Robotic Infrared Vision: Enabling operation in fog and low-light environments

Leonard Günzel
German Aerospace Center (DLR)
Institute for the Protection
of Maritime Infrastructures
Germany, Bremerhaven Fischkai 1
Email: leonard.guenzel@dlr.de

Dr. Enno Peters *
German Aerospace Center (DLR)
Institute for the Protection
of Maritime Infrastructures
Germany, Bremerhaven Fischkai 1
Email: enno.peters@dlr.de

Jendrik Schmidt
German Aerospace Center (DLR)
Institute for the Protection
of Maritime Infrastructures
Germany, Bremerhaven Fischkai 1
Email: jendrik.schmidt@dlr.de

***Abstract*—The operation of robots in fog or smoke is a challenging task for both autonomous and remote-controlled operations. To increase the situational awareness and observational abilities, a legged robot was equipped with a thermal camera, which was integrated into the controlling unit to enable operation in environments which would not have been accessible without. The interface between the robot and thermal camera control will be addressed in this context, as well as the physical and mechanical setup. Additionally to the communication between Spot and the thermal camera, an option was developed to access the data stream from an external computer within the network. The software developed for this purpose will be introduced as well.**

The performance of the system will be discussed on the basis of a field test in fog. Demonstrating that the implemented system enhanced the situational awareness of the operator and allowed for operation in a scenario, which would have been impossible for remote robot operation without the additional capability of thermal imaging. In particular, the system enabled the operator to detect a person in an environment with high fog density. The robot was further able to approach the person in an obstacle filled environment and therefore gather more detailed information about the scene.

I. Introduction

Hazardous environments impose severe challenges ranging from high temperatures to toxic or radiological dangers. Operating in these environments poses a threat to human health, which is a limiting factor for various activities like inspection tasks or emergency response (e.g.in the case of fire). To enable such activities without the need of human presence, the notion towards robotic systems, designed to operate in these special environments, has seen a steady increase in popularity.

With more sophisticated emergency robots, humans no longer need to be present physically in dangerous, extreme or health-affecting scenarios. However, autonomy is limited in complex situations and requires the human to overtake direct control. Therefore, the operator depends on a comprehensive awareness of the situation provided by sensors. Digital imaging sensors usually produce images similar to what the human eye would perceive and are therefore the key sensor in most robotic operations. However, similar to the human counterpart, their functionality suffers in scenarios of low light or reduced visibility due to fog. With additional infrared sensors, this loss can not only be compensated, but also exceed the naturally given limits of human perception. More precisely, due to a thermal camera, the operator is able to perceive thermal radiation (heat) emitted by objects and is not dependent on an external light source. It therefore enables operation in low light conditions and is barely affected by scattering from fog or smoke, providing an increased perception in low visibility conditions [1]–[4]. The resulting advantages can be best seen in Figure 1. With an increase in fog density, the boat depicted in the images taken by the consumer camera is almost unrecognizable due to the reflection and scattering by the fog. The thermal image is unchanged and the scene is clearly observable.

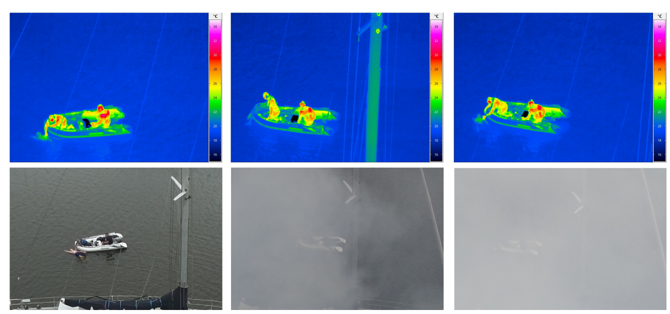

Fig. 1. Comparison of two sets of images. The upper set displays images taken from a thermal camera and the lower set is taken by a digital consumer camera. The digital camera's view is affected by increasing fog while the thermal camera is almost unaffected.

Wheeled robots, often used for extreme risk scenarios (ranging from bomb disposal to maintenance of radioactive reactors), are able to move on even terrain in an efficient and fast manner, while carrying large payloads at the same time. However, in uneven terrain wheeled robots often fail when confronted with obstacles such as rocks. Legged robots on the other side, offer a flexible transversion of complex terrain, with a degree of mobility comparable to living creatures. The ability to master these terrains gives legged robots a bigger field of operation with fewer restrictions to consider.

In this study, the legged robot Spot (produced by Boston Dynamics [5]) was equipped with an additional high-resolution thermal camera. The objective was to stream data from this camera directly to Spot's software interface and from there to the remote control of the robot. Furthermore, it was called for the possibility to stream the data inside the network to another computer (which is part of a situational awareness room). This allows for a centralized surveillance of the mission, especially for operations involving several systems (drones, surveillance cameras, etc.). The added value in terms of situational awareness under strong smoke conditions obtained from the combination of these two systems (thermal camera and Spot) was investigated through a field test.

II. System

The System consists of three main parts, the robotic platform (Spot), a computer (Next Unit Computing (NUC) [6]) and the sensor (thermal camera). The thermal camera is equipped with a 640 by 480 pixel (10.88 mm by 8.16 mm) micro-bolometer image sensor and is able to detect infrared radiation ranging from $7.5\,\mu m$ to $14\,\mu m$ [7]. It is equipped with a telephoto-lens with a focal length of $60\,mm$, resulting in an AFOV (Angular Field Of View) of $13°$ diagonal. The camera is placed in a protective housing and mounted on the back of Spot. The weight of the payload is distributed close to Spot's center of gravity, as additional undistributed weight impacts the robot's ability to maneuver. The robot not only functions as a mobile platform, but does also come equipped with multiple sensors. The data produced by these sensors can be evaluated by the operator through the remote control. To implement the thermal camera's video stream into the remote control it is necessary to establish a physical and software interface between both. The data sent by the camera needs to be compressed and formatted to be displayed on the remote control. This kind of computing is not supported on Spot's own system, further it its prohibited to use external software on Spot's internal computer. Therefore, the NUC is needed to connect to the camera as well as to Spot (as shown in Figure 2). Furthermore, it is used to transmit the live video to a situational awareness room for a centralized mission surveillance.

A. Power supply

The payloads that need power supply are the thermal camera and the NUC. On average, the thermal camera in combination with the heat element consume $11\,W$ with a supply voltage of $24\,V$. The NUC uses approximately $28\,W$ of power while also requiring a supply voltage of 12 to $24\,V$. Therefore, the overall power consumption of the setup is approximately $40\,W$. The energy consumption depends slightly on the outside temperature, because of the heat element stabilizing the housing temperature.

With an estimated operation time of 1.5 hours, which is the typical operation time of Spot in normal usage, the needed capacity adds up to approximately $60\,Wh$. In order to provide both devices with the required energy, two options must be weighed up: either using an external battery pack or connecting the devices directly to Spots power supply.

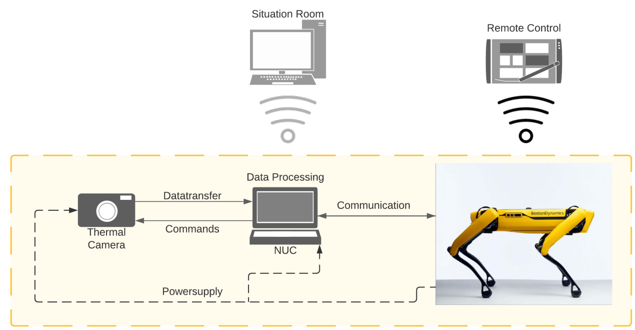

Fig. 2. Schematic overview of the setup. The yellow box includes all devices which are connected to each other physically. The wired network connections are illustrated with solid lines and the power supply which is provided by Spot with dashed lines. The remote control connects to Spot via WLAN. The situational awareness room utilizes a WLAN connection to communicate with the NUC.

Spot is equipped with a $605\,Wh$ battery and has two ports on its back which can each provide up to $150\,W$. Equipped with the GXP Adapter (General Expansion Payload [8]) (which provides several breakout ports from Spot's main payload port) the power outlet is split up into $5\,V$, $12\,V$ and $24\,V$. With a maximum power output of up to $150\,W$ per payload port, the camera and the NUC can be powered by a single payload port.

Keeping in mind that the additional weight decreases the efficiency and operating time of Spot, a battery pack would not only add another complex part requiring maintenance and charging, but also would shorten mission operation time and take up additional space on Spot's back. On the other hand, if the devices are connected to Spot's power outlets, they drain the battery faster and consequently shorten the operation time by approximately $10\,\%$. Crucial for this decision was the fact that a battery pack would further complicate the system design and would take up needed space on Spot's back. Therefore, it was decided to dismiss the idea of a battery pack and to use the power outlets provided by Spot.

B. Mechanical mounting

Three devices need to be mounted to Spot physically, the thermal camera, the NUC and the GXP-adapter. Furthermore, it is of importance to supply the first two devices with power and to keep the weight as close to the center of gravity as possible. Spot is equipped with a mounting rail to which devices can be screwed using the M5 threads. The thermal camera is the heaviest of all devices ($7\,kg$ including a standard housing and custom mounting bracket) and consequently needs to be placed as close to the middle of Spot's back as possible. In this proof of concept, no special effort was taken in designing a compact, lightweight housing. Potentially, the thermal camera setup can be significantly reduced in mass and size. It was decided to have the camera sit on a fixed mounting pointing forward (in the direction of travel). The different possibilities to adjust and position Spot allow to orientate the camera. The GXP-adapter is placed in the front, as it is the most suitable position regarding the payload port connection. The NUC is placed between the thermal camera and the GXP adapter. To save space, it was decided to place the NUC's interfaces above the GXP interfaces allowing for the cables to be fed through below the NUC. The final physical setup can be seen in Figure 3.

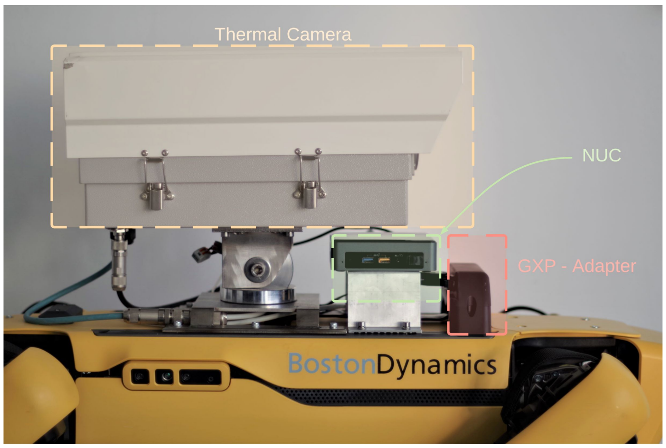

Fig. 3. Image of the three devices (NUC, thermal camera and GXP adapter) mounted to Spot. The NUC and thermal camera each needed a custom fit mounting to attach to Spot.

C. Network interfaces

The general network structure is shown in Figure 2. The thermal camera is an IP-camera and can therefore only be accessed with an Ethernet connection. The NUC is the binding element between the thermal camera and Spot. It is WLAN capable and equipped with a single Ethernet port. The Ethernet port is used to communicate, via a static Ethernet connection, with the thermal camera. An additional Ethernet port was added, using an adapter, to establish a connection to Spot, that is more stable than WLAN. The WLAN capability is used to communicate between the NUC and the situational awareness room. Spot is accessible over Ethernet or WLAN. The Ethernet connection is used to communicate with the NUC and the WLAN connection is utilized to communicate with the remote control.

D. Software development

The software developed for this project comprises of two Python scripts, running on the NUC which are depicted schematically in Figure 4. The Control Script is used to establish the connection between the thermal camera and the computer, set up a Flask server and process the requests made from the Flask server. The Spot Script is provided by Boston Dynamics and forwards the data it receives to the remote control. Further, in Figure 4 the Flask Server is emphasized as a separate block. It allows for the interaction between the clients (Spot and situational awareness room) and the thermal camera over a network. Flask is a micro web framework written in Python. It enables the user to easily setup a development server in a network and link it to Python functions.

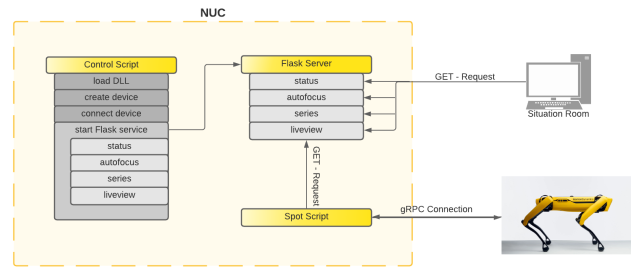

Fig. 4. Schematic overview of the software architecture: On the left-hand side the Control Script is displayed with its functions. After going through the script continuously it sets up a Flask Server, which is illustrated as a separate function in the upper middle part of the schematic. This server provides several functions which can be accessed by Spot and the situational awareness room. The functions which are placed inside the yellow box are all executed on the NUC.

In the Control Script, first, Dynamic linked libraries (DLLs) are loaded. The libraries include camera control functions, e.g. adjusting the focus, and are provided by the manufacturer. After connecting to the camera and initializing the data stream, the script waits for commands from the client through the Flask Server (illustrated as a loop in the schematic Figure 4). It is also depicted as individual script in the middle of the schematic called Flask Server. This is due to the fact, that the Clients (Spot and situational awareness room) interact with the server to execute functions, rather than with the code in Control Script directly. If the Spot Script accesses the address of the liveview function (see Figure 4), the underlying function in the Control Script is executed.

The liveview function compresses the data stream from the thermal camera and streams the compressed images to the Flask Server. The liveview function receives the data from the thermal camera in a 16-bit raw format at a $30\,\mathrm{Hz}$ framerate resulting in a data rate of $\approx 150\,\mathrm{Mbit/s}$. To reduce the data stream between the NUC and the remote control, the data is compressed to JPEG (Joint Photographic Experts Group standard) and saved on disk in a given time interval. For the compressed images a framerate of $5\,\mathrm{Hz}$ was determined and for the high-resolution images the requirement was set to save them locally once every second. The bottleneck for the data transmission is the Spot Script and the framerate has to be adjusted to the restrictions of this script. The data that arrives in the meantime is discarded to reduce storage space and data stream.

The Spot Script (displayed in the lower bottom of Figure 2) forwards the image stream from the liveview function to Spot using a gRPC-Protocol (google Remote Procedure Call). While Spot is only accessing the video stream, the situational awareness room can also interact with the camera through the

Flask Server. In contrast to the Spot Script, the situational awareness room can access all functions provided by the Flask Server. The protocols used for this request are GET and POST methods based on the Hypertext Transfer Protocol (HTTP)

III. Field Test

A test was carried out to provide further information on the extent to which a thermal camera can support the robot in scenarios of poor visibility. In addition, it was tested whether the system is functional as a unit.

A. Field test setup

The robot was confronted with a limited visibility scenario. The objective was to find a person (right side of the tent in Fig. 5 A) in a smoke-filled area in which the robot has to be navigated. Additionally, a dummy was placed in the left corner of the tent and a contrast panel on the back. Behind the person, a smoke machine was placed, filled with a water-based fog fluid.

Spot was started and the connection between the thermal camera and Spot was established. In this process the focus of the camera was set to a given distance. A set focus was used since the integrated autofocus of the thermal camera was not rated as satisfactory. The test was performed twice with different foci, once with a $2\,\mathrm{m}$ focus and once with a $5\,\mathrm{m}$ focus. During each test, the tent was filled with fog until the inside of the tent was not observable anymore with the unaided eye (see Figure 5 B). At the point when no RGB-camera or human eye could observe the scene in the tent anymore, the mission started. The scene was observed through the robot from approximately $20\,\mathrm{m}$ and the robot was steered towards the tent.

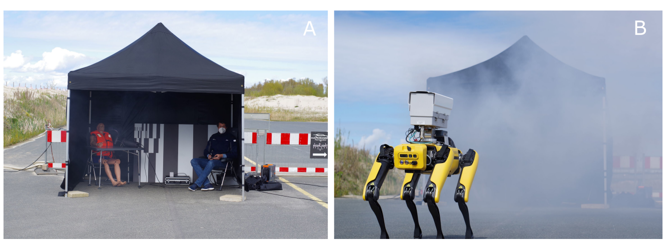

Fig. 5. A) Test setup with a living person sitting inside on the right side and a dummy on the left side. The smoke machine is located outside the tent on the right, but the nozzle is positioned inside the tent.
B) The same test setup can be seen with the fog machine activated. For the human eye or RGB-camera the scene is not observable anymore. In the foreground on the left, Spot is approaching the scene.

B. Results from the field test

The thermal images appear to be unaffected from fog (Figure 6). The scene is clearly observable and allows for a perception of the situation where the consumer camera, set up outside the tent for comparison purposes, fails (Figure 5 B). The images shown in Figure 6 represent the heat radiation in a gray scale image (hot regions white, cold regions black). The images were acquired in different distances to the tent and with two different foci. Especially image D in Figure 6 shows the capabilities of the thermal camera. The operator can not only distinguish between the person (Figure 6 A, B and C) and the dummy, but also observe obstacles and their features (table, tent sides etc.).

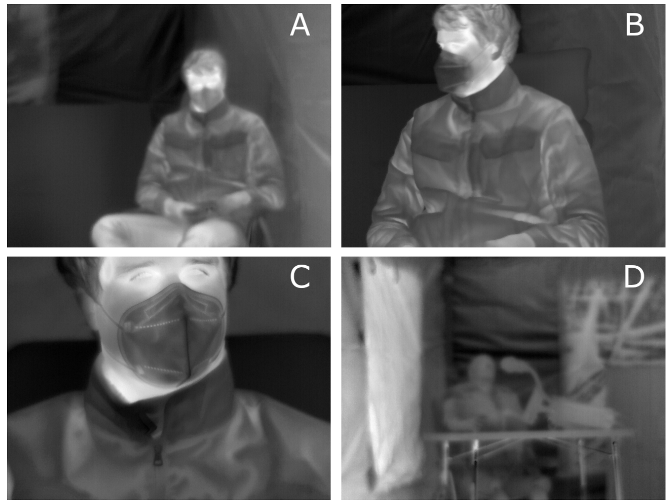

Fig. 6. Thermal images from the field test.

The system successfully completed the field test. The overall performance was very good and the system enhanced the observational capacities of the operator. All components (Spot, thermal camera, NUC) functioned as a unit and allowed the operator to steer Spot inside an area of decreased visibility. Furthermore, the images show that a thermal camera can provide added value in these scenarios. A more detailed discussion of the advantages and limitations can be found in the following section.

IV. Discussion

A. Advantages

The thermal camera enhanced the visual perception of the operator and the implementation of the video stream on the tablet allowed for more direct feedback when controlling the robot (see Figure 7). The connection between the camera and the remote control was stable and gave the operator a better understanding of the situation. During the test, no latency issues of the video stream occurred.

Fig. 7. A) Thermal camera image displayed on the remote control.
B) Stereo camera image provided by Spot displayed on the remote control. The operator can switch between the different perspectives using a drop down menu.

The concept for the data acquisition worked well. The compressed images were saved at a $5\,\mathrm{HZ}$ interval and required approximately 6 KB hard-disk storage each. The high-resolution images were simultaneously saved at a 1 Hz interval and took up approximately 90 KB each. The battery lasted through the whole mission time of approximately one hour.

B. Limitations

Even though the tent as a test environment worked well, the space for the robot to move in it was limited. However, the space was definitely sufficient to cover the robot completely in fog and hence corrupt its stereo sensor data, which are used for autonomous navigation (see Figure 8). Due to the limited space and the risk of the robot falling, it was not possible to perform tests to pass objects in the tent.

Fig. 8. The view of Spot from inside the tent at approximately 2 m distance. With increasing fog, the visibility decreases up to a point where spot is not observable any more.

For further testing, it would be interesting to fill the tent with smoke instead of fog. Due to the different composition of fog and smoke, different behavior of both the visible cameras and the thermal camera can be expected.

The field test concentrated on the functionality of the connection between the thermal camera and the remote control. The connection between the robot and the situational awareness room works if Spot is in the same network as the situational awareness room. Due to the limited reach and high security standards of the network, this was tested succesfully in advance at the institute.

The telephoto-lens used in this case is best suited in a scenario requiring an observation from afar. For the purpose discussed here a different lens with a wider AFOV would have been favorable, providing a better overview to ease steering Spot in the limited space of the tent. Further the fixed focus impaired the perception of the operator. A beneficial solution for this would be an option to set the focus manually from the remote control according to the needs of the situation.

V. Conclusion

The perception of the thermal camera was hardly influenced by fog and enhanced therefore the operator's awareness of the situation. For system operation in closed rooms, a different lens will be needed. Further tests need to be carried out to determine the thermal camera's performance in smoke.

The thermal camera images are integrated in Spot's hand-held remote control and therefore extend the observational capacity of the operator beyond the visible light perception. Additionally, the thermal camera can also be controlled via the situational awareness room, even though a solution needs to be found to access the thermal camera from outside the network. The devices are all powered using the battery pack provided by Spot and offer an integrated and simple handling solution.

In the future, emphases should be given on object recognition. An example can be seen in Figure 9 where some of the images from the field test where analyzed using an object recognition algorithm trained on RGB data [9].

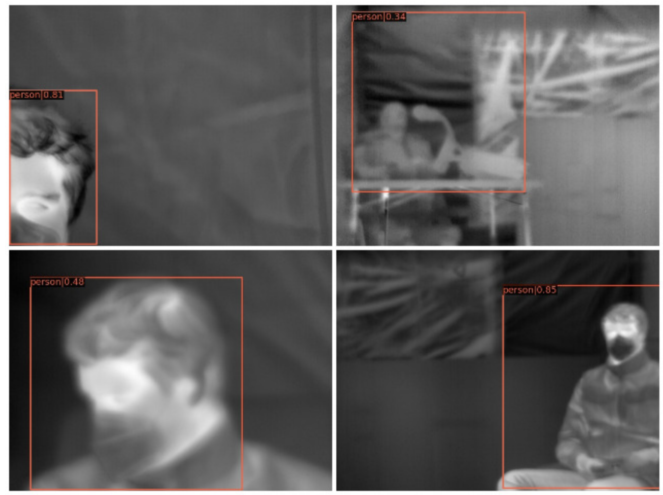

Fig. 9. Thermal images analyzed by object recognition. From upper left to bottom right: human head (recognized as person 81 %), dummy with table (person 34 %), human head and shoulders (person 48 %), full body shot of human (person 85 %).

The information gained from thermal images is not only helpful in environments of decreased visibility, it does also add new information to scenes which have been captured only with an RGB consumer camera before. Furthermore, with improved object recognition, the autonomous capabilities of the robot could expand.

Acknowledgment

The work presented here was funded by DLR (German Aerospace Center) and has been performed as part of a bachelor thesis. We acknowledge Prof. Dr. Axel Bochert (Hochschule Bremerhaven) for his support and supervising the bachelor thesis. We further acknowledge support during the field test, especially Susanne Wollgarten and Matthias Mischung. Borja Jesus Carillo Perez (DLR) provided images with object recognition applied. Marco Berger (DLR) helped with construction drawings and mounting of the thermal camera on Spot. Finally, we acknowledge Boston Dynamics for helpful discussions and support with integration of the video stream to Spot's remote control.

References

[1] Jong-Hwan Kim and Brian Y. Lattimer, *Real-time probabilistic classification of fire and smoke using thermal imagery for intelligent firefighting robot* Fire Safety Journal, 2015, volume 72, DOI:10.1016/j.firesaf.2015.02.007.

[2] Alfred Weigert, Heinrich J Wendker, and Lutz Wisotzki, *Astronomie und Astrophysik* WileyVCH Verlag, 2005.

[3] William C. Malm, *Visibility* Elsevier Science, 2016, page 37-41.

[4] Houghton, H. G., *The Size and Size Distribution of Fog Particles* Physics, 1932, volume 2, pages 467-745 DOI:10.1063/1.1745072.

[5] Boston Dynamics, *Spot Specifications*, [online accessed] 16.07.2021, URL: https://shop.bostondynamics.com/spot.

[6] Intel, *Intel Specifications NUC*, 2021.

[7] Infratec, *InfraTec Specifications*, [online accessed] 16.07.2021, URL: https://www.infratec.eu/thermography/infrared-camera/variocam-hdx-head-600/.

[8] Boston Dynamics, *GXP Specifications*, [online accessed] 16.07.2021, URL: https://shop.bostondynamics.com/spot-gxp.

[9] Ren, Shaoqing and He, Kaiming and Girshick, Ross and Sun, Jian, *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, volume 39, DOI:10.1109/tpami.2016.2577031.