Skoltech
Skolkovo Institute of Science and Technology

Thesis Changes Log

Name of Candidate: Ivan Kalinov

PhD Program: Engineering Systems

Title of Thesis: Development of a heterogeneous robotic system for automated inventory stocktaking of industrial warehouse

Supervisor: Associate Professor Dzmitry Tsetserukou

The thesis document includes the following changes in answer to the external review process.

Dear Jury Members,
I am grateful for your helpful suggestions and comments. Here I provide answers and thesis corrections.

Reviewer: Prof. Andrey Somov

Question #1. Abstract. It is hard to follow Abstract. The thesis includes too many contributions that it is hard to understand how are they connected and address the problem as well as what is the exact scientific value of this work.
Answer. I revised and rewrite abstract according to recommendation.

Question #2 p.22, Method 3 - 'manned UAV' = manned Unmanned Aerial Vehicles. 'Manned Unmanned' sounds like an improper phrase.
Answer. The misprints and inappropriate terms were corrected throughout the thesis.

Question #3 Section 1.2.5 - 'The main limitations of the stocktaking are the presence of errors...'. Which errors are you referring to? Please specify.
Answer. The main limitations of the inventory are the presence of a large number of errors as a result of manual inventory using methods "Method 1 - the complete removal of pallets by workers" or "Method 2 - high-altitude pallet scanning". During the inventory using "Method 1", an error most often occurs as a result of pallet position changes during the inspection. For example, a pallet is removed for inspection from 5-th tier of the rack. And after inspection, it is placed on a lower vacant tier. Errors of this kind also occur in the process of collecting orders, when, after removing a part of the goods, the pallet can be returned to a symmetrical place on the other side of the row. In "Method 2" of warehouse stocktaking, workers unintentionally scan barcodes of adjacent pallet locations. All these errors directly affect the company profit, the duration of the inventory, forcing the warehouse to stop operating activities for several days.

Question #4 Section 1.3 - Why do you need to calculate the UAV global coordinates via the URG coordinates? Which 'well-known' methods do you use for the URG global coordinates calculation. Same paragraph: 'In addition, my method does not imply... ' - you have not formulated your method yet.
Answer. I absolutely agree that I should have provided links to these methods, as well as a link to a part of the thesis, where it will be explained in detail. “For localization and navigation of URG, we use a combination of well-known methods (Gálvez-López et al. [2016], Grisetti et al. [2010]) with optimizations for my system, which is explained in more detail in Section 5.1.”

Question #5 In next paragraph you switch to the problem of bar code scanning. What is the actual problem of UAV localization - is it the localization in space? As far as I understand, you get XY coordinates from the UGR and Z coordinate from the bar codes. What is about wireless technologies for localization, e.g. UWB secures accurate localization. Is a bar code scanning the only option?
Answer. You are right, to get the global coordinates in the coordinate system associated with the warehouse, we use the coordinates of the mobile platform. Further, using the infrared patterns system, we get the coordinates of the UAV relative to the mobile platform. And this makes it possible to solve the localization problem quite accurately. However, this does not solve the problem of barcode scanning, and as discussed in the first paragraph of this chapter, barcodes are used in 87% of warehouses. But in the process of creating a method for scanning barcodes, I also realized that it is possible to improve the localization of the UAV by using barcodes as landmarks, which was done using the method of active perception described in Chapter 6. Of course, solving the localization problem can be easier by using an array of active sensors (UWB and ultrasonic), but this contradicted the thesis concept and the requirements of warehouses for the implementation of additional active infrastructure. These requirements are placed in chapter 3.1 and taken into account in the impact model of my research.

Question #6 Also, you do not need to provide the solution in this section - this section is about Problem. Please break down the problem and discuss it in more details, so that it would be clear what is the exact problem you address in your thesis.
Answer. In this section, I wanted to briefly formulate the problem, slightly touching on the presented solution, so that external readers of my thesis, after reading the abstract, can immediately evaluate on the basis of this section how relevant reading my thesis is for them. And then, in Chapter 2, the entire background of my research is devoted to studying the problem and analyzing existing solutions. And in Chapter 3, on the basis of this, the final reference and impact models of my research are built.

Question #7 Introduction - What is the exact contribution (move Section 8.2 in Intro) and novelty of your thesis (it is described later on p.68, but must be in Intro)? What do you put on top of the state-of-the-art? I strongly recommend to discuss it in Intro.
Answer. In one of the first versions of my thesis, I put section 8.2 in the introduction. But following the recommendations of the individual committee member, the following narrative chain was chosen. Because that was expected that I stated the problem, the gap, etc., then explained the proposed methodology, presented the literature and patent surveys, presented my solutions in a global scientific perspective showing clearly how my thesis addressed the gaps and solved the scientific problem, and then summarized my contributions and novelty with some solid arguments, not before this was all done.
In terms of the state-of-the-art works in this field at the moment, there is no universal, fully working solution (Section 2.6). As I noted in section 2.6.1, works in this direction began to appear not so long ago (2018-2019) and were fragmentary, which made it possible to close certain parts of the described inventory problem (robot localization, reading tags, manual control).

Question #8 Section 2.2 - I hope you are aware of the point that a PhD degree is the academic degree. It is not an engineering/commercial/marketing project - it is about research. I am surprised that the first requirements are identified by a company. I do not see Literature review. I found just about 6 pages (pp. 54-59). Please make sure that a PhD thesis is about research and it is a good practice when the structured literature review comprises up to 30% of the thesis. Market review etc. is a nice add-on, but not the main part of literature review. Carrying out the literature review helps one to make the conclusions on the novelty and the gap(s) in the state-of-the art. If the problem is localization you might want to consider various methods based on RFID, UWB, ZigBee, FM radio, video cameras, methods for data processing/analysis, optimization methods, etc.
Answer. Since my dissertation is industrially oriented and based on the industry problem, which is quite typical for theses in the field of robotics, that is why the first requirements came from the industry, because it was important to solve an applied problem. Therefore, a fairly detailed analysis of industrial solutions, along with patents in the selected area, is provided. As mentioned in the answer to the previous question, there is no perfect solution at the moment. Therefore, in the literature review, the most relevant research to solve the problem of stocktaking and building a robotic solution are given, which are still few. Nevertheless, an overview of individual methods, even with the direct implementation of some parts of my work, is given in Subsections 5.5.2, 4.4.2, 6.1.2. For better understanding, I have corrected these section titles, gave links to them in Section

2.6. I also reorganized the literature review section and divided it into subsections to clarify how technology was needed to create the system presented in the thesis.
The review of such technologies as RFID, UWB, radio, etc., was not provided since they imply installing additional active infrastructure, which is unacceptable from the industry. Nevertheless, I consider it important to add them to the literature review and clarify this point in the thesis.

Question #9 p.68 - 'The main research issue will be just to study the impact of the proposed system on the stocktaking process'. I am puzzled with this statement. 'Main Research question: How autonomous heterogeneous robotic system will improve quality of stocktaking and decrease it's duration?'. Sounds like the robotic system is not the goal of this work. I would recommend to title section 3.4 as 'Thesis goal and objectives'. Almost all listed objectives are technical - what are the research objectives?
Answer. Section 3.4 was renamed as 'Thesis goal and objectives' according to the recommendation. The main question really needs to be corrected "How the creation of autonomous heterogeneous robotic system will improve quality of stocktaking and decrease its duration?", because one of the strongest sides of my thesis is a comprehensive analysis of the problem and the design of the most promising concept of the heterogeneous robotic system. The concept of the robotic system, i.e., its structure, operation scenario, and required functionality, is the main research focus. To validate the proposed solution, we had to resolve several technical objectives. Succinctly, three main scientific objectives of the thesis can be formulated as follows (also added to the thesis):

1. Design the concept of novel heterogeneous robotic system for automated inventory stocktaking
2. Develop a control system for the heterogeneous robotic system which does not require significant infrastructure changes and could provide:
- Autonomous and Precise (±2 cm accuracy) localization and navigation for UAV and UGR.
- Continuous operation with a quick start.
3. Create a novel method of effective and robust scanning for the autonomous heterogeneous robotic system, that could work:
- Faster than people (mean stocktaking speed).
- More precisely than people (% of errors during stocktaking).

Question #10 Figure 4-2 is pointless. It would make sense to put it in Appendix and substitute with a block diagram in Section 4.1.
Answer. I agree that the electrical drawing can be moved to the appendix, however I would like to mention it in the chapter on mobile platform design, because I have not provided any other materials for the appendix.

Question #11 Section 4.4. Why do you develop VR operation system if the heterogeneous URG-UAV is supposed to be autonomous?
Answer. It is absolutely true that the system should work completely autonomously. Nevertheless, during industrial research, there was often a request from the industry for the ability to supervise the operation of the system remotely. One of the directions of development of this system is the automatic detection of damaged pallets by ML. We assume that the workers could mark damaged pallets via the presented interface and thereby create a marked dataset for us while supervising the system. In addition to regular and partial stocktaking, it is sometimes necessary to conduct a detailed inspection of a specific pallet place. This can be done manually by calling the work team or through a developed interface that simplifies the control of the robotic system.

Reviewer: Prof. Jun Miura

Question #1 Related robotics and computer vision technologies cited in Sec. 2.6 should be organized using subsubsections based on the types of technologies (e.g., UAV control, path planning, SLAM, object recognition). Please summarize what technologies are necessary for the specific applications (automated stocktaking) at the end of this subsection.
Answer. I agree, the literature review was divided into sections technologies (UAV control, path planning, SLAM, object recognition), I also added a conclusion about what technologies are needed to create such a system.

Question #2 Why do you use the abbreviation URG for Unmanned Ground Robot? For the consistency with UAV, 'UGV' (unmanned ground vehicle) looks appropriate ("UGR" is actually used in Fig. 5-7).
Answer. I checked and found both terms UGR and UGV are often found in the literature, however, UGV is more often used, so I changed and use UGV in my thesis.

Question #3 About the notation of citation of papers, [Grisetti et al., 2010] (both in the parentheses) looks better than "Grisetti et al. [2010]".
Answer. I agree that it would have looked better, I tried to fix it, but unfortunately, I could not fix the Latex citation style.

Question #4 What is the strategy for covering all the areas of a warehouse? How do the heterogeneous robots operate collaboratively? There is a description of the sequence of robots' movements in p. 134. If this is the strategy, this should appear earlier in the section describing the method(s), not in the section describing experimental results.
Answer. Page 134 outlines the strategy for running a particular experiment, therefore a description has been placed in this section. I agree that the general description of the fly-over strategy during the inventory should have been posted in the thesis earlier. I added it after section 5.1. It is worth noting that with the implementation of the active perception method, the last steps of this strategy change according to Algorithm 2 from section 6.2.

Question #5 Please describe why you use a segmentation method (U-net) for barcode detection instead of detector-type methods such as SSD.
Answer. Initially, we wanted to classify each pixel from the UAV camera to understand whether it is a barcode or not. This is exactly the task of segmentation. This approach gives us the opportunity for more flexible post-processing, including in terms of changing lighting. In addition, in post-processing, we can cluster pixels in the area of the intended barcodes and add their centers as waypoints of the UAV route. It is also worth noting that according to preliminary benchmarks running on the selected processor (Nvidia Jetson Nano), SSD ResNet and U-net shown the same FPS at lower image resolutions at the SSD input [1], therefore, with the equal resolution at the input of both, the FPS of the SSD will be lower than that of the U-net.
[1] https://developer.nvidia.com/embedded/jetson-nano-dl-inference-benchmarks

Question #6 Please assess the quality of the generated map by SLAM (e.g., Fig. 5-1).
Answer. Since the purpose of this work was not a comparative analysis of SLAM methods, we focused more on assessing the accuracy of the UGV and evaluated the created map qualitatively. Figure 5.1 shows a high-quality map of the warehouse since it shows the same number of corners as the warehouse environment. Moreover, it can be seen that the map is fully enclosed. Such qualitative metrics are used in [2] to compare SLAM algorithms. It is also important to note that the presented map matches the geometry and dimensions of the warehouse map where this experiment was carried out. It is important to note that the final map is of sufficient quality, nevertheless, at the beginning of its construction process, the quality was worse and there were unclosed areas.
[2] A. Filatov, A. Filatov, K. Krinkin, B. Chen, and D. Molodan, “2D SLAM Quality Evaluation Methods,” Aug. 2017.

Question #7 Since UAVs have a barcode detection ability, the X and Y direction's precise localization may not be necessary. Please describe clearly the required accuracy in localization and assess the experimental results for it. This assessment is also related to showing the advantage of a heterogeneous robot team over a team of (homogeneous) UAVs.
Answer. The UAV can accurately detect barcodes and their position by combining data from its camera and accurate data about its location coming from the UGV (Algorithm 2). Therefore, a separate UAV cannot perform the task of localization in space without knowing its global coordinates in the coordinate system associated with the warehouse. A group of UAVs can be used together with one UGV, which will increase the speed of inventory. The development of such an approach is discussed in Section 8.5.

Question #8 I do not fully understand the "Localization improvement with active perception." The active perception is for locating the UAV to an appropriate position for barcode recognition. In this case, the absolute localization (in the world coordinate system) does not matter very much. However, in Sec. 7.3.2, the IR-based localization and the active perception-based localization are compared. What does this result show?
Answer. Initially, the active perception was designed to fly to a position suitable for reading a barcode. Only on the basis of the coordinates of the barcodes, the UAV will not be able to accurately determine its position at the current level of development presented in the thesis. During experiments without barcodes, it was found that positioning accuracy of IR-based localization decreases with increasing altitude. Therefore, barcodes with their coordinates are taken as an additional source of information. And to determine the position of barcodes and flight to them, we must use active perception. The experimental results in section 7.3.2 just show that the use of barcodes can improve the results of IR-based localization accuracy by means of active perception. This is an important result, which in the future will allow us to abandon the use of the global coordinates of the UAV relative to the UGV, and will allow the UAV to fly independently in a much larger working area, using the UGV only to recharge and deliver a swarm of UAVs to one alley.

Question #9 The third objective of the thesis (p. 69) is to ensure a long operating time of the system in terms of repeatability of experiments. Did you conduct experiments to show that?
Answer. To achieve the repeatability of the experiments, the soft impedance landing method described in section 5.5 was developed. As part of its testing, a series of experiments were carried out, where we evaluated the repeatability and softness of the landing. As part of the experiments in the warehouse, we also looked at the repeatability of experiments, but did not report the results of evaluations separately. Nevertheless, it should be noted that such an experiment will be carried out when the contact charging of the UAV on the UGV is completed.

Reviewer: Prof. Clement Fortin
Question #1 My only request for a potential change is the acronym of the term “Unmanned Ground Robot” which is represented by the acronym URG instead of UGR. I leave it to the candidate to decide about this very small change.
Answer. I have fixed this misprint. Moreover, according to the recommendation of Prof. Jun Miura, I changed the term “Unmanned Ground Robot” to “Unmanned Ground Vehicle” (UGV), which is used more often in scientific papers.

Reviewer: Prof. Riichiro Tadakuma
Question #1. If there are some more videos that recorded real motions of the unmanned aerial vehicles and the unmanned ground robots used in this dissertation, they will enhance its academic contents. But of course the present video of WareVision introduced in YouTube is quite enough to understand the scientific significance of the results obtained in this dissertation
Answer. To enhance the academic contents of the dissertation, according to the recommendation, direct links to video materials of experiments with the presented system were added at the end of section 8.1.

Other changes:
Section 4.4 was modified and extended by the user study.
Two accepted papers in co-authorship were added:

1. A. Petrovsky, Ivan Kalinov, P. Karpyshev, M. Kurenkov, V. Ilin, V. Ramzhaev, and D.Tsetserukou. Customer behavior analytics using an autonomous robotics-based system. In IEEE 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), 2020. Core2020 A, Scopus and WoS, H-index(SJR)=19.
2. P. Karpyshev, V. Ilin, I. Kalinov, A. Petrovsky, and D. Tsetserukou. Autonomous Mobile Robot for Apple Plant Disease Detection based on CNN and Multi-Spectral Vision System. In IEEE/SICE International Symposium on System Integration (SII 2021), IEEE, 2020. Scopus and WoS, H-index(SJR)=12.