

25   addition, the database may also be hierarchically organized and/or divided into different parts. For a territory such as Germany for instance, different partial databases for the individual federal states could be available, in which, in turn, the towns are entered. Hierarchically arranged under the towns are the administrative districts of the towns and then, under the individual administrative districts of the towns, the street names, etc.

30   In addition, certain geographical criteria, such as relationships between the individual data entries DE, e.g. the proximity of two towns to one another, may also be stored in the database 8. In particular, the database 8 may exhibit an area which records which geographical criteria can be determined from the database without great effort, or for which geographical criteria ready-prepared data records are available.

11

Simultaneously with the prompt output command PB, the dialog manager 3 outputs a list compilation command LB to a word list generator 9, which, according to the geographical criteria currently being sought, retrieves the data entries DE from the geographical database 8 and compiles from these a word list WL, which comprises the active

5   vocabulary for a voice recognition of a subsequent spoken response S by the user. In addition, the dialog manager 3 transmits a start command AS to the voice recognition device 6 and the analysis device 7, which are shown here as one block. The word list generator 9 may be a separate module. However, it may also be a sub-routine of the voice recognition device 6, as shown in Fig. 1 by way of example.

10   The voice recognition device 6 then determines evaluated hypotheses for the spoken response S following the prompt P, wherein the spoken responses are each compared with the stored acoustic models of the words contained in the word list WL compiled by the word list generator 9. Since this is a relatively restricted word list WL, a higher recognition probability is possible than with a complete word list of all geographical proper names.

15   The best-evaluated recognition result EE or multiple recognition hypotheses EH are then checked again by the analysis device 7, where applicable, for consistency with the data entries DE in the geographical database 8 and/or with previous recognition results and with the preceding prompt. To this end, the analysis device 7 where applicable retrieves consistency check data KCD from the database 8. If a recognition result EE is certain, the

20   database 8 is, where applicable, restricted for the further course of the dialog if, for instance, it can be reliably ruled out, on the basis of the recognition result EE or the hypotheses EH, that certain data entries DE in the database 8 can no longer occur in the case of subsequent spoken responses. So, for example, given a reliable recognition of the words "Lower Saxony" in response to the input request "Please enter the federal state in which your destination is

25   located", all location data in other federal states can be dispensed with for the following steps of the dialog.

The recognition result EE is also reported back to the dialog manager 3 and is there entered in a "slot filling module" 4, which documents the current overall state of knowledge of the system. This slot filling module 4 of the dialog manager 3 decides when the

30   information is sufficient, i.e. when all query points have been clarified, in order that, for example, the destination or the start point can be precisely defined. If the information is not yet sufficient, a further dialog step follows, in which a prompt output command PB is again output to the prompt generator 5, a list compilation command LB to the word list generator 9 and a start signal AS to the voice recognition system 6, in order for the next spoken response

12

to be recognized. In this step, only the previously restricted database 8 is then used, so the entire system can operate considerably faster in the following dialog steps.

If the slot filling module 4 establishes that all necessary information is present, the prompt generator 5 is induced to issue a corresponding prompt confirming the desired

5   destination, and the destination is transmitted to the further components (here once again shown as a block 13) of the navigation system 1 for further processing.

This sequence is described more specifically below with reference to two examples:

In the first example, it is assumed that the dialog sequence is hierarchically

10   structured according to geographical criteria. In this case, a prompt, for example "In what country is your destination?" is output by the prompt generator 5 in a first step. As only the names of different countries are expected as the spoken response to this input request, a word list with the possible countries is generated by the word list generator 9 on the basis of the database 8. This word list is then available during the voice recognition for the subsequent

15   spoken response. This is the first hierarchical level of the dialog. Once the matching country has been correctly recognized, e.g. if the country "Germany" has been stated, the prompt "In which federal state is your destination?" is then generated in a second hierarchical level. A word list is then compiled, listing all the federal states in Germany. Then, in a further hierarchical level, the town is queried or, if applicable, a particular region is queried in an

20   intervening hierarchical step. Once a town has been established, the administrative district may be queried in the case of larger towns, and finally, at one of the lower stages, the street name and a house number, or a particular building, etc.

In the second embodiment example, it is assumed that the individual database entries in the database 8 are provided with markers, which represent particular types of

25   database entries or particular relationships between the database entries. With this variant, the dialog sequence is not strictly hierarchically structured from large down to small geographical units *per se*, being relatively flexible. A dialog sequence of this kind may, in certain circumstances, i.e. under good recognition conditions, reach the destination in fewer steps than a dialog sequence with a strictly hierarchical structure. In this case, the dialog

30   control unit 3 firstly selects, for example, a prompt "To which town do you want to travel?". Then, if applicable, a word list with all town entries available in the database 8 will be compiled. To the extent that no further restrictions have been undertaken previously, this will, of course, be a relatively long list. In other words, the active vocabulary of the voice recognition system is extremely broad, making voice recognition considerably harder than if

13

the word list had previously been restricted by prior queries as to countries, federal states, etc. If the voice recognition system yields an acceptable recognition result, the town of the destination would have been clarified with just one query and the inputting of the further address data, such as street and house number, could then take place in a subsequent dialog

5   step.

If, however, the system is not sure of the result, e.g. because a calculated confidence level of the various recognition hypotheses is insufficient, this fact may be delivered back to the dialog control device 3 as a (preliminary) recognition result. The latter could then, in a subsequent dialog step, output a further prompt, e.g. "What large conurbation

10   is in the vicinity?". An active word list restricted to large conurbations will then be compiled. This is possible relatively easily in that all data entries DE containing the marker <large conurbation> are searched for from the database 8. This word list is considerably smaller than the word list in the first dialog step, so, owing to the smaller active vocabulary, the recognition result EE will be better in the case of this second query than for the first.

15   Using the recognition result EE, all data entries DE in the database 8 that are located in the vicinity of the large conurbation sought can then be extracted. If applicable, all data entries DE that fulfil the condition of being located near the recognized large conurbation may also be marked in a first step. The new word list is then compiled, containing all towns that fulfil the condition. If the spoken response of the user to the

20   previous query as to the desired town has been stored, it is now possible to undertake a voice recognition for this first spoken response once again with the restricted word list in order to arrive at a better recognition result. Alternatively, the dialog manager 3 may also induce the prompt generation device 5 to output the first prompt "To which town do you want to travel?" once again and then to undertake the voice recognition of the subsequent spoken

25   response with the restricted word list.

In conclusion, it is pointed out once again that the invention is not restricted to the above-described embodiment examples – in particular the precise structure of the voice-data user interface or the precise sequence of the explained dialogs – but may be varied to a large degree by a person skilled in the field without exceeding the bounds of the invention.

30   In particular, it is possible to call upon further criteria, especially supplementary geographical knowledge, for the voice recognition. For example, distances from the current location may also be taken into account in the evaluation of recognition hypotheses and/or the compilation of word lists. The number of times a destination has been visited hitherto by the particular user may also be taken into account, since, in many cases, a

14

user makes frequent journeys to the same locations. It is also possible for multiple queries to be covered by one prompt, e.g. "State the desired destination and a large town in the vicinity". In the subsequent voice recognition, the geographical relationships of various towns can then be used in order that the recognition hypotheses are better evaluated.

5   Furthermore, the slot filling module could, for example, be arranged in the analysis device rather than in the dialog manager.

The invention has been described largely with reference to examples in which a destination has to be determined. In the same way, however, the starting point or other location data, such as intermediate stops etc., may also be determined in a dialog between the

10   user and the system. In other words, multiple similar dialogs may be conducted successively.

For the sake of completeness, it is also pointed out that the use of the indefinite article "a" or "an" does not exclude the possibility that the features in question may also be present more than once, and that the use of the term "comprehensive" does not preclude the existence of further elements or steps.

15

CLAIMS:

1.   A method of operating a voice-controlled navigation system (1) in which, within an automatically conducted dialog:

- input requests (P) are generated, taking account of geographical criteria (GK), and output to a user,

- responses spoken by the user (S) are detected,

- the spoken responses (S) are analyzed for recognition of location data using an automatic voice recognition method, taking account of the geographical criteria (GK).

2.   A method as claimed in claim 1, characterized in that, using the geographical criteria (GK) taken into account in the generation of the input request (P), and/or as a function of recognition results (EE) of previous and/or subsequent spoken responses (S) within the dialog by the user, a word list (WL) is compiled from a geographical database 8 to serve as an active vocabulary during the voice recognition of a subsequent spoken response by the user.

3.   A method as claimed in claim 1 or 2, characterized in that recognition hypotheses (EH) that have been determined during a voice recognition of a spoken response (S) by the user are evaluated, using a geographical database (8), by means of the geographical criteria (GK) taken into account in the generation of a previous input request (P) and/or as a function of recognition results (EE) of spoken responses (S) by the user previously and/or subsequently within the dialog.

4.   A method as claimed in any one of claims 1 to 3, characterized in that input requests (P) hierarchically structured according to geographical criteria are generated within the dialog.

5.   A method as claimed in any one of claims 1 to 4, characterized in that the input requests (P) concerning geographical criteria are generated within the dialog as a function of recognition results of previous spoken responses by the user.

16

6.   A method as claimed in any one of claims 1 to 5, characterized in that a geographical database (8) is used within the dialog for compiling a word list (WL) and/or for evaluating a recognition hypothesis, which database has been restricted on the basis of a
5   previous input request (P) and/or a previous recognition result (EE) of a spoken response to certain data entries (DE).

7.   A method as claimed in any one of claims 2 to 6, characterized in that, in order to compile a word list (WL) and/or to evaluate a recognition hypothesis (EH), a geographical
10   database (8) is used, with data entries (DE) that each have assigned to them one or more markers (M) representing a type of the data entry (DE) concerned and/or a geographical hierarchy level and/or a relationship to other data entries (DE) and/or other geographical features.

15   8.   A voice-data user interface (2) for a navigation system (1) with:
            - an output device (12) for outputting input requests (P) to a user,
            - a voice-inputting device (11) for detecting spoken responses (S) by the user,
            - a dialog-control device (3) for controlling a dialog with the user, taking
account of geographical criteria (GK),
20   - a prompt-generation unit (5) for generating input requests (P), taking account
of geographical criteria (GK),
            - a voice recognition device (6) and an analysis unit (7) for analyzing the
detected spoken responses (S) for the recognition of location data, taking account of
geographical criteria (GK),
25   - a geographical database (8) and/or a data interface for accessing a
geographical database, which makes available geographical criteria (GK) and/or geographical
data for the dialog-control device (3) and/or the prompt-generation device (5) and/or the
voice recognition device (6) and/or the analysis unit (7).

30   9.   A navigation system (1) with a voice-data user interface (2) as claimed in
claim 8.

17

10.   A computer program with program-coding means in order to execute all steps
of the method as claimed in any one of claims 1 to 7 when the program is executed on a
processor of a voice-data user interface of a navigation system.

5   11.   A method of generating a geographical database (8) for use in a method as
claimed in any one of claims 1 to 7, in which the individual data entries (DE) each have
assigned to them one or more markers (M) representing a type and/or a relationship to other
data entries (DE) and/or a geographical hierarchy level and/or other geographical features of
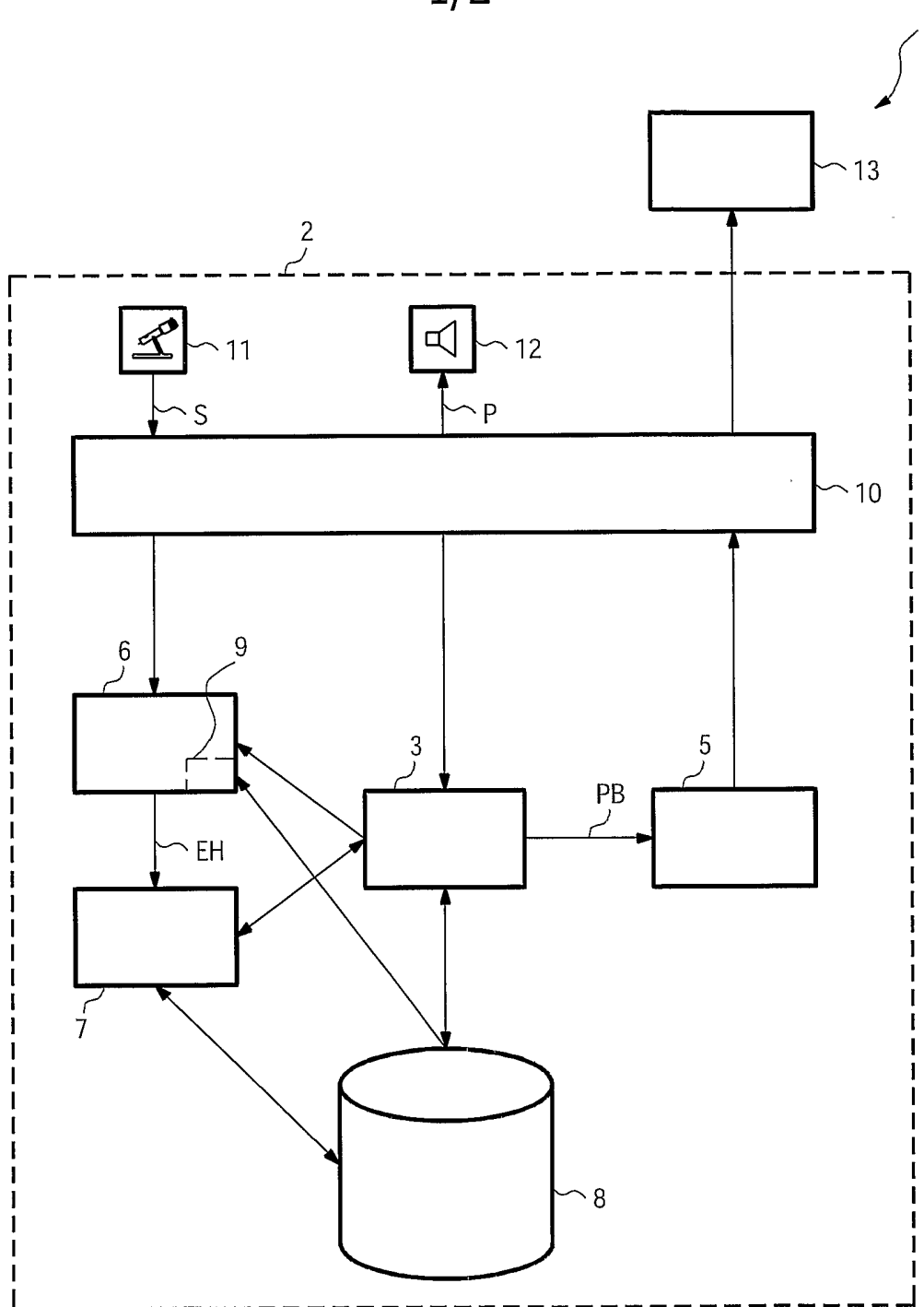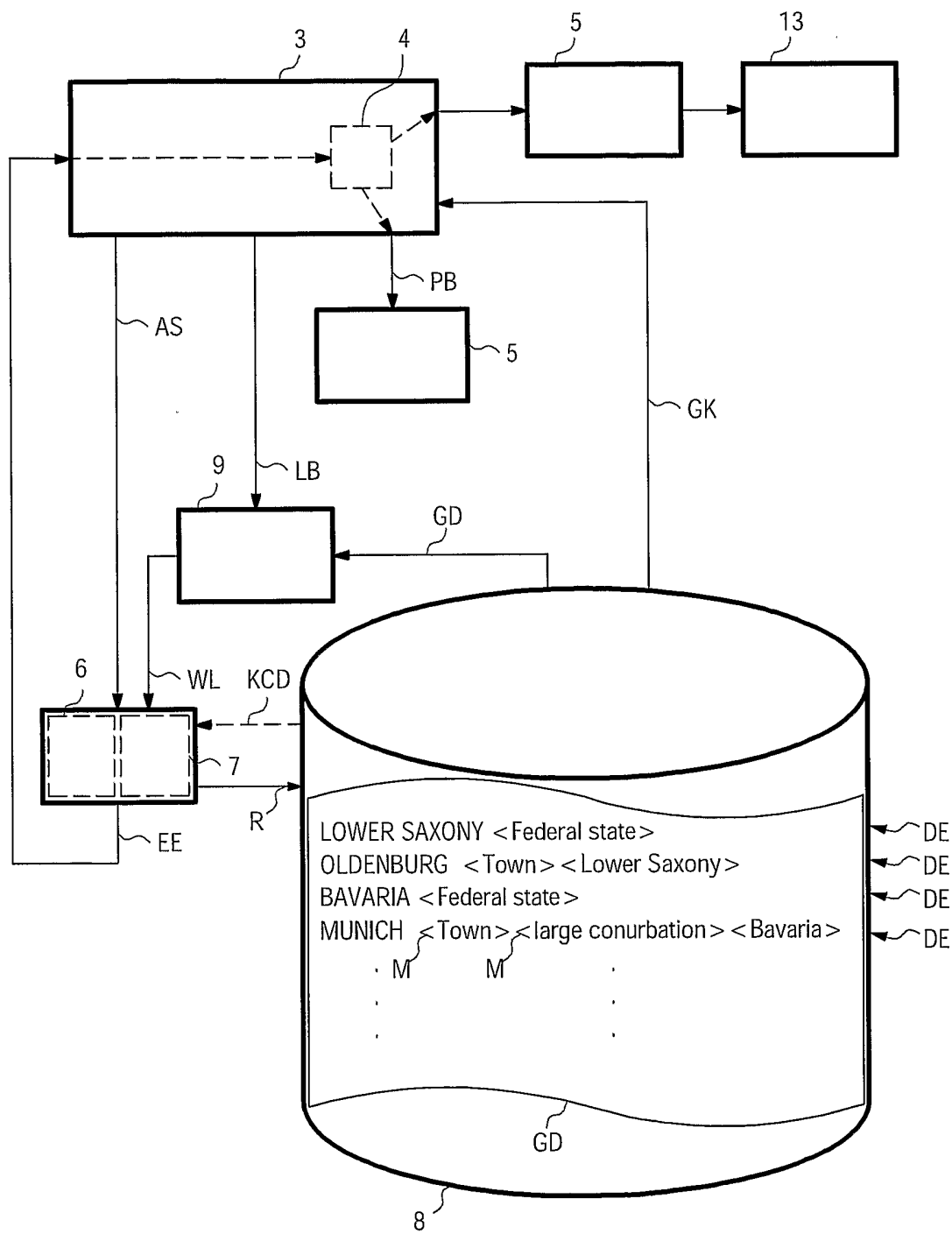the data entry (DE) concerned.

1/2

FIG. 1

2/2

FIG. 2

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7   G01C21/36   G10L15/22

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7   G01C   G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 230 132 B1 (CLASS ET AL) 8 May 2001 (2001-05-08) column 7, line 60 - column 8, line 23 column 3, line 14 - line 45 column 10, line 39 - column 11, line 60 column 13, line 5 - line 67; figures 2,7 -----	1-11
X	DE 199 62 048 A (ZUENDORF DETLEF) 12 July 2001 (2001-07-12) column 2, line 1 - line 36 -----	1,2,8-11
A	EP 1 233 407 A (NAVIGATION TECH CORP) 21 August 2002 (2002-08-21) abstract -----	1
A	EP 1 298 415 A (BOSCH GMBH ROBERT) 2 April 2003 (2003-04-02) paragraphs '0006! - '0008! -----	1

☐ Further documents are listed in the continuation of box C.   ☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
1 September 2004	08/09/2004

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Hoekstra, F

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 6230132	B1	08-05-2001	DE	19709518 C1	05-03-1998
			CA	2231851 A1	10-09-1998
			EP	0865014 A2	16-09-1998
			JP	11015493 A	22-01-1999
DE 19962048	A	12-07-2001	DE	19962048 A1	12-07-2001
EP 1233407	A	21-08-2002	US	2002111810 A1	15-08-2002
			EP	1233407 A1	21-08-2002
			JP	2002318136 A	31-10-2002
EP 1298415	A	02-04-2003	DE	10147734 A1	10-04-2003
			EP	1298415 A2	02-04-2003
			US	2003125941 A1	03-07-2003