INTELLIGENT COCKPIT INTERACTION METHOD AND SYSTEM BASED ON MULTIMODAL SPATIOTEMPORAL ATTENTION MECHANISM

Field of the Invention

The present invention relates to the field of intelligent cockpits, and in particular to an intelligent cockpit interaction method and system based on a multimodal spatiotemporal attention mechanism.

Background to the Invention

Current vehicles' prevention and control of window fogging mainly rely on a passive response mechanism based on fixed temperature and humidity thresholds. The system starts the full-cabin air conditioning for defogging only when the sensor detects that fog has formed, which has inherent defects such as delayed safety intervention, high energy consumption, and poor passenger comfort.

Existing technologies lack the abilities of fusion perception and dynamic prediction of multi-dimensional information such as occupant behavior, in-cabin microclimate, and window surface temperature field. Therefore, the present invention provides an intelligent cockpit interaction method and system based on a multimodal spatiotemporal attention mechanism to solve the above problems.

Statement of Invention

Aiming at the deficiencies of existing technologies, an objective of the present invention is to provide an intelligent cockpit interaction method and system based on a multimodal spatiotemporal attention mechanism to solve the problems existing in the above background.

The present invention is realized as follows: an intelligent cockpit interaction method based on a multimodal spatiotemporal attention mechanism includes the following steps:

collecting multimodal data in real time through a sensor array distributed in a cockpit, where the multimodal data include occupant position, breathing characteristics, window glass temperature field, ambient temperature and humidity, and air conditioning status, and the sensor array includes an infrared thermal imaging sensor, a visual sensor, a temperature and humidity sensor, and a microphone array;

conducting fusion analysis on the multimodal data based on a pre-constructed spatiotemporal attention prediction model to obtain a spatiotemporal probability heatmap capable of representing fog condensation on windshields and side windows within a predetermined future time window;

identifying high-risk fogging areas with a condensation probability exceeding a threshold and located within a driver's field of vision based on the spatiotemporal probability heatmap, and generating a predictive interaction instruction for the area, where the instruction is used for setting a wind direction, wind speed, and duration of corresponding air conditioning outlets; and

executing the predictive interaction instruction to perform airflow intervention on the high-risk fogging areas before condensation occurs.

The collecting multimodal data in real time through a sensor array distributed in a cockpit specifically includes the step of:

periodically scanning to obtain a window glass temperature field on inner surfaces of the windshields and side windows through the infrared thermal imaging sensor;

acquiring images containing occupants through the visual sensor, and determining and outputting precise three-dimensional position coordinates of each occupant in a cockpit coordinate system using image recognition and positioning algorithms as occupant position data;

extracting and quantifying breathing characteristic data representing breathing intensity and frequency from the breathing characteristics collected by the microphone array;

collecting ambient temperature and humidity data of different zones in the cockpit and air conditioning status data of an air conditioning system through the temperature and humidity sensor; and aligning and fusing the temperature distribution matrix, occupant position, breathing characteristics, ambient temperature and humidity, and air conditioning status data with a unified timestamp and a spatial reference system.

The identifying high-risk fogging areas with a condensation probability exceeding a threshold and located within a driver's field of vision based on the spatiotemporal probability heatmap, and generating a predictive interaction instruction for the area specifically includes the steps of:

parsing the spatiotemporal probability heatmap, identifying all continuous areas in a map where the condensation probability exceeds a preset safety threshold, and defining these areas as high-risk fogging areas to be treated;

for each high-risk fogging area, calculating three-dimensional coordinates of a geometric center in a window coordinate system, and predicting a diffusion direction vector of a condensation trend of the area based on changes in the heatmap over a time dimension;

dynamically calculating and allocating the required wind direction adjustment angles, wind speed levels, and action durations for multiple air conditioning outlets based on the pre-stored air conditioning outlet spatial position and wind direction adjustment capability model, as well as the three-dimensional coordinates and diffusion direction vector; and

optimizing and adjusting the calculated wind direction, wind speed, and duration parameters according to a current vehicle operation mode, and generating a predictive interaction instruction containing target air outlet identifiers and a set of control parameters.

The method further includes:

continuously collecting actual images of the high-risk fogging areas through the visual sensor within the predetermined future time window, and quantitatively evaluating a visibility level of the area using image analysis algorithms;

comparing the visibility level with the spatiotemporal probability heatmap to obtain a predicted local error value;

integrating complete multimodal fusion data, the generated predictive interaction instruction, and the predicted local error value into a training sample and storing the sample in a historical interaction case library; and

performing incremental training and parameter adjustment on the spatiotemporal attention prediction model using all samples in the historical interaction case library.

Another objective of the present invention is to provide an intelligent cockpit interaction system based on a multi-modal spatiotemporal attention mechanism, including:

an environment perception module for collecting multimodal data in real time through a sensor array distributed in a cockpit, where the multimodal data include occupant position, breathing characteristics, window glass temperature field, ambient temperature and humidity, and air conditioning status, and the sensor array includes an infrared thermal imaging sensor, a visual sensor, a temperature and humidity sensor, and a microphone array;

a risk prediction module for conducting fusion analysis on the multimodal data based on a pre-constructed spatiotemporal attention prediction model to obtain a spatiotemporal probability heatmap capable of representing fog condensation on windshields and side windows within a predetermined future time window;

an instruction generation module for identifying high-risk fogging areas with a condensation probability exceeding a threshold and located within a driver's field of vision based on the spatiotemporal probability heatmap, and generating a predictive interaction instruction for the area, where the instruction is used for setting a wind direction, wind speed, and duration of corresponding air conditioning outlets; and

an intervention execution module for executing the predictive interaction instruction to perform airflow intervention on the high-risk fogging areas before condensation occurs.

The instruction generation module includes:

an area identification unit for parsing the spatiotemporal probability heatmap, identifying all continuous areas in a map where the condensation probability exceeds a preset safety threshold, and defining these areas as high-risk fogging areas to be treated;

a trend analysis unit for calculating, for each high-risk fogging area, three-dimensional coordinates of a geometric center in a window coordinate system, and predicting a diffusion direction vector of a condensation trend of the area based on changes in the heatmap over a time dimension;

a parameter allocation unit for dynamically calculating and allocating the required wind direction adjustment angles, wind speed levels, and action durations for multiple air conditioning outlets based on the pre-stored air conditioning outlet spatial position and wind direction adjustment capability model, as well as the three-dimensional coordinates and diffusion direction vector; and

an instruction optimization unit for optimizing and adjusting the calculated wind direction, wind speed, and duration parameters according to a current vehicle operation mode, and generating a predictive interaction instruction containing target air outlet identifiers and a set of control parameters.

The system further includes a feedback optimization module, and the feedback optimization module includes:

an effect evaluation unit for continuously collecting actual images of the high-risk fogging areas through the visual sensor within the predetermined future time window, and quantitatively evaluating a visibility level of the area using image analysis algorithms;

an error calculation unit for comparing the visibility level with the spatiotemporal probability heatmap to obtain a predicted local error value;

a sample construction unit for integrating complete multimodal fusion data, the generated predictive interaction instruction, and the predicted local error value into a training sample and storing the sample in a historical interaction case library; and

a model optimization unit for performing incremental training and parameter adjustment on the spatiotemporal attention prediction model using all samples in the historical interaction case library.

Compared with the prior art, the present invention has the following advantageous effects.

By fusing multi-dimensional data such as occupant behavior, window temperature field, and in-cabin microclimate, and conducting joint analysis and prediction using a spatiotemporal attention mechanism, the present invention can predict the specific location and time of fog condensation before the physical process of condensation occurs. This enables targeted and localized airflow intervention to be implemented in advance, fundamentally eliminating the safety hazard of temporary blind spots in the driver's field of vision caused by fogging. At the same time, optimizing the traditional defogging mode into precise intervention for high-risk areas greatly reduces energy consumption and avoids discomfort to occupants caused by strong winds. In summary, the present invention enables the intelligent cockpit to have the abilities of forward-looking and intelligent interaction with the environment and occupants.

Brief Description of the Drawings

FIG. 1 is a flowchart of an intelligent cockpit interaction method.

FIG. 2 is a flowchart of collecting multimodal data in the intelligent cockpit interaction method.

FIG. 3 is a flowchart of generating a predictive interaction instruction for an area in the intelligent cockpit interaction method.

FIG. 4 is a flowchart of performing incremental training and parameter adjustment on a spatiotemporal attention prediction model in the intelligent cockpit interaction method.

FIG. 5 is a flowchart of an intelligent cockpit interaction system.

FIG. 6 is a flowchart of an instruction generation module in the intelligent cockpit interaction system.

FIG. 7 is a schematic structural diagram of a feedback optimization module in the intelligent cockpit interaction system.

Detailed Description

The present invention will be further described in detail below with reference to the accompanying drawings and specific embodiments.

As shown in FIG. 1, an embodiment of the present invention provides an intelligent cockpit interaction method based on a multimodal spatiotemporal attention mechanism, including the following steps:

S100, collecting multimodal data in real time through a sensor array distributed in a cockpit, where the multimodal data include occupant position, breathing characteristics, window glass temperature field, ambient temperature and humidity, and air conditioning status, and the sensor array includes an infrared thermal imaging sensor, a visual sensor, a temperature and humidity sensor, and a microphone array;

S200, conducting fusion analysis on the multimodal data based on a pre-constructed spatiotemporal attention prediction model to obtain a spatiotemporal probability heatmap capable of representing fog condensation on windshields and side windows within a predetermined future time window;

S300, identifying high-risk fogging areas with a condensation probability exceeding a threshold and located within a driver's field of vision based on the spatiotemporal probability heatmap, and generating a predictive interaction instruction for the area, where the instruction is used for setting a wind direction, wind speed, and duration of corresponding air conditioning outlets; and

S400, executing the predictive interaction instruction to perform airflow intervention on the high-risk fogging areas before condensation occurs.

By fusing multi-dimensional data such as occupant behavior, window temperature field, and in-cabin microclimate, and conducting joint analysis and prediction using a spatiotemporal attention mechanism, the present invention can predict the specific location and time of fog condensation before the physical process of condensation occurs. This enables targeted and localized airflow intervention to be implemented in advance, fundamentally eliminating the safety hazard of temporary blind spots in the driver's field of vision caused by fogging. At the same time, optimizing the traditional defogging mode into precise intervention for high-risk areas greatly reduces energy consumption and avoids discomfort to occupants caused by strong winds.

As shown in FIG. 2, the collecting multimodal data in real time through a sensor array distributed in a cockpit specifically includes the step of:

S101, periodically scanning to obtain a window glass temperature field on inner surfaces of the windshields and side windows through the infrared thermal imaging sensor;

S102, acquiring images containing occupants through the visual sensor, and determining and outputting precise three-dimensional position coordinates of each occupant in a cockpit coordinate system using image recognition and positioning algorithms as occupant position data;

S103, extracting and quantifying breathing characteristic data representing breathing intensity and frequency from the breathing characteristics collected by the microphone array;

S104, collecting ambient temperature and humidity data of different zones in the cockpit and air conditioning status data of an air conditioning system through the temperature and humidity sensor; and

S105, aligning and fusing the temperature distribution matrix, occupant position, breathing characteristics, ambient temperature and humidity, and air conditioning status data with a unified timestamp and a spatial reference system.

The window glass temperature field on the inner surfaces of the windshield and side windows is directly obtained by periodic scanning with the infrared thermal imaging sensor deployed in the cockpit. At the same time, the visual sensor continuously captures images containing occupants, and runs built-in image recognition and positioning algorithms to real-time output the precise three-dimensional position coordinates of each occupant in the cockpit coordinate system, which are used as the occupant position data. To obtain breathing characteristic data, the original audio signals collected by the microphone array are processed to extract and quantify digital features representing breathing intensity and frequency. Meanwhile, temperature and humidity sensors distributed at different positions in the cabin directly collect ambient temperature and humidity data of each zone, and air conditioning status data of the air conditioning system is directly read from the bus. Finally, all data (window glass temperature field, occupant position data, breathing characteristic data, ambient temperature and humidity data, and air conditioning status data) are sent to a data fusion core, which strictly aligns and fuses all data with a unified timestamp and spatial reference frame to form a standardized data set that can be directly processed by subsequent models.

As shown in FIG. 3, the identifying high-risk fogging areas with a condensation probability exceeding a threshold and located within a driver's field of vision based on the spatiotemporal probability heatmap, and generating a predictive interaction instruction for the area specifically includes the steps of:

S301, parsing the spatiotemporal probability heatmap, identifying all continuous areas in a map where the condensation probability exceeds a preset safety threshold, and defining these areas as high-risk fogging areas to be treated;

S302, for each high-risk fogging area, calculating three-dimensional coordinates of a geometric center in a window coordinate system, and predicting a diffusion direction vector of a condensation trend of the area based on changes in the heatmap over a time dimension;

S303, dynamically calculating and allocating the required wind direction adjustment angles, wind speed levels, and action durations for multiple air conditioning outlets based on the pre-stored air conditioning outlet spatial position and wind direction adjustment capability model, as well as the three-dimensional coordinates and diffusion direction vector; and

S304, optimizing and adjusting the calculated wind direction, wind speed, and duration parameters according to a current vehicle operation mode, and generating a predictive interaction instruction containing target air outlet identifiers and a set of control parameters.

First, the spatiotemporal probability heatmap is parsed, and all continuous pixel areas in the map where the condensation probability exceeds a preset safety threshold are identified through image segmentation algorithms and defined as pending high-risk fogging areas. Then, for each high-risk fogging area, the three-dimensional coordinates of the center of its geometric shape in the window coordinate system are calculated, and the morphological evolution of the area in the spatiotemporal probability heatmap of continuous time frames is analyzed to calculate the diffusion direction vector representing its condensation trend. The pre-stored air conditioning outlet spatial position and wind direction adjustment capability model (a data model including the physical position, adjustable angle range, and wind speed range of each air outlet) is called, and the combination of the required wind direction adjustment angle, wind speed level, and action duration for one or more air conditioning outlets is dynamically calculated and allocated through geometric projection and optimization algorithms based on the three-dimensional coordinates and diffusion direction vector. Finally, the initially calculated wind direction adjustment angle, wind speed level, and action duration parameters are finally optimized, adjusted, and verified for compliance in combination with the strategy constraints defined in the current vehicle operation mode (such as comfort mode, energy-saving mode, or long-distance mode), thereby generating a predictive interaction instruction that includes specific target air outlet identifiers and a complete set of control parameters and can be directly issued for execution.

As shown in FIG. 4, the intelligent cockpit interaction method based on a multimodal spatiotemporal attention mechanism further includes:

S501, continuously collecting actual images of the high-risk fogging areas through the visual sensor within the predetermined future time window, and quantitatively evaluating a visibility level of the area using image analysis algorithms;

S502, comparing the visibility level with the spatiotemporal probability heatmap to obtain a predicted local error value;

S503, integrating complete multimodal fusion data, the generated predictive interaction instruction, and the predicted local error value into a training sample and storing the sample in a historical interaction case library; and

S504, performing incremental training and parameter adjustment on the spatiotemporal attention prediction model using all samples in the historical interaction case library.

In the embodiment of the present invention, within the predetermined future time window after executing the predictive interaction instruction, the system will continuously call the visual sensor to collect images of the high-risk fogging area where intervention has been implemented; and by real-time processing these actual images, the visibility level of the area is quantitatively evaluated using image analysis algorithms (such as a calculation model based on contrast or edge sharpness) to quantify the intervention effect. Then, the visibility level is inversely mapped and compared with the predicted condensation probability of the corresponding spatiotemporal grid in the spatiotemporal probability heatmap on which the instruction is based, and a predicted local error value is calculated through a preset error function, which directly represents the accuracy of the model's current prediction. All original data and decision records triggering the complete interaction cycle are encapsulated in a standardized format to form a structured training sample, and the historical interaction case library is updated with this sample. Finally, the system schedules a machine learning framework in the background to periodically or when the case library reaches a certain scale, perform incremental training and parameter adjustment on the spatiotemporal attention prediction model using all samples in the historical interaction case library, and adjust the internal weights of the model through algorithms such as backpropagation to make its prediction of similar spatiotemporal scenarios more accurate in the future.

As shown in FIG. 5, an embodiment of the present invention also provides an intelligent cockpit interaction system based on a multimodal spatiotemporal attention mechanism, including:

an environment perception module 100, for collecting multimodal data in real time through a sensor array distributed in a cockpit, where the multimodal data include occupant position, breathing characteristics, window glass temperature field, ambient temperature and humidity, and air conditioning status, and the sensor array includes an infrared thermal imaging sensor, a visual sensor, a temperature and humidity sensor, and a microphone array;

a risk prediction module 200, for conducting fusion analysis on the multimodal data based on a pre-constructed spatiotemporal attention prediction model to obtain a spatiotemporal probability heatmap capable of representing fog condensation on windshields and side windows within a predetermined future time window;

an instruction generation module 300, for identifying high-risk fogging areas with a condensation probability exceeding a threshold and located within a driver's field of vision based on the spatiotemporal probability heatmap, and generating a predictive interaction instruction for the area, where the instruction is used for setting a wind direction, wind speed, and duration of corresponding air conditioning outlets; and

an intervention execution module 400, for executing the predictive interaction instruction to perform airflow intervention on the high-risk fogging areas before condensation occurs.

As shown in FIG. 6, the instruction generation module 300 includes:

an area identification unit 301, for parsing the spatiotemporal probability heatmap, identifying all continuous areas in a map where the condensation probability exceeds a preset safety threshold, and defining these areas as high-risk fogging areas to be treated;

a trend analysis unit 302, for calculating, for each high-risk fogging area, three-dimensional coordinates of a geometric center in a window coordinate system, and predicting a diffusion direction vector of a condensation trend of the area based on changes in the heatmap over a time dimension;

a parameter allocation unit 303, for dynamically calculating and allocating the required wind direction adjustment angles, wind speed levels, and action durations for multiple air conditioning outlets based on the pre-stored air conditioning outlet spatial position and wind direction adjustment capability model, as well as the three-dimensional coordinates and diffusion direction vector; and

an instruction optimization unit 304, for optimizing and adjusting the calculated wind direction, wind speed, and duration parameters according to a current vehicle operation mode, and generating a predictive interaction instruction containing target air outlet identifiers and a set of control parameters.

As shown in FIG. 7, the intelligent cockpit interaction system based on a multimodal spatiotemporal attention mechanism further includes a feedback optimization module 500, and the feedback optimization module 500 includes:

an effect evaluation unit 501, for continuously collecting actual images of the high-risk fogging areas through the visual sensor within the predetermined future time window, and quantitatively evaluating a visibility level of the area using image analysis algorithms;

an error calculation unit 502, for comparing the visibility level with the spatiotemporal probability heatmap to obtain a predicted local error value;

a sample construction unit 503, for integrating complete multimodal fusion data, the generated predictive interaction instruction, and the predicted local error value into a training sample and storing the sample in a historical interaction case library; and

a model optimization unit 504, for performing incremental training and parameter adjustment on the spatiotemporal attention prediction model using all samples in the historical interaction case library.

The above is only a detailed description of the preferred embodiments of the present invention and is not intended to limit the present invention.