ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

III Международная конференция

«Когнитивная Робототехника»

в рамках Международного форума

«Интеллектуальные системы
4-й промышленной революции»

КОГНИТИВНАЯ
РОБОТОТЕХНИКА

СБОРНИК ТЕЗИСОВ КОНФЕРЕНЦИИ

Спонсор

Организаторы

21-23 ноября 2018 года

Томск
Издательский Дом Томского государственного университета
2018

be changed on the go like Keras. In the end, for a person who is just starting to learn ANNs, it comes down to personal preference to make the choice. However, if the framework is being chosen with a specific project in mind, then it is worth paying attention to what framework would best fit the task at hand.

Overall, two frameworks do tend to stand out in their performance, and that is TensorFlow and Caffe. Most outlets tend to agree that these frameworks have the most developer and community support as well as high performance when applied properly. They tend to be relatively user friendly, with plenty of learning material for beginners and enormous library variety for more advanced use. Their functionality is still being expanded by the developers and the frameworks have a bright future.

This study was supported by the Russian Foundation for Basic Research (grant No. 16-29-04388).

References

[1] Maruti Techlabs. Best deep learning frameworks for data scince enthusiasts // Meduim. 2018 April 5th Retrieved on: November 10th 2018 Available at:https://medium.com/the-mission/8-best-deep-learning-frameworks-for-data-science-enthusiasts-d72714157761

[2] González-Gutiérrez C, Santos JD, Martínez-Zarzuela M, et al. Comparative Study of Neural Network Frameworks for the Next Generation of Adaptive Optics Systems. Sensors (Basel). 2017;17(6):1263. Published 2017 Jun 2. Retrieved on November 9th 2018

GESTURE BASED TERRAIN MAPPING SYSTEM-OVERVIEW

M. Mondal, V.I. Syryamkin

National Research Tomsk State University
mainakme2140@gmail.com

Up until a few years ago, the only way to get an aerial overview of a designated area was to fly over it with a manned aircraft and physically inspecting the areas or taking pictures of it. Unmanned Aerial Vehicles, have been a blessing to all the industries which relied on aerial images or drawings for their work. Unmanned aerial vehicles also knows as drones, in more colloquial terms, is a blessing to all industries in the world. The cost of manufac-

turing and assembling drones have gone down significantly as majority of people adapted to using it.

The film industries are using it to get excellent footage with accurate direction and lighting. Civilians can use it as toys. Students can use it to learn avionics and flight mechanisms. Youtubers use it for professional high resolution footage. Rescue teams use it for surveying the area. Armies use it to get a tactical advantage in a firefight and recently, Amazon is using drones for superfast delivery service. The possibilities are endless if properly invested on. Machine Learning and Artificial Intelligence are next best thing in the field of Computer Science and that's why I plan to merge these two fields and present something which fundamentally adds more features and a hive mind facility to even work without a commanding signal.

The current real time terrain mapping, usually carried out by Self Drive Cars usually use the overlay of the map to find directions. The on board computer of the car, usually uses a bunch of sensor like: Proximity Sensor, Stereoscopic Cameras to identify objects and cars in its trajectory. This method very viable for cars because it has less to no space constraints to house the compute power. I propose a method to identify terrain defects for tactical combat and rescue operations by dividing the work into terrain recognition and terrain mapping. The Operator is the person who controls the drones and provides the drones with directional data for it to follow. There is also an autopilot which takes over the controls when no gesture commands are received for a while.

All the mapping systems implementations, if implemented in real-time cost huge computing resources and cannot map accurately. The most accurate is the LIDAR based Mapping as it is used in self drive cars and other heavy vehicles with ample computing power.

There are several attempts for real-time mapping using drones but most of the projects are pretty much at its infancy as they all try to map the terrain in real time. They all require huge computing power since terrain mapping requires huge resources.

My plan is to divide the work into 4 parts: -

1. **Gesture Recognition** – handled by an independent flight controller.
2. **Drone Pilot –** The schedule to be run when no gesture commands are received.
3. **Terrain recognition –** The schedule to take high resolution images of the area and recognize the pits and outline of the area for tactical use.
4. **Terrain Mapping –** The Schedule which runs on the a separate high – end computer which uses its compute power and the high resolution images from the previous schedule and maps the area with accuracy.

For the actual system, there should be 2 working computer units.

1. **Drone Controller System** – A computer specifically to pilot the drone. This computer also listens to the gestures control commands given by the operator and decides on the which schedule to follow, Autopilot or Gesture.

2. **Mapping System** – The Computer which takes in all the sensor data and segregates the data into the what's required immediately and what will be sent to the EC service for further processing. This unit also communicates to the operator and will relay the terrain recognition data to consumption device.

The distinguishing part of this system will be the fact that this device does not have to process high resolution images in real-time. The computer handling the terrain detection will only produce a depth map using numeric values after calculating the depth and a thermal imaging data can be used for overlay. This data can be sent over to the human operator for physically exploring the terrain. An object detection algorithm like tensor flow maybe implemented in the middle which takes in camera images and recognizes objects and overlays it in the map.

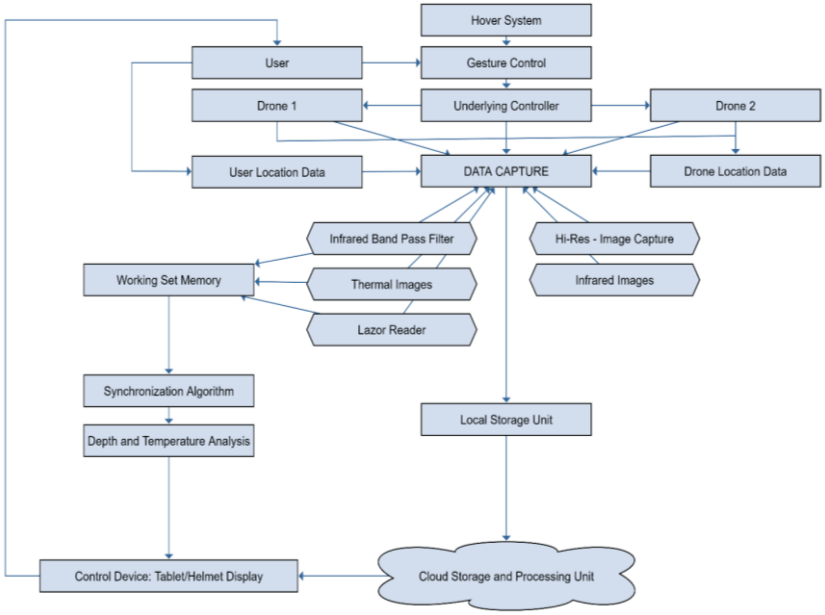

Fig. 1. Structure Diagram

The high resolution images captured can be uploaded to a cloud server for processing. The cloud server can use a commercial or an open source mapping software which used geolocation data and combines it with all the sensor data received and spits out an accurate map of the given area.

As we see here, there are 2 major zones of work:

1. Which goes to the working set memory (after storing the data to the non-volatile memory) which is required for Real-time data processing to assist the operator of the system navigate through the terrain.
2. The second zone revolves around the data capture and processing the captured data using an elastic cloud service (Scalable Computing). The Scalable Compute Service will process all the data and produce a 3d map of the area which can be used for future endeavors.

Categorization of data in the given concept:

1. Thermal Images
2. Stereoscopic camera images (high-resolution image capture)
3. Depth sensor data
4. Infrared Camera Data (for night vision)
5. Generic Camera footage (for relay to the user and use for object detection by Tensor flow)
6. Geolocation data of the Drone as well as the operator.

Conclusion and Future Work

The sole idea of this project was to differentiate between mapping an area and terrain detection to utilize computing resources in an efficient way. The research currently being done in this sphere is phenomenal and when we combine all the research we get different views on how to aliment an idea and the sphere of Unmanned Ariel Vehicle is just at its infancy. We have many different challenges ahead like increasing flight time, increasing signal range and many other things but with the insight I provided, it is certain that all the problems can be solved given time. This article shows the current work being done in the sphere of terrain mapping and drone automation, and combines the two for the inception of a new concept. Further, these tasks can be sub-divided into more tasks for a simpler and faster approach to tackle the problem of terrain mapping and recognition.

This study was supported by the Russian Foundation for Basic Research (grant No. 16-29-04388).

References

[1] Scheding S., Leal J., Bishop M., Sukkarieh S. Terrain Mapping in Real-Time: Sensors and Algorithms / Australian Centre for Field Robotics (ACFR), The University of Sydney. URL: http://www.regional.org.au/au/gia/10/298scheding.htm
[2] ROS-based UAV control using hand gesture recognition – 2017 : 29th Chinese Control and Decision Conference (CCDC), IEEE. URL: https://ieeexplore.ieee.org/document/7978402
[3] Simultaneous localization and mapping // IEEE Robotics & Automation Magazine. URL: https://ieeexplore.ieee.org/document/1638022
[4] Panin V.E., Syryamkin V.I., Deryugin Ye. Ye. et al. Optical-television techniques for research and diagnostics of materials at mesolevel // Physical Mesomechanics and Computer-Aided Design of Materials / ed. by V.E. Panin. Novosibirsk : Nauka, 1995. Vol. 1. P. 176–194.
[5] Simultaneous localization and mapping (SLAM). URL: https://ieeexplore.ieee.org/document/1678144
[6] Rumyantseva T.B., Solinthone Ph., Syryamkin V.I., Vaganova E.V., Zinov V.G. Technology management. Part 2: Strategic management of intellectual property / ed. by V.I. Syryamkin. Tomsk, 2016.
[7] Gesture Control of Drone Using a Motion Controller : 2016 International Conference on Industrial Informatics and Computer Systems (CIICS). URL: https://www.researchgate.net/publication/301800528_Gesture_control_of_drone_using_a_motion_controller

POSSIBILITIES FOR APPLIED JOINT SPEECH PROCESSING AND COMPUTER VISION SYSTEMS

E. Solis Romeu, V.I. Syryamkin

National Research Tomsk State University
edgar7sagan@gmail.com

This report addresses the current state of the Computer Vision and the Speech Processing fields. It explains some of the most used algorithms in these fields and provides an overview of the current use of these technologies in the market. The objective of this report is to find new possible industrial and commercial applications that use both Computer Vision and Speech Processing.

Computer Science has many fields and one of them is Speech Processing. The aim of this discipline is to create computer systems capable of recognizing spoken words. The main idea is to detect the words spoken by a person and process that data to identify what the person is saying. To do this, a microphone or similar device is used to input the audio data into the system. One of the main objectives in this field is the implementation of accurate Automatic Speech Processing, this is a system that is able to convert spoken words into digital data to identify individual sounds and use mathematical models to select particular words or phrases.