FIG. 1

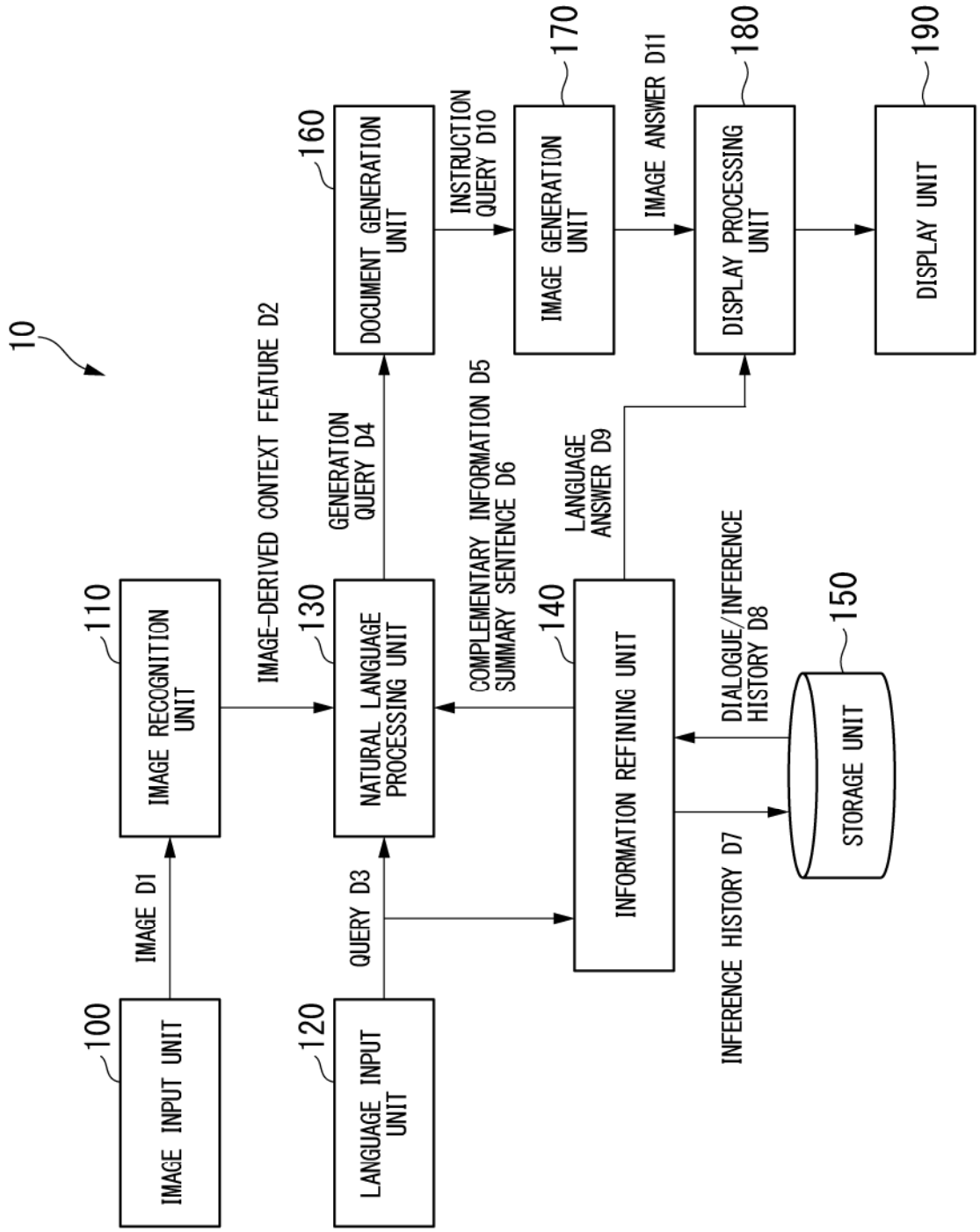

10
100
IMAGE INPUT UNIT
110
IMAGE RECOGNITION UNIT
120
LANGUAGE INPUT UNIT
130
NATURAL LANGUAGE PROCESSING UNIT
140
INFORMATION REFINING UNIT
150
STORAGE UNIT
160
DOCUMENT GENERATION UNIT
170
IMAGE GENERATION UNIT
180
DISPLAY PROCESSING UNIT
190
DISPLAY UNIT
IMAGE D1
QUERY D3
IMAGE-DERIVED CONTEXT FEATURE D2
GENERATION QUERY D4
COMPLEMENTARY INFORMATION D5
SUMMARY SENTENCE D6
INFERENCE HISTORY D7
DIALOGUE/INFERENCE HISTORY D8
LANGUAGE ANSWER D9
INSTRUCTION QUERY D10
IMAGE ANSWER D11

FIG. 2

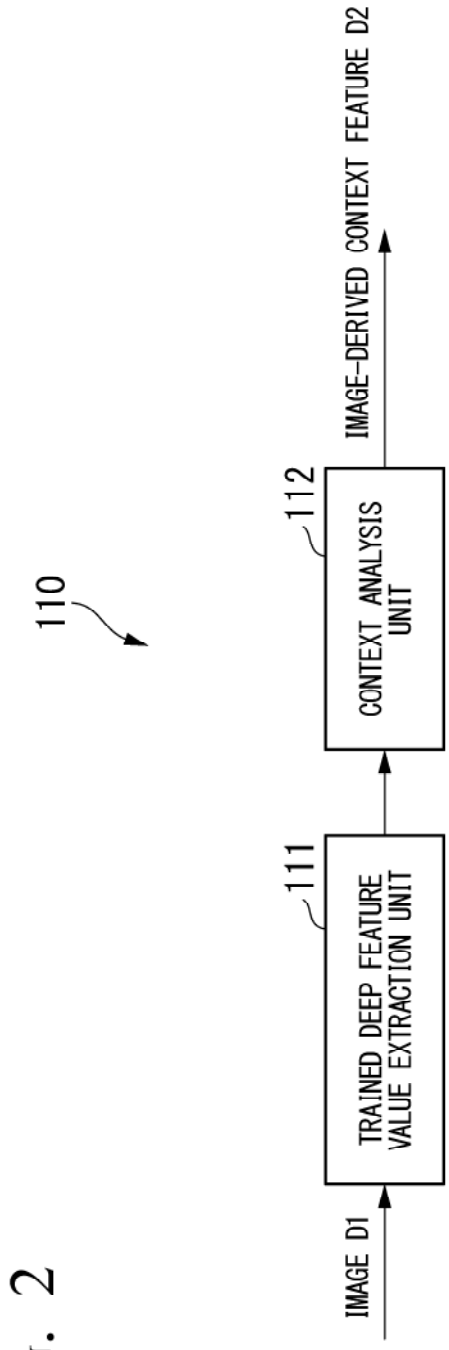

110
111
TRAINED DEEP FEATURE VALUE EXTRACTION UNIT
112
CONTEXT ANALYSIS UNIT
IMAGE D1
IMAGE-DERIVED CONTEXT FEATURE D2

FIG. 3

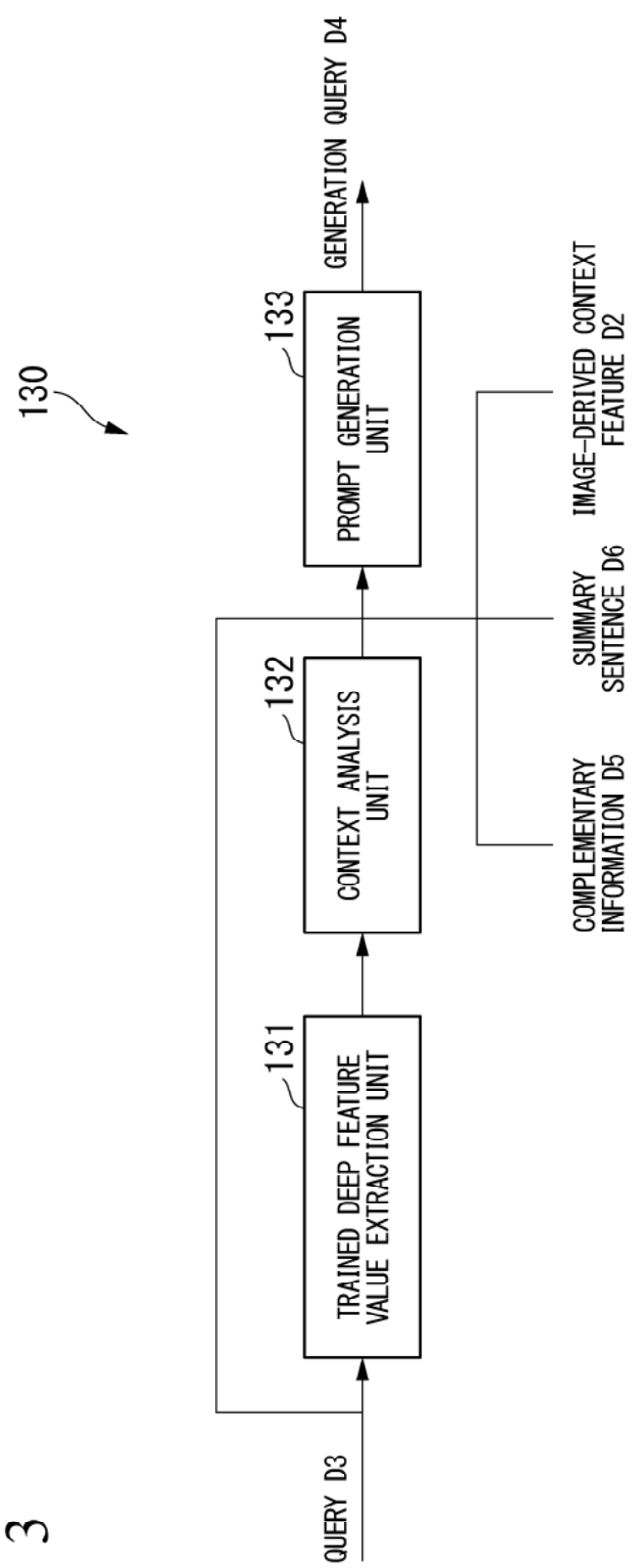

130
131
TRAINED DEEP FEATURE VALUE EXTRACTION UNIT
132
CONTEXT ANALYSIS UNIT
133
PROMPT GENERATION UNIT
QUERY D3
COMPLEMENTARY INFORMATION D5
SUMMARY SENTENCE D6
IMAGE-DERIVED CONTEXT FEATURE D2
GENERATION QUERY D4

FIG. 4

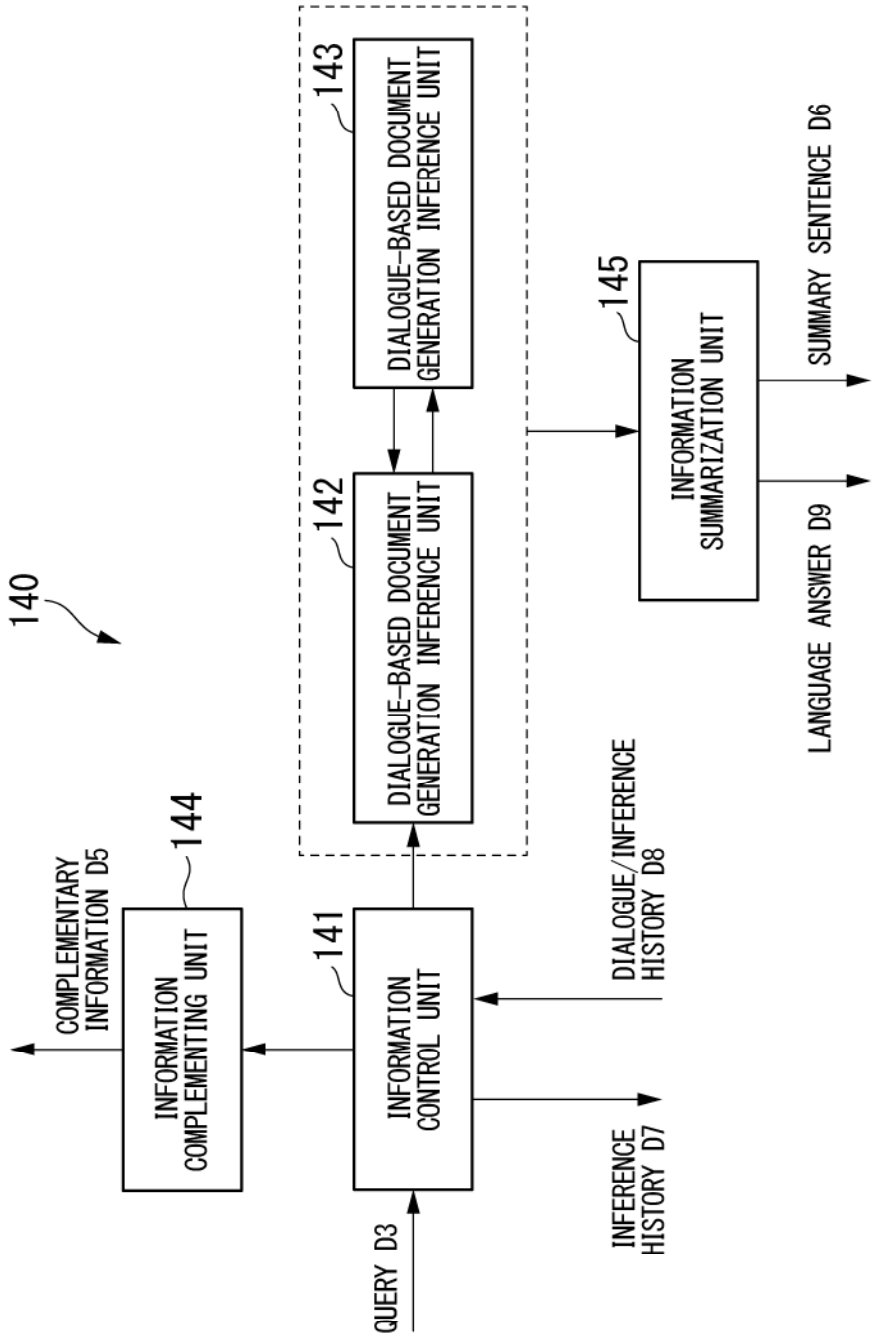

140
141
142
143
144
145
INFORMATION CONTROL UNIT
INFORMATION COMPLEMENTING UNIT
DIALOGUE-BASED DOCUMENT GENERATION INFERENCE UNIT
DIALOGUE-BASED DOCUMENT GENERATION INFERENCE UNIT
INFORMATION SUMMARIZATION UNIT
QUERY D3
COMPLEMENTARY INFORMATION D5
INFERENCE HISTORY D7
DIALOGUE/INFERENCE HISTORY D8
SUMMARY SENTENCE D6
LANGUAGE ANSWER D9

FIG. 5

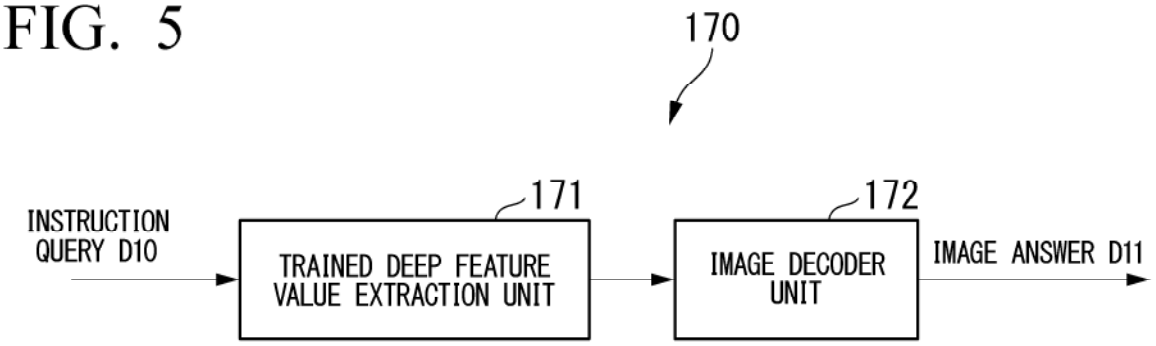

170
171
172
INSTRUCTION QUERY D10
TRAINED DEEP FEATURE VALUE EXTRACTION UNIT
IMAGE DECODER UNIT
IMAGE ANSWER D11

FIG. 6

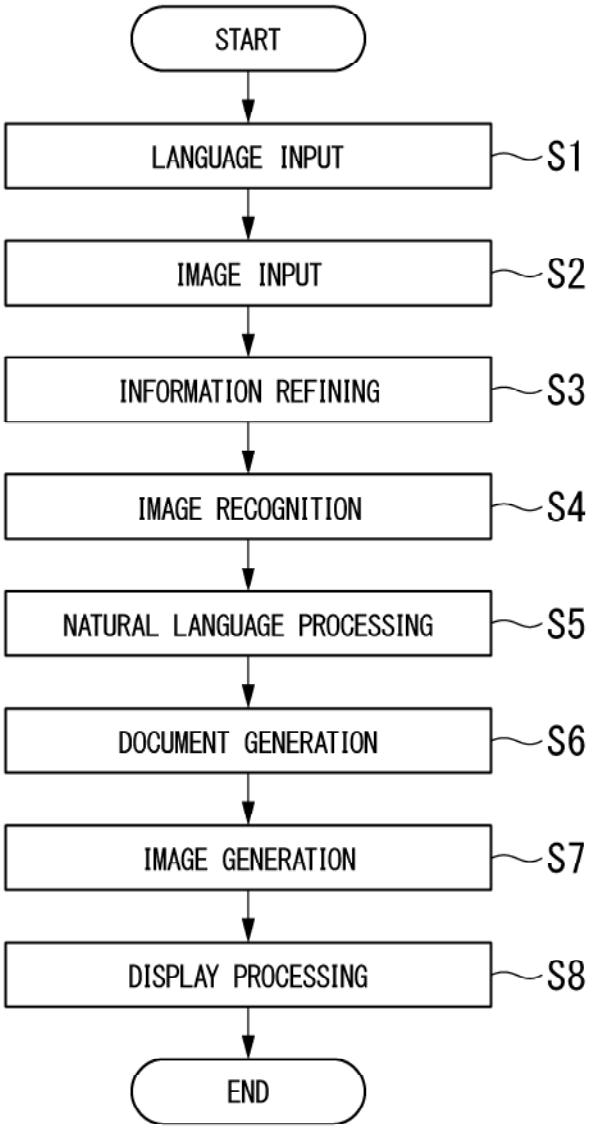

START
LANGUAGE INPUT
S1
IMAGE INPUT
S2
INFORMATION REFINING
S3
IMAGE RECOGNITION
S4
NATURAL LANGUAGE PROCESSING
S5
DOCUMENT GENERATION
S6
IMAGE GENERATION
S7
DISPLAY PROCESSING
S8
END

FIG. 7

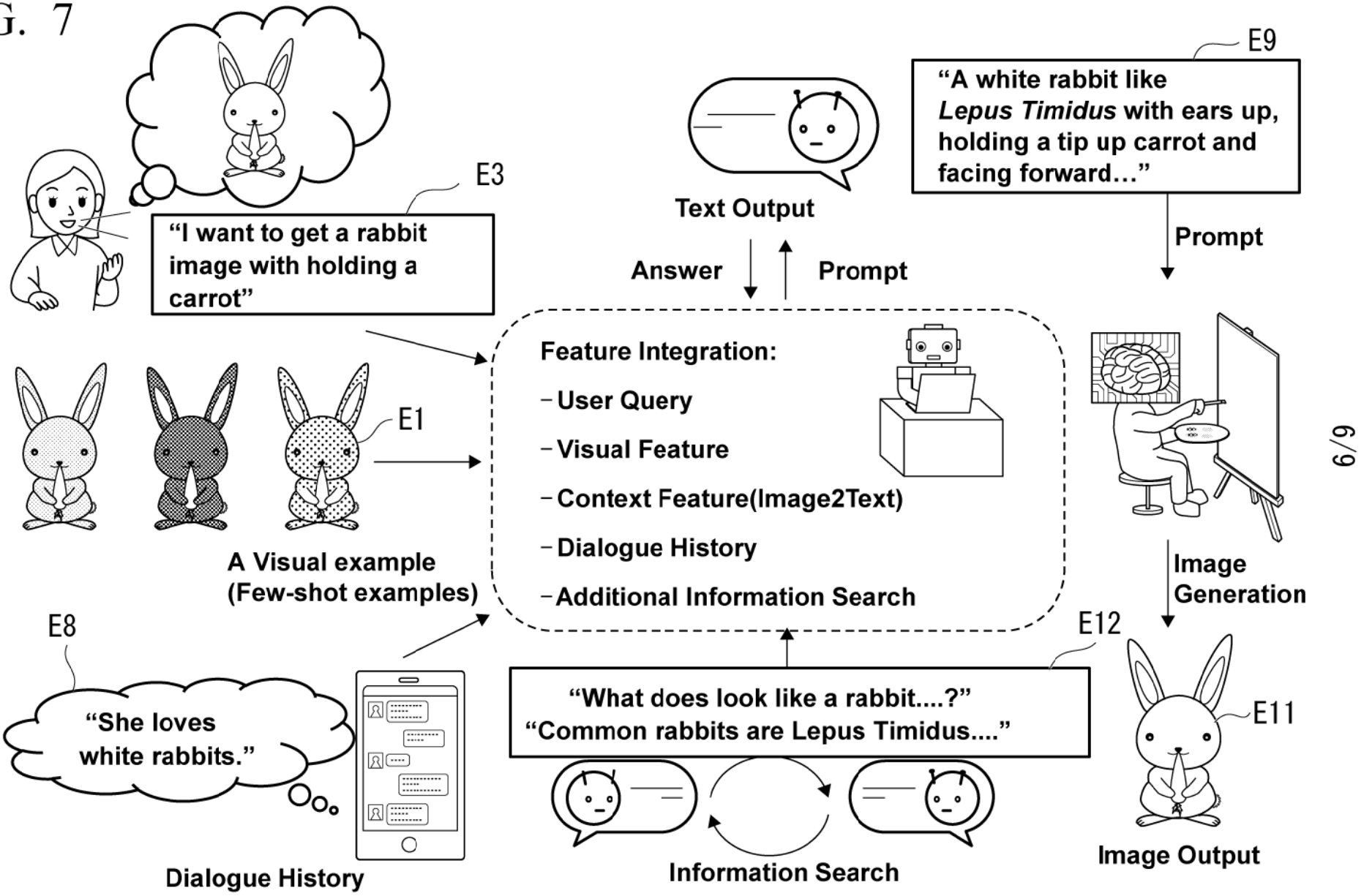

E3
"I want to get a rabbit image with holding a carrot"
Text Output
E9
"A white rabbit like *Lepus Timidus* with ears up, holding a tip up carrot and facing forward…"
Prompt
Answer
Prompt
Feature Integration:
- User Query
- Visual Feature
- Context Feature(Image2Text)
- Dialogue History
- Additional Information Search
E1
A Visual example
(Few-shot examples)
Image Generation
E8
"She loves white rabbits."
Dialogue History
E12
"What does look like a rabbit....?"
"Common rabbits are Lepus Timidus...."
Information Search
E11
Image Output

FIG. 8

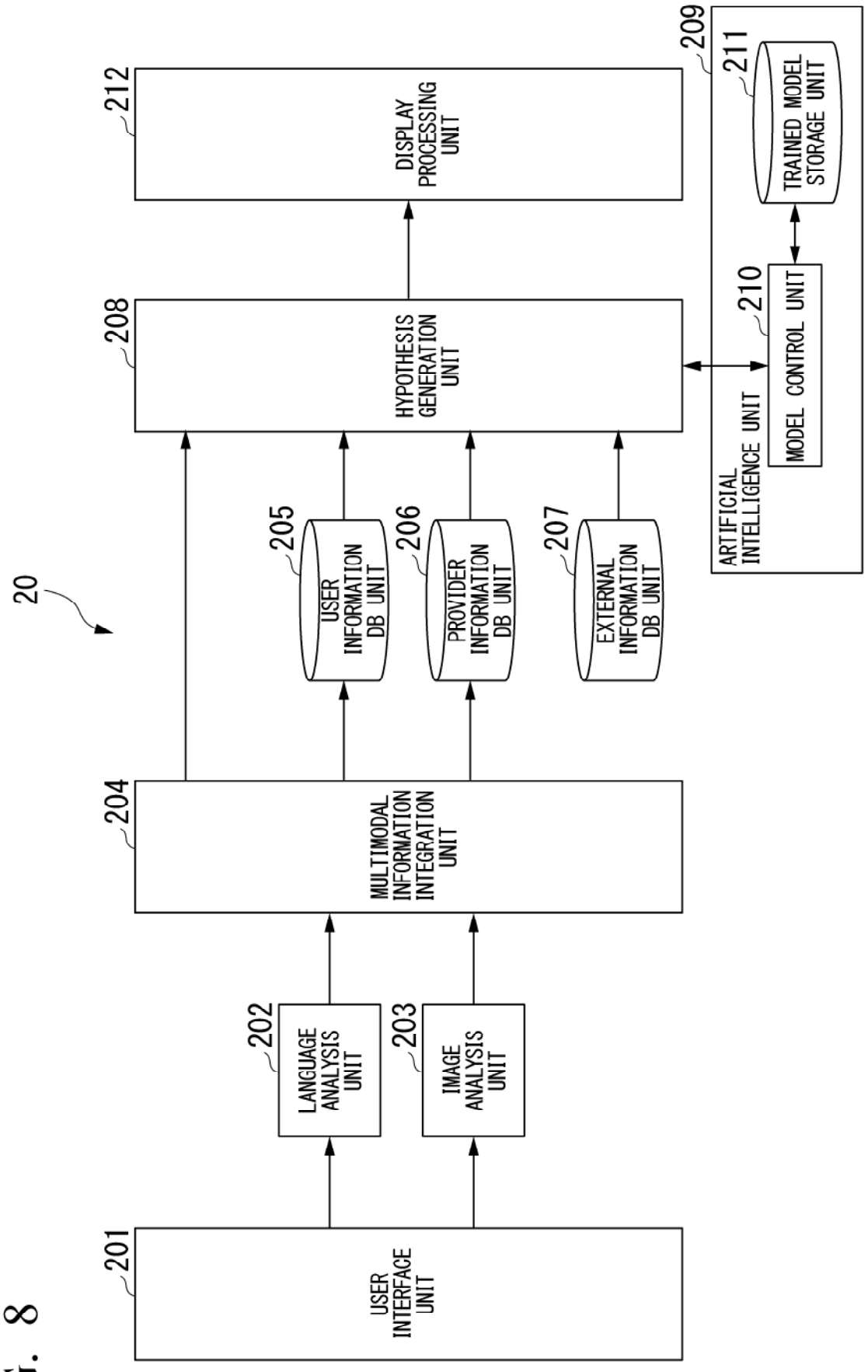

20
201
USER INTERFACE UNIT
202
LANGUAGE ANALYSIS UNIT
203
IMAGE ANALYSIS UNIT
204
MULTIMODAL INFORMATION INTEGRATION UNIT
205
USER INFORMATION DB UNIT
206
PROVIDER INFORMATION DB UNIT
207
EXTERNAL INFORMATION DB UNIT
208
HYPOTHESIS GENERATION UNIT
209
ARTIFICIAL INTELLIGENCE UNIT
210
MODEL CONTROL UNIT
211
TRAINED MODEL STORAGE UNIT
212
DISPLAY PROCESSING UNIT

FIG. 9

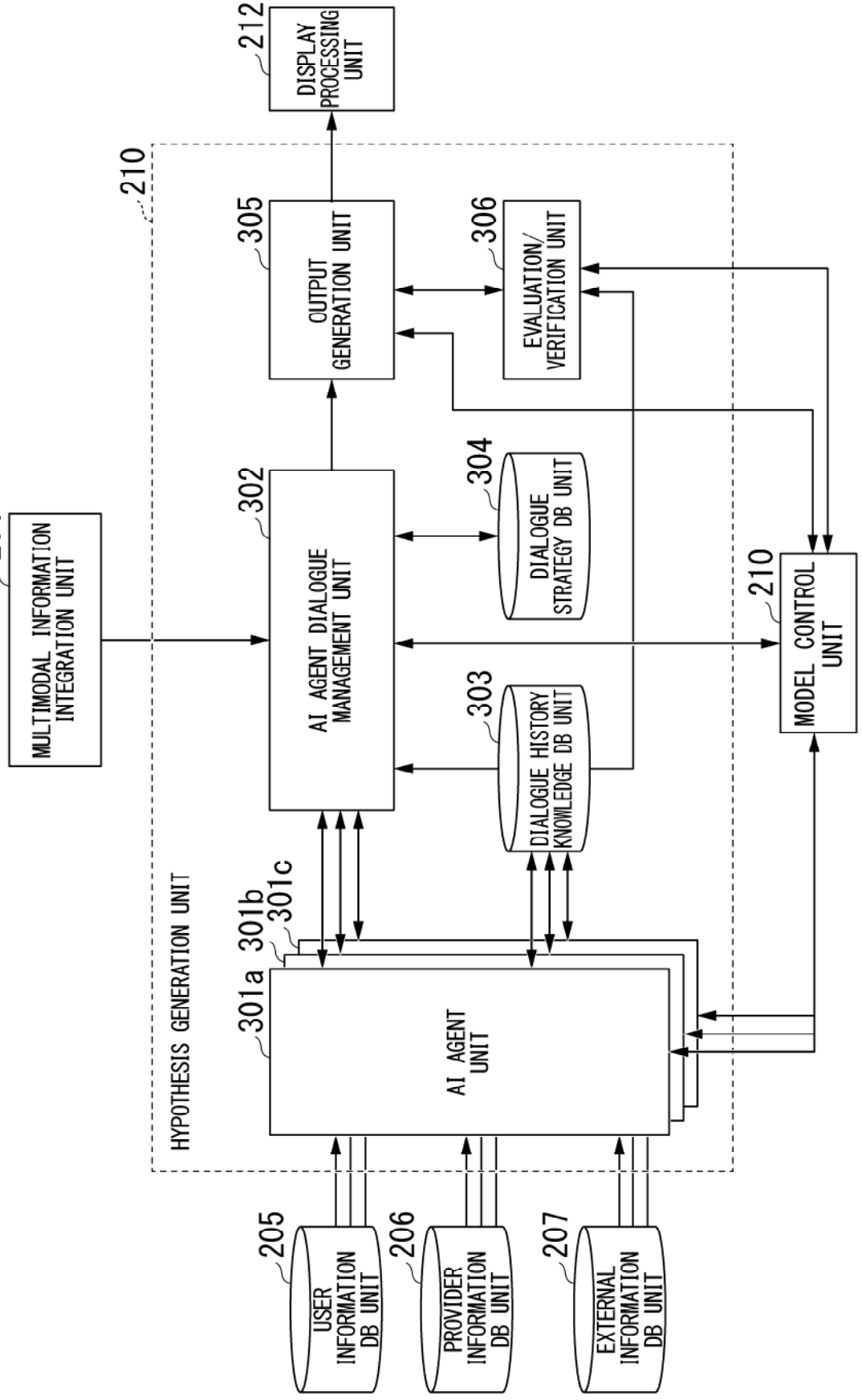
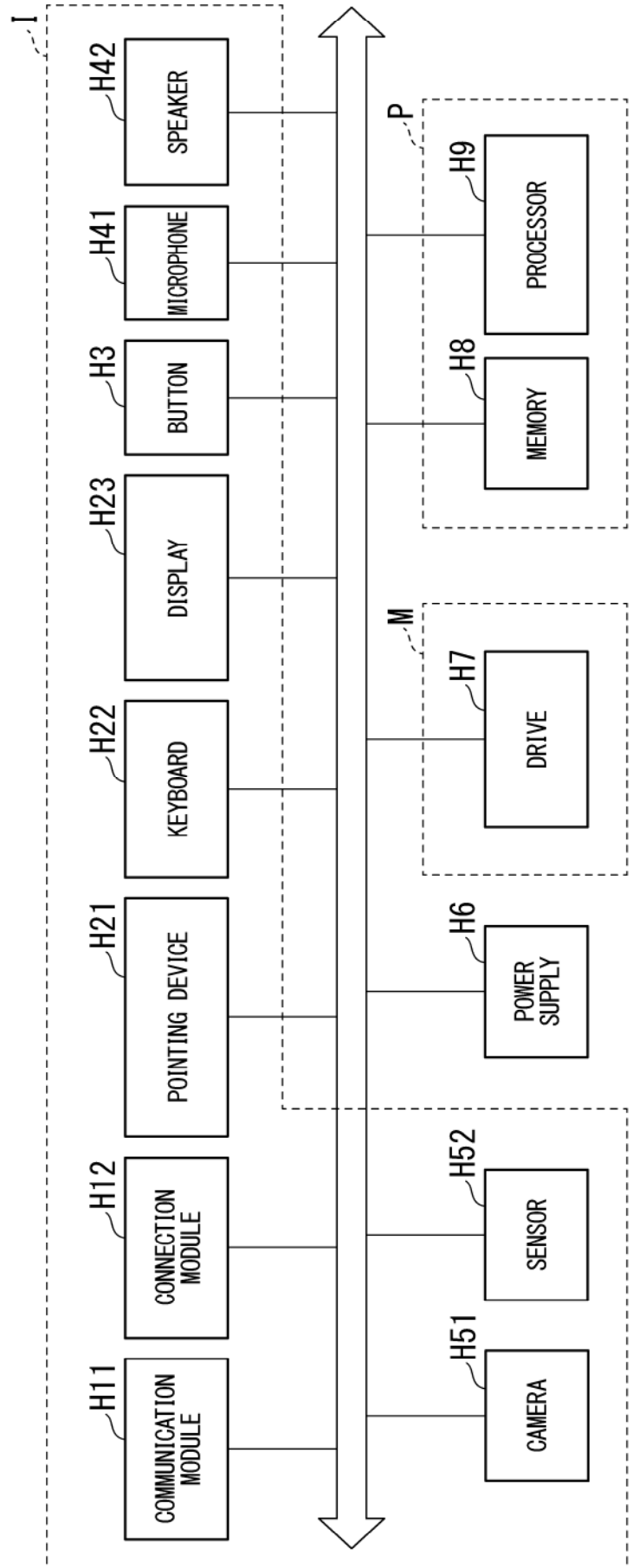

212
DISPLAY PROCESSING UNIT
210
305
OUTPUT GENERATION UNIT
306
EVALUATION/ VERIFICATION UNIT
204
MULTIMODAL INFORMATION INTEGRATION UNIT
302
AI AGENT DIALOGUE MANAGEMENT UNIT
304
DIALOGUE STRATEGY DB UNIT
210
MODEL CONTROL UNIT
303
DIALOGUE HISTORY KNOWLEDGE DB UNIT
HYPOTHESIS GENERATION UNIT
301a
301b
301c
AI AGENT UNIT
205
USER INFORMATION DB UNIT
206
PROVIDER INFORMATION DB UNIT
207
EXTERNAL INFORMATION DB UNIT

FIG. 10


I
H42
SPEAKER
H41
MICROPHONE
H3
BUTTON
H23
DISPLAY
H22
KEYBOARD
H21
POINTING DEVICE
H12
CONNECTION MODULE
H11
COMMUNICATION MODULE
P
H9
PROCESSOR
H8
MEMORY
M
H7
DRIVE
H6
POWER SUPPLY
H52
SENSOR
H51
CAMERA