

Vision-Based Robot Control in the Context of Human-Machine Interactions

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Andrzej Nycz
August 2012

Copyright © 2012 by Andrzej Nycz
All rights reserved.

Acknowledgment

I would like to thank Dr. William R. Hamel, my advisor for his guidance and patience during my graduate studies. I would like to thank my parents and my family, especially my brother Stanislaw for the ongoing support and motivation. I would like to thank Dr. Mark Noakes for years of great cooperation and friendship in the laboratory and his valuable input in the area of electrical engineering and teleoperation. I would also like to thank fellow graduate students: Renbin Zhou, Heather Humphreys and, especially Matt Young for his long and successful collaboration in the developing of the Tracking Fluoroscope System.

I would like to thank the inspiring and motivating group of people I had a pleasure to work with at the Industrial Institute for Automation and Measurement for introducing me to the world of robotics and research.

I would also like thank the whole crew of Skydive East Tennessee for introducing me the sport and providing the unforgettable entertainment and recreation during my study years.

Abstract

This research has explored motion control based on visual servoing – in the context of complex human-machine interactions and operations in realistic environments. Two classes of intelligent robotic systems were studied in this context: operator assistance with a high dexterity telerobotic manipulator performing remote tooling-centric tasks, and a bio-robot for X-ray imaging of lower extremity human skeletal joints during natural walking. The combination of human-machine interactions and practical application scenarios has led to the following fundamental contributions: 1) exploration and evaluation of a new concept of acquiring fluoroscope images of musculoskeletal features of interest during natural human motion, 2) creation of a generalized framework for tracking features of interest in visual data, and 3) creation and experimental evaluation of a vision based concept of object acquisition suitable for efficient teleoperation.

Several methods were proposed for motion control based on image sensing and processing. These methods were implemented and experimentally evaluated in both application contexts. The fluoroscopy tests included emulated joint tracking using a mechanized mannequin, and actual skeletal joint tracking trials conducted on 30 human subjects. The teleoperation assistance framework was tested using a full-scale telerobotic remote manipulator system with realistic task geometries, loads and lighting conditions.

The proposed methods and approaches produced promising results in both cases. It was demonstrated that the vision-based servo control using fluoroscopy images is an effective way to track human skeletal joints during natural movements. The telerobotic demonstration showed that visual servoing is a feasible mechanism to assist operators in acquiring and handling objects of interest in complex work scenes.

Key words: motion control, visual servoing, human-machine interaction, telerobotics, fluoroscopy, teleoperation.

Table of Contents

A. Introduction.....	1
X-ray Tracking Bio-Robot.....	5
Fluoroscopy description	5
Current Fluoroscopy Procedure	5
New Approach in Fluoroscopy	7
Potential advantages.....	7
Challenges	8
Framework for tracking features of interest in human fluoroscopic images	9
High dexterity telerobotic tasks suitable for vision servoing	11
B. Fundamental Contributions.....	15
Motivation.....	15
Fundamental contributions	16
B. Literature review	17
Visual servoing basics.....	17
Image analysis basics	17
Simultaneous camera or object of interest motion.....	17
Frameworks for visual servoing	18
Human-Machine Interface.....	18
Robotics for medical applications	19
Fluoroscopy imaging for surgery navigation.....	19
Fluoroscopy based guided surgery.....	20
Mobile robotics guidance	20

Early Tracking Fluoroscope System Study.....	20
Vision master control	21
Other.....	21
C. Concepts.....	23
Visual servoing agents and strategies	25
Human subject body tracking	25
Network distributed data processing and communication.....	25
Locating the object of interest in the image	25
Safety systems	25
Analysis of the gait and acquisition speed for tracking	26
Agent structure.....	26
Control interface.....	26
State machine analysis.....	26
Dual vision feedback mechanism.....	26
Human control of the test stand	27
X-ray video display and recording.....	27
Testing procedure	27
Analysis of the video feedback	27
Post analysis of the images.....	28
D. Experimental Environment and Constraints.....	29
Hardware and Software Overview	29
Tracking Fluoroscope System.....	29
Teleoperation	30
Tracking Fluoroscope Limitations	31

General operator's and development constraints for the teleoperation	31
E. Concept Execution.....	33
Visual servoing agents and strategies – Tracking Fluoroscope System.....	33
Image to real world coordinates	34
Vision tracking feedback adaptive gains	35
Asynchronous mode of execution.....	37
Position units.....	39
Preventing excessive motion	39
An example of full motion sequence	40
Modes of operation	41
Visual servoing agents and strategies – Teleoperation System.....	42
Vision servoing	42
High level Controller development	43
Master control	45
Network distributed data processing and communication – TFS and teleoperation.....	48
Networking.....	48
Inter Process Communication (IPC).....	51
Locating the object of interest in the image	52
Tracking Fluoroscope System.....	54
Tracking.....	55
Teleoperation.....	60
Template matching	62
Defining threshold score.....	64
Tracking Fluoroscope Safety Systems.....	65
Analysis of the gait and acquisition speed for the TFS	69
Knee displacement calculation	69
Additional data.....	70
Agent structure - TFS.....	74
Camera and network management	74
Video processing agent	75

Video transfer agent	77
Video watchdog agent	77
Main communication loop agent	77
Image processing computer to remote host.....	77
Remote host to the image processing computer.....	78
Laser data processing agent.....	78
Laser control agent.....	79
Video recording agent.....	79
Watchdog agent.....	80
Wheel control agent.....	80
Joint tracking control agent	81
MCU communication agent.	82
Interface agent.....	82
Agent structure – Teleoperation	82
Control interface.....	83
Tracking Fluoroscope System.....	83
Teleoperation system.....	86
Fluoroscope state machine analysis.....	89
Network vision data state machine.....	91
Network vision state machine.....	91
Teleoperation system state machines.....	92
Dual vision feedback mechanism in the fluoroscope system.....	92
Human control of the TFS test stand.....	94
X-ray video display and recording.....	95
Analysis of the X-ray video feedback	99
Image tracking analysis	101
Human subject body tracking in the TFS	103
Choice of measurements	104
Scanner setup.....	104
Method of finding the position and orientation	105

F. Experimental Results.....	110
Evaluation of the vision based approach of acquiring fluoroscopy images	110
Body Tracking.....	110
Body angle estimation	110
Body distance estimation.....	112
Active distance control.....	113
Knee tracking with a camera	114
Stationary tracking	114
Dynamic tracking.....	115
Knee tracking with the use of X-ray	115
Stationary tracking	115
Dynamic tracking.....	116
Framework results for tracking features of interest in human fluoroscopic images	117
Range of motion	117
Range of activity	118
Range of applications	118
Range of image features	118
Additional framework functionality	118
Data flow	119
Teleoperation framework results	121
Whole Arm Manipulator	121
Behavioral agent	121
Vision based agent	122
Manual control interface	124
Hybrid modes	124
G. Summary and Future Work	125
Summary	125
Review of Contributions	125
Future work.....	126
E. References	128

<i>Appendices.....</i>	<i>137</i>
Main Control Unit sample code	138
Image processing unit sample code	141
<i>Vita</i>	<i>145</i>

List of Figures

FIG. 1. PATIENT PERFORMING A DEEP KNEE BEND IN A C –ARM.....	6
FIG. 2. KNEE IMPLANT X-RAY IMAGE.	9
FIG. 3. AN EXAMPLE OF X-RAY IMAGE OF CROSSING LEGS.	9
FIG. 4. SHOULDER IMPLANT.	10
FIG. 5. ANGIOPLASTY EXAMPLE (BEFORE AND AFTER).....	11
FIG. 6. TYPICAL BOLTING/UNBOLTING OPERATION.	12
FIG. 7. TYPICAL OBJECTS FOR D&D TASKS.	13
FIG. 8. EXAMPLE OF AN INSERTING TASK.	14
FIG. 9. X-RAY INTENSIFIER MOCKUP WITH A CAMERA MOUNTED.	29
FIG. 10. DUAL ARM TELEOPERATION SYSTEM WITH A MOCKUP.	30
FIG. 11. LINEAR DRIVES WITH THE ENCODER SHOWN.	33
FIG. 12. RELATION OF IMAGE TO REAL WORLD OBJECTS.....	34
FIG. 13. PROJECTION OF THE OBJECT OF INTEREST	34
FIG. 14. RELATION OF MEASURED ERROR AND ERROR USED TO FOR CONTROL SIGNAL	36
FIG. 15. IMAGE AND MOTION UPDATE ON THE TIMELINE.....	37
FIG. 16. MOTION OUTPUT RAMPING.	38
FIG. 17. MOTION AGENT INTERFACE.....	42
FIG. 18. HLC DATA PROCESSING.	43
FIG. 19. TITAN II LOW LEVEL CONTROLLER.....	44
FIG. 20. WAM COMMUNICATION DIAGRAM.....	45
FIG. 21. NETWORK TOPOLOGY.	48
FIG. 22. TELEOPERATION SYSTEM TOPOLOGY.....	50
FIG. 23. ARTIFICIAL MARKERS.....	54
FIG. 24. X-RAY IMAGE OF ARTIFICIAL AND NATURAL KNEE.....	55
FIG. 25. KNEE TRACKING WITH A MARKER.....	56
FIG. 26. X-RAY PROJECTION DIAGRAM.	56
FIG. 27. X-RAY IMAGE PREPROCESSING SEQUENCE	57
FIG. 28. ARTIFICIAL KNEE TRACKING.....	58
FIG. 29. KNEE TEMPLATE.	58
FIG. 30. KNEE WITH THE REGION OF INTEREST MARKED.....	59

FIG. 31. REGION OF INTEREST EXTRACTED FROM THE IMAGE.....	59
FIG. 32. MANUAL DEFINITION OF THE REGION OF INTEREST.	60
FIG. 33. SERIES OF IMAGES FEATURE LEARNED AUTOMATICALLY.	61
FIG. 34. CORRELATION PROCESS OF THE TEMPLATE AND THE IMAGE.	62
FIG. 35. CREATION OF THE MATCHING TEMPLATE.	63
FIG. 36. TRACKING FLUOROSCOPE SYSTEM DIAGRAM.....	66
FIG. 37. SONAR AND BUMPER SETUP.....	67
FIG. 38. EMERGENCY STOP BUTTON CONFIGURATION.....	68
FIG. 39. X-RAY CONTROL INTERFACE.....	69
FIG. 40. HUMAN GAIT CYCLE.	70
FIG. 41. GEOMETRY OF FEATURES OF INTEREST IN THE X-RAY IMAGE.....	71
FIG. 42. TFS SYSTEM TOPOLOGY. COLORS ARE REPRESENTING THE DATA EXCHANGE AND SOURCE.....	75
FIG. 43. WHEEL CONTROL SYSTEM.....	81
FIG. 44. TELEOPERATION SYSTEM TOPOLOGY. COLORS REPRESENTS THE DATA EXCHANGE AND SOURCE.	83
FIG. 45. TFS CONTROL INTERFACE.....	84
FIG. 46. HIGH LEVEL CONTROLLER GUI.....	86
FIG. 47. MAIN TELEOPERATION CONTROL INTERFACE.....	87
FIG. 48. REMOTE CONSOLE STATION.	88
FIG. 49. TYPICAL STATE MACHINE STRUCTURE.	91
FIG. 50. TFS WITH THE KNEE ROBOT.	92
FIG. 51. IMAGE WITH FULL OVERLAY: OBJECT OF INTEREST AND BODY POSITION.	96
FIG. 52. DISPLAY AND RECORDING INTERFACE.	96
FIG. 53. IMAGE MATCHING ANALYSIS TOOL.	101
FIG. 54. KNEE IMAGE WITH OVERLAY INFORMATION ADDED.	102
FIG. 55. X-RAY IMAGE WITH AND WITHOUT THE OVERLAY INFORMATION.	103
FIG. 56. TRACKING FLUOROSCOPE SYSTEM.....	103
FIG. 57. AN EXAMPLE OF SCANNERS POINTING AT MODEL’S SHOULDERS, MAXIMUM DISTANCE.....	105
FIG. 58. DATA BEFORE AND AFTER FILTERING. TOP - UNFILTERED, BOTTOM – FILTERED.....	106
FIG. 59. LINEAR REGRESSION LINE.....	107
FIG. 60. LINEAR REGRESSION APPLIED TO DATA POINTS.	108
FIG. 61. ORIENTATION LINE OBTAINED THROUGH SECOND AREA MOMENT METHOD.	109
FIG. 62. ORIENTATION ESTIMATION ERROR.....	111
FIG. 63. LASER TRACKING DISTANCE ERROR.....	114
FIG. 64. IMAGE POSITIONING ERROR IN THE DIRECTION OF WALKING.	114

FIG. 65. IMAGE POSITIONING ERROR IN THE VERTICAL DIRECTION.....	115
FIG. 66. HORIZONTAL TRACKING ERROR WITH THE USE OF X-RAY.	116
FIG. 67. VERTICAL ERROR WITH THE USE OF X-RAY.	116
FIG. 68. FEATURE TRACKING DATA FLOW DIAGRAM	119
FIG. 69. VISION SERVOING DATA FLOW DIAGRAM.	120
FIG. 70. IMAGE AND DATA STORAGE DIAGRAM.	120
FIG. 71. THE MAIN MOTION CONTROLLER.	138
FIG. 72. WHEEL MOTION CONTROLLER.....	138
FIG. 73. COMMUNICATION LOOP.....	139
FIG. 74. INPUT/OUTPUT AND INTERFACE LOOP.....	140
FIG. 75. VISION COMMUNICATION LOOP.....	141
FIG. 76. IMAGE TRANSFER LOOP.	141
FIG. 77. CONNECTION MANAGER LOOP.....	142
FIG. 78. LASER DATA PROCESSING LOOP.....	143
FIG. 79. LASER DATA COMMUNICATION LOOP.	143
FIG. 80. VISION ACQUISITION AND PROCESSING LOOP.....	144
FIG. 81. VIDEO WATCHDOG LOOP.....	144

List of Tables

TABLE 1. LIST OF CONCEPTS AND RELATIONSHIPS TO GENERAL PROBLEMS.....	23
TABLE 2. MOTION OUTPUT RAMPING.....	38
TABLE 3. LIST OF TYPES OF MATCHING TEMPLATES AND THEIR PROPERTIES.	64
TABLE 4. AN EXAMPLE OF THE VISION CONFIGURATION FILE.	76
TABLE 5. ORIENTATION ANGLE TEST. VALUES MEASURED AND CALCULATED.	110
TABLE 6. ORIENTATION ANGLE OBTAINED USING TWO DIFFERENT METHODS.	112
TABLE 7. BODY DISTANCE ESTIMATION.....	113

A. Introduction

Robotics and artificial intelligence emerged as very promising science disciplines that could lead to dramatic increases of the human life comfort and safety. The popular literature and cinema were presenting a future full of intelligent humanoidal robots. However, those high expectations were never met, and robots have not become intelligent companions of humans, neither servants or fully self aware toys. One of the reasons of this failure is robot's inability to acquire process and reason the information necessary for navigation and task execution in random dynamically changing environments. Among five senses possessed by humans only, three provide such information: vision, hearing and touch. Among these three, vision is the one with the highest possible range of applications and perspective. This fact is consistent with routine observations of nature and daily experience.

Without vision, a human can still achieve his/her goals but it becomes tremendously harder than with vision. There are examples of creatures relying on acoustics (bats) as the main sense for navigation but, among mammals and land animals, it is of lesser importance.

The development of vision guided navigation started with the introduction of digital cameras and microprocessors. The computational power and quality of cameras increased drastically, however the practical usage of vision information remains limited. The algorithms used to process image data are highly advanced but usually only applicable to certain fields, environments or light conditions; and one can be almost certain to face situations in which any algorithm fails.

Another challenge of current trends is increased image resolution leading to increased computational power, but without corresponding progress in the area of image understanding or reasoning. The human eye is capable of acquiring millions of pixels but the focusing mechanism allows us to see clearly within the cone of 20 degrees [1]. The rest of the image is blurred. Even further data elimination takes place in the brain where only useful data is really processed and the majority of the vision input is simply ignored [2]. Another important fact is

that the understanding of the image also comes from the database built during our life. People with vision limited to one eye are still able to estimate sizes distances and three-dimensional geometry. Temporary memory also plays a key role in vision navigation. Every movement decision takes into consideration the information acquired in the past, its prediction and provides predictions of what could happen during the next step. Finally it can be said, that nature forced a safety feature in order to constantly monitor, improve, adjust or abort decisions based on visual data. For example, the eye blinks as well as auto focuses in adaptation to light intensity. When these fail it is natural for a human to stiffen the body, take a more stable posture, enhance sound perception and dynamically increase the tactile signal input while constantly trying to acquire clear vision data.

There are many types of intelligent robotic systems. These types are differentiated by factors such as fundamental capabilities, complexity, and levels of human interaction, to name a few. The synergy and integrated-effect of such factors, in great part, define specific underlying research challenges associated with the realization of such systems. Human robot interaction, particularly the class that involves human proximity to operating mechanisms, leads to very complicated research issues pertaining to human safety, and human-machine physical and cognitive interfaces. In this research two specific classes of intelligent robotics systems involving significant human interactions are considered; 1) a high dexterity telerobotic manipulator system, and 2) a bio-robot that can X-ray image human skeletal joints during normal walking. Even though these are very different systems, they have common challenges that are to be explored. These challenges are:

- Motion control based on visual servoing
- Machine vision when both the camera and/or the object of interest are undergoing dynamic motion
- Machine vision object detection/recognition for tracking
- Practical realistic environment
- Human–robot physical interaction
- Human–robot cognitive interaction

- Human–robot interfacing.
- Framework development
- Human and hardware safety

Visual servoing is a way to control a robot, manipulator or simple actuator's position through providing a vision based control signal. Important parts of such systems are: camera, interface, processing unit and motion controller with appropriate interface. The camera provides the image through an appropriate interface to the processing unit (PU). Once the image reaches the PU, it is processed to extract the features of interest and establish their position in the image. Then the image coordinates are transformed into the robot's coordinates. The next step is the creation of the motion control signal error by comparing the desired and actual feature of interest coordinates. Since the frequency of image acquisition is usually slower than motion control, it requires a proper synchronization and adaptation. The image may be acquired continuously or on demand, which may cause the image processing part may not to be synchronized, i.e., significant time delays may occur and the motion controller must to be able to compensate accordingly.

Continuous object tracking based on visual servoing presents additional challenges. When the object of interest is not holding its position relative to the camera the tracking algorithm has to be able to adapt to the new conditions. This may be caused by the camera motion or object motion. As a result the projected image of the object may change its shape, color, texture, or size. Additionally occlusions, illumination changes and reflections may occur. When objects of interest have irregular shapes or of non distinct visual features, object tracking becomes very difficult.

The human machine interface (HMI) is a crucial part of a robot system when behavior is directly determined or dependent on human interaction. Two HMI forms may exists: physical hardware interface (like buttons or touch screens), that directly and unambiguously provides a control signal to the robot or a system of sensors that is intelligently trying to acquire and interpret human intentions. Both types can form a hybrid system as well. The first type is the most common. An example of such interface is a keyboard or a joystick. When used with a

proper software or hardware there is no uncertainty regarding operator's intention. The design of this type of interface focuses on the ease of use, intuitiveness and minimum fatigue. The second type has different challenges. The main task is to sense/interpret/infer the intention of the human. It may be in the form of camera sensors on an intersection or an active servoing agent helping the surgeon stabilize the motion of the surgical tool. This type may involve human motion/body language or nuances that are very task specific.

One of the most challenging tasks in machine vision is object recognition/tracking. The goal is to accurately identify the object of interest in the image and provide its position or orientation. The object image size is usually relatively smaller than a whole image itself. There might be objects of similar shapes in the background that can be falsely recognized. Another key factor is the processing time. Many robotics applications require near or real-time performance which can be anywhere from a few cycles per second to a couple hundred Hz. There are several techniques for locating the object in the image. One of the most common one is pattern matching. It is usually based on a correlation factor. A reference image is compared with the actual image and through best fit criteria the location of the object is found. This can be achieved through image intensity matching or shape matching. Such techniques show good performance with blurred images or varying intensity. Depending on the object visual nature, color matching or morphology analysis may be used as well. If the object of interest has a unique color that distinguishes it from the background color matching can provide very good and fast results. In a case when the object has proper contrast and certain unique features like illuminated holes, openings or edges, morphological analysis can be very efficient, although a downside of this method is sensitivity to changing illumination.

Creation of a framework for robotic vision application is very type specific. On the common level there is usually a vision sensor, vision interface, PU and motion interface. When it comes to more specific task more factors come into play. Starting with the camera acquisition, the system may require more than one camera or interface. Independent of the amount of hardware sensors the system should be able, in timely manner to process the data and react accordingly. When the vision task is changing but still belongs to the same class of

problems the framework should be versatile enough to function with it. It also applies to human interaction-it should be able to interact with more than one operator or user. On the development side it is beneficial when it is easily upgradable, extendable or modular.

The safety of human and hardware in robotics systems is a necessity. Importance of emergency switches, fuses or contact sensors is obvious. However in advanced system there is a need for more sophisticated devices like sonars or light fences. The use of advanced sensors requires extra reasoning in interpreting them. For example, in a case of a scanning and tracking system, the loss of an object might be an indicator of a coming collision. Similarly in a case of a measuring system, values that are out of expected range might suggest a risk of damaging the hardware or harming living objects. That applies also to human control over the system. The operator may unintentionally bring the system into a dangerous state. Especially when it comes to human machine interaction extra care should be taken in analysis of possible human and machine errors and unexpected behavior.

X-ray Tracking Bio-Robot

Fluoroscopy description

Fluoroscopy is an imaging technique used to obtain real-time images of the internal structure of a human body. The fluoroscope is a device consisting of X-ray tube, receiver in a form of an image intensifier or flat panel and interface hardware and software necessary for presenting and/or storing the images. The frame rate can vary anywhere from a few frame per second to sixty frames in modern system. The common areas of fluoroscopy procedures include: Investigations of the gastrointestinal tract, urological surgery, angiography, orthopedic surgery, implantation and placement of lifesaving devices (i.e. pacemakers) and pre and post surgery evaluation of bone implants.

Current Fluoroscopy Procedure

The primary interest of this research is acquiring fluoroscopy images of the musculoskeletal system of human patients in motion. The conventional way of accomplishing it



Fig. 1. Patient performing a deep knee bend in a C –Arm.

usually involves a C-arm (Fig. 1), an active participation of the assistive personnel, X-ray technician and the fixtures necessary to obtain desired type of motion. C-arm is a fluoroscopy system mounted on a C-shape frame structure to ensure alignment of the X-Ray tube and the receiver. Due to its nature it is a massive device in size and dimensions, intended to be used stationary.

Very often during the procedure, a person instructing and watching the patient is necessary. The patient can be asked to perform a set of specific activities which may involve: bending the knee, natural simple walking, climbing up/down the stairs or going up or down the slope. Obtaining dynamic, not stationary images is crucial since an object acts and looks different in motion. That allows for better reasoning about kinematics and dynamics of the bone/joint. Since the C-arm is stationary it may require several attempts with changing the gear and repeating the procedure. Due to the limited area of the imaging device the patient has to execute the activity within the field of view of the device. This may cause fatigue and lack of comfort, especially with elderly patients. It also involves many hours of the assistive personnel and X-ray technicians which directly translates into high costs. The complexity and time required to spend on the test may also cause the patients not to act as they are requested to, resulting in poor images without the features of interest. Another factor is the total time of X-

ray exposure. The obvious goal is to do minimum exposure to the patient while achieving all the goals of the procedure. Easy, straightforward short and low dose fluoroscopy evaluation of the object of interest is always in the best interest of the patient and hospital crew.

New Approach in Fluoroscopy

In 2004 the University of Tennessee started working on a new concept of acquiring fluoroscopy images. After years of research in the area of orthopedic surgery and implant design and evaluation a new idea was formed. The proposed key change was in mounting the X-ray components on a mobile robot and making it autonomous. Its main role is be able to position the fluoroscopy system in such a way that the patient can execute a natural motion without any physical distraction while the onboard actuation, sensing and computation systems are responsible for moving the robot, keeping the objects of interest in the field of view and most importantly, continuously checking for safety of people and hardware. The robot has to be able to intercept the imaging signal, process it and reason about the position of the object of interest with respect of the robot frame. It also has to sense and keep the whole body in a safe distance. In order to do such a complex actuation, sensing and computing system is necessary.

Potential advantages

The benefits and features of such a device would possibly include:

- Fully automatic operation
- Precise tracking of features
- Improvement in the quality of the images
- Decrease in the time required to complete the procedure in its total length and numbers of attempts/visits in the hospital
- Limitation of the exposure time for the patient and personnel

- Satisfaction of the patient
- Ease of use
- Wireless and untethered and mobile design
- Decrease in the cost of the procedure

Challenges

This new concept brings new challenging problems and aspects to explore and evaluate besides the technical difficulties. One of the most important is the safety of the patient. It includes human-machine interaction in motion. It involves high accelerations and speeds, complex motion and fast dynamic reaction of the system. Since the device is creating radiation this aspect has to be covered meticulously, considering all the details. Another problem to address is the dynamic side of the concept. The device, besides being safe has to offer certain level of performance characteristics. The third aspect to explore is of imaging/machine vision nature. The feature of interest in the image has to meet at least four criteria: it has to be present in the field of view, be possible to identify and find its location, be of good quality and processed in real time.

Processing of the X-ray image presents several new problems to study. The joint of interest will be slightly different from a person to person. Due to motion it can change its shape and size (Fig. 2). The brightness of the bone can change as well and several artifacts like blurring may occur. One of the challenging problems in tracking knees is the leg crossing issue. During natural walk legs are passing each other from the point of view of the camera. For a brief moment they are aligned and overlapping each other (Fig. 3). This may cause the system to track the wrong joint in the following tracking cycle or cause instability in the tracking algorithm.



Fig. 2. Knee implant X-ray image.



Fig. 3. An example of X-ray image of crossing legs.

Framework for tracking features of interest in human fluoroscopic images

The average age of human population increased significantly in the last century. As a result the number of people with musculoskeletal system injuries rose as well. Hence, the demand for improving the quality of life is becoming obvious. A similar situation applies to athletes. The constant pursuit of greater results and record breaking does not come without an impact on health condition. Many of those problems can be investigated deeper if physicians were equipped with more advanced diagnostics tools. Among them, there is a group of problems that could be approached with the use of active fluoroscopy. It includes but is not limited to pre-and post-surgery evaluation of: knee, ankles, hips, shoulders, elbows, wrists (Fig. 4).

All the mentioned body parts are joints used for locomotion or manipulation. In general in order to diagnose the injury or discomfort several common steps have to be made. In order to interactively and dynamically observe the feature of interest a proper type of robot/device is necessary. Assuming that, the following tasks are to be executed: safely placing the patient within the workspace of the device, initial analysis of the feature of interest, adaptation/calibration of the feature search engine, engaging the fluoroscopy and actuation system, acquiring the image, processing, reasoning, control loop activation and initial alignment motion, patient's predefined set of motions with real-time tracking, stopping the motion of the patient, actuation system, fluoroscopy and releasing the patient from the workspace. To a certain extent this concept can be applied to other areas where the range of motion is not high but it is important to provide a stable online imaging. In those cases, the motion is a result of a natural body function like respiration, body fluid flow or muscle stretch. An example of angioplasty is shown in Fig. 5. A vision based fluoroscopy system could be potentially used to control the motion and positioning of the catheter.



Fig. 4. Shoulder implant.



Fig. 5. Angioplasty example (before and after). Attribution: JHeuser at http://en.wikipedia.org/wiki/File:HWI_PTCA.jpg.

Research in cases like knee implant evaluation could provide information that is not available today; in other cases it can improve the imaging quality. Developing a generalized, framework for creation of new fluoroscopic systems with active human-machine interaction seems beneficial for all the future research.

High dexterity telerobotic tasks suitable for vision servoing

A similar idea can be introduced in teleoperation, which is a field of robotics concentrating on remote control and task execution. Both the object of interest and the camera can be in motion. A standard teleoperation system consists of [3]: operator's station equipped with master arm its control unit, control panel, speakers and displays, the remote arm with cameras and microphones, communication system and an operator. In general tasks can be executed through teleoperation, automatically or semi-automatically. In most cases the first one is the prevailing one. Unfortunately, it requires very well trained and experienced operators. Furthermore, controlling the arm for extended period of time is an exhausting task [4] that can lead to operator's errors and inaccuracies. It is very unlikely to eliminate the operator from the system in the near future [5], however developing a smart vision servoing system could in many situations improve the speed and quality of the execution, reduce fatigue as well as allowing for less experienced human operators to perform tasks. This is the case especially during operation requiring precise positioning and orientation/ alignment like bolting or assembly operation.

The proposed research can be used in many teleoperation activities like: bolting/unbolting, picking up objects, cutting, breaking or inserting objects during an assembly process. It is essentially limited only to the physical potential of the machine vision and tool availability. The mentioned tasks fall into the category of decontamination and dismantlement tasks (D&D) tasks, tooling and positioning, as well as space and underwater manipulation. From the interface point of view it also can have an impact on the augmented reality field. The types of tasks were chosen based on their rate of recurrence among other teleoperation tasks [6], usage of vision by the operator and potential gain in automating them.

Typical bolting operation may look like Fig. 6. The difficulty comes from the limited perception of distance and orientation/position when doing remote teleoperation. It is not easy to align the tool and the object. That is not an issue when doing hand operations. However, when it comes to doing the same task remotely with a robotic arm on the slave side and an input device on the master side, experience and a 3D imagination are very important. Very often it is necessary to perform the same operation on different objects. Vision assistance could greatly lessen the fatigue associated with this task as well as decrease the execution time.

A different task is handling or picking up objects of similar shape but not necessarily

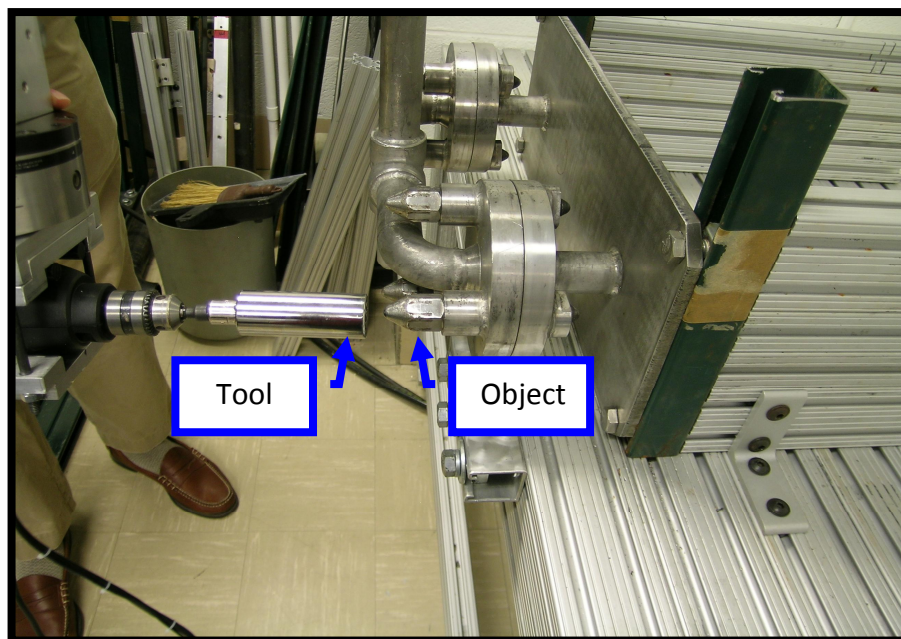


Fig. 6. Typical bolting/unbolting operation.

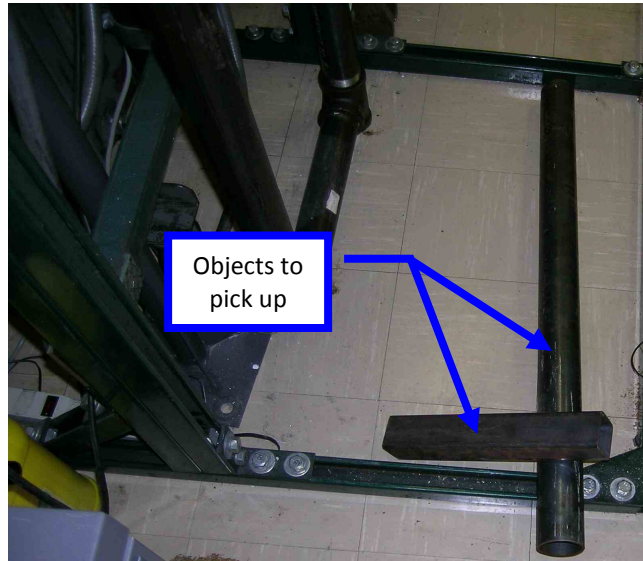


Fig. 7. Typical objects for D&D tasks.

specified color, texture, lighting condition, poor perspective or background. Those objects share the same grasp type [7]. However from machine vision perspective each object of that kind is completely different. Practical applications include cleanup after disassembly or decommission. The objects being split from a structure or plant needs to be picked up and stored in a bin or container for further transport. Very often the shape of the object is a cylindrical type with a large ratio of length to width (Fig. 7). The operator may need to pick up hundreds of pieces of that kind. It may include objects like pipes, extrusions c-shaped framing or unistruts.

Those operations, if executed with assistance of a smart visual servoing agent, could increase the performance of the D&D process.

A third type of task discussed in this proposal is an assembly task of inserting. It is a very common task that requires proper alignment of the object being inserted, for example a pin and the matching opening in the structure (Fig. 8). It can be a pipe, bolt or aluminum extrusion as well. Other scenarios may include insertion of a thermocouple into a well, pin into hole or bolt into thread hole. It is a very tedious operation, requiring high levels of experience and training.

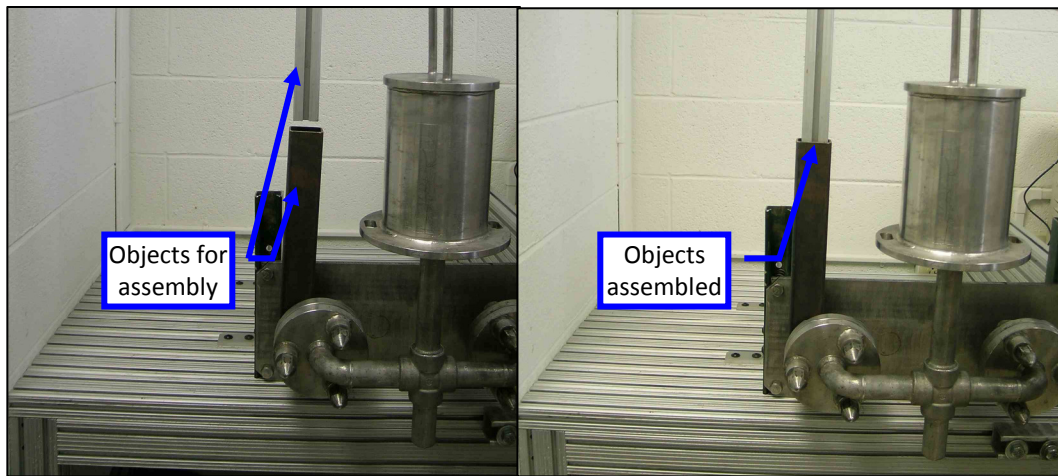


Fig. 8. Example of an inserting task.

B. Fundamental Contributions

Motivation

The fundamental goal of this research is to explore the combination of active motion control and visual servoing in machines that involve simultaneous camera and object of interest motion, all in the context of human machine interaction. This interaction is very important in the context of teleoperation. In developing advanced telerobotic systems it is important to properly interpret human operator intentions so the efficiency of the robot can be optimized. The lack of direct presence may result in loss of performance due to decreased precision, longer time required to execute tasks, decreased information about the environment, robot and its interaction. It also requires more experience and skills from the operators as well as greater sophistication in the mechanical part of the interface. This also has an economical effect in applications of robotics. For these reasons the need to improve teleoperation is important and can be accomplished by the introduction of vision servoing into the system.

Human-robot interaction is obviously an important issue in many types and applications of robotic systems. In addition to teleoperation, another application domain is considered in this research that can be considered biorobotics.

The application of visual servoing in human subject body tracking is a new area that can improve fluoroscopy. Currently, in order to diagnose a patient, design an implant or analyze its performance after surgery a standard C-Arm is used. It allows for very limited joint motion that lacks the natural dynamics of walking. Therefore, the researchers do not have complete data useful for implant development. In order to achieve greater range of motion, the X-ray technician can manually move the C-arm to track the patient's joint but that is a very difficult task and can only provide limited results due to poor tracking and short distances. Hence, it seems reasonable to explore a concept of building a robot that can perform real time imaging on a naturally walking subject. This can be achieved through usage of vision servoing with respect to the x-ray images from the fluoroscope.

The general problems that will be investigated in this research are summarized below.

- Motion control based on visual servoing
- Machine vision with dynamic motion
- object detection for tracking
- Practical realistic environment
- Human–robot physical interaction
- Human–robot cognitive interaction
- Human–robot interfacing
- Framework development

Fundamental contributions

In addition to the general explanation of the system engineering/integration aspects of human-robot interaction systems, this research has produced fundamental contributions in these three topical areas.

- Exploration and evaluation of a new concept of acquiring fluoroscopy images of musculoskeletal features of interest during natural human motion with active human-machine interaction.
- Creation of a generalized framework for tracking features of interest in human fluoroscopic images
- Creation and experimental evaluation of a vision based concept for autonomous and assistive objects acquisition during teleoperation procedures in unstructured environment.

B. Literature review

Visual servoing basics

The robotics discipline of visual servoing is a growing field gaining a lot of interest. A thorough view on the fundamentals was presented in [8]. The authors described basic coordinate transformations with respect to a camera and a robotic arm, two major classes of systems and general methods for tracking as well as common problems with existing systems. A detailed discussion on the class of image based servoing can be found in [9]. Aspects of feature extraction can be found in [10].

Based on this study it was decided that Image Based Visual Servoing is the right choice for the systems research in this work. This method does not require precise camera calibration and converges even with a poor calibration. Also, based on this study, a simple proportional controller was chosen for the implementation.

Although the general concepts of the visual servoing will be applied, the specific area of research will require solutions customized for the classes of robots used in this research.

Image analysis basics

A detailed study of machine vision and its application was presented in [11]. Vision concepts and its applications were also presented in [12]. These works had direct influence on the type of image tracking techniques used in this research. The main ones were: template matching, image filtering and illumination concerns.

Simultaneous camera or object of interest motion

In many applications, detection and tracking of objects must be performed while the camera is also in motion. A dynamic analysis of this problem is presented in [13]. A new way to estimate camera motion is described in [14]. Instead of trying to calculate one solution several hypothesis are being simultaneously computed and statistically evaluated. A real-time system using this statistical approach was also presented in [15].

The reviewed research did not address the problems of an X-ray in motion or an unstructured environment. The aspects of human joints in motion being tracked in close proximity to an X-raying device were not found in the literature.

Frameworks for visual servoing

Due to high complexity and variety in applications, frameworks for visual servoing vary significantly. A system structure that allows for use of various configurations is shown in [16]. The main goal was to create a framework that allows for the use of plug and play sensor-based components. A different type of framework was presented in [17]. Authors discussed a new method of combining image based and position based visual servoing. They used a probabilistic method for integration of both servoing spaces. A framework for relative positioning is demonstrated in [18]. It used an Affine Visual Servoing Technique, observing changes of shape in the image contours based on planar objects. A task based framework is illustrated in [19]. The problem of grasping is distributed by five agents that depending on the stage of the task are being instantiated. Task-based frameworks are also explored in [20]. The system uses learning algorithms for docking with mobile robots. The task of homing is presented in [21]. A mobile robot with an omnidirectional camera was considered in the research to study the homing ability. The results were tested in a simulation. A surgery type application-based frameworks is shown in [22]. Authors investigate a visual servoing framework for laparoscopic surgery.

The concept of a framework for tracking human fluoroscopy features in human joints in motion was not attempted in the literature. The concept of frameworks for an unstructured environment was not fully explored, either.

Human-Machine Interface

The Human-Machine interface plays an important role, even in systems with a substantial amount of autonomy. A discussion on the impact of human actions is presented in [23]. The paper presents a way to address the impact of a human error prior to or during the interface design phase. The role of visualization in developing an interface is studied in [24]. A

special consideration was given to colors and human perception of colors. A human-machine interface for guided surgery was proposed in [25]. An augmented reality with gesture recognition is explored for surgical interventions. An interface for recording human gait is described in [26]. The system uses reflective markers and a camera system to grab the image and perform measurements. The experimental part was executed with the help of an electric motor bar and markers simulating the gait.

Although literature described many types of HMIs, the published results are not applicable here. Incorporation of a fluoroscopic device and a robotic platform created new challenges that were not attempted in the literature. The uniqueness of the application and novelty of the concepts being considered dictates further research.

Robotics for medical applications

In recent years the impact and interest of using robotics for medical application has grown rapidly. Main applications of robotics include rehabilitation robotics, surgery and biorobotics. Rehabilitation robotics can be in the form of intelligent wheelchair systems, active prostheses or customized hospital systems. Robotics for surgery play important parts in: orthopedics and brain, eyes prostate or spine surgery. When looking from the tool point of view it includes endoscopic devices and minimally invasive surgery. The bio-robotics aspects is related to artificial life, human like robots or artificial systems for investigating physiology of biological systems. Bio-robotics is considered to be the future of medical applications [27].

The area of medical robotics is developing rapidly, however the extensive study of robotic diagnostic devices interacting with people was not found. The application of mobile robotic platforms with X-ray capabilities is a new field.

Fluoroscopy imaging for surgery navigation

Visual servoing combined with fluoroscopy and active human interaction is a new area of bio-robotics without vast contributions in the literature. X-ray images are often used as 2D projections that are compared with 3D models of bones acquired through preoperative CT scanning and modeling. In this case the fluoroscopy and its analysis allows for 3D navigation

during surgery ([28], [29], [30], [31], [32], [33], [34]).

The fluoroscopy used in the cited research is used for supportive navigation and not active and dynamic feedback control. Its purpose is closely related with the augmented reality and visualization.

Fluoroscopy based guided surgery

A similar concept is computer aided or guided surgery where the image choreography is used for guidance [35]. X-ray Images may serve to close the feedback loop for medical tool placement [36] or as a means for trajectory planning in long bone fracture therapy, or it can serve as a part of a guiding system in radio-surgery treatment [37].

The area of using fluoroscopy for active motion control and joint tracking is not covered in this works. It is believed that some of the experience from the works presented in this dissertation can be applicable in this field in the future.

Mobile robotics guidance

Image sensing plays crucial roles in mobile robots. It is used in autonomous, semiautonomous or pure teleoperation tasks. A fully autonomous robot is described in [38]. It computes the displacements through tracking the pixels in consecutive frames of the stereo system. Autonomous operations are also presented in [39], where a vision system is responsible for the deploying and docking of a small robot scout. A different approach to vision is presented in [40]. A camera system is used to capture human gesture to control a mobile robot. In [41], though, the vision is coupled with a force feedback system and helps the operator to center and grab the object in the view of a mobile robot equipped with an arm.

An application of a mobile platform tracking human body during natural walk and joints of interest in close proximity was not found.

Early Tracking Fluoroscope System Study

The work presented in [42], [43] and [44] offered valuable information. The first prototype platform was built and initial tests of body tracking were conducted. A detailed

analysis was presented on the motion feedback performance. Based on this, the decision was made to use laser range systems for position feedback instead of wheel encoder feedback. Also the initial safety system structure was proposed. This led to the current TFS safety methods. The mechanical structure was proposed, and the geometric aspects were developed. A detailed study of the human gait was presented. Further analysis of [45] led to the analysis and performance requirements of X-ray the image intensifier.

Vision master control

An interesting aspect of vision in robotics is vision master control. In this approach, a camera system is placed in the master loop as an essential key link between the source of the master signal and slave arm controller. Such a solution is presented in [46]. It describes a markerless camera system for observing and analyzing human arm motions. Based on the image analysis, the slave arm is driven to perform the required tasks. A vision system without markers is also presented in [47]. It relies on the bare human hand and stereo-vision cameras to drive the slave arm. A technique with markers is used in [48]. In this paper the operator wears a glove with highly contrasting markers. The motion of the human hand is being captured by a camera system, and then it is analyzed to project the motion of the glove to the three finger gripper.

Research related to active control with the use of knee or ankle joints was not found. None of the systems used X-rays for imaging. X-ray imaging brings additional challenges that had to be addressed in this dissertation.

Other

Literature in the area of microassembly, space exploration, underwater operation, tooling, positioning and augmented reality in the context of visual servoing was also reviewed in a search of possible solutions and inspirations. Even though progress in the recognition and tracking of objects in motion has been made, the studies cases are often planar structures, polyhedral objects or objects with a CAD model. Those strategies struggle to offer

generalization. A detailed survey describing these issues are presented in [49]. More recent review was presented in [50]. The authors pointed out lack of generality and robustness. It appears especially in sensitivity to changing light conditions and unstructured or cluttered environment. The progress is noted in online learning, which also is reported in [51]. The strategy for self learning was employed in this research.

C. Concepts

The presented worked was done at the Robotics and Electromechanical Systems Laboratory of the Mechanical Aerospace and Biomedical Engineering Department of the University of Tennessee. Two kind of robotic platforms were used: the Tracking Fluoroscope System and a teleoperation system.

The described work was done in several major research areas. The general research topics were applicable to a number of concepts. These concepts were investigated and analyzed and then proper methods were executed on the two platforms. The relationship between them is presented in the table below:

Table 1. List of concepts and relationships to general problems.

General research topics	Concepts relationships to general topics	Scope
Motion control based on visual servoing →	Visual servoing agents and strategies →	TFS, Teleoperation
	Analysis of the gait and acquisition speed for tracking →	TFS
Machine vision with dynamic motion →	Visual servoing agents and strategies →	TFS, Teleoperation
	Analysis of the gait and acquisition speed for tracking →	TFS
object detection for tracking →	Locating the object of interest in the image →	TFS, Teleoperation
	Image tracking analysis →	TFS
	Analysis of the video feedback→	TFS

General research topics	Concepts relationships to general topics	Scope
Practical realistic environment →	Human subject body tracking →	TFS
	Dual vision feedback mechanism →	TFS
	X-ray video display and recording →	TFS
	Locating the object of interest in the image →	TFS , Teleoperation
	Testing procedure →	TFS
	Analysis of the video feedback→	TFS
Human–robot physical interaction →	Human subject body tracking →	TFS
	Control interface →	TFS, Teleoperation
	Human control of the test stand →	TFS
	Testing procedure →	TFS
Human–robot cognitive interaction →	Control interface →	TFS, Teleoperation
	Testing procedure →	TFS
Human–robot interfacing →	Agent structure →	TFS, Teleoperation
	Dual vision feedback mechanism →	TFS
Framework development →	Human subject body tracking →	TFS
	Network distributed data processing and communication →	TFS, Teleoperation
	Agent structure →	TFS, Teleoperation
	Control interface →	TFS, Teleoperation
	State machine analysis →	TFS
	Dual vision feedback mechanism →	TFS
Human and hardware safety →	Safety systems →	TFS

Visual servoing agents and strategies

The use of camera as a feedback sensor for motion control brings several challenges. First, the data the camera outputs are not compatible (units in pixels) with the rest of the system (output in volts). It requires processing and reasoning. Additionally, the coordinate frame is dissimilar and the data acquired have a noise of different nature than resolvers or encoders.

Human subject body tracking

The problem of tracking was divided into two independent tasks. One of them is tracking the body of a walking person with respect to the mobile platform. It includes the choice of a motion sensor and tracking techniques.

Network distributed data processing and communication

The systems described in this work are distributed data systems with data gathered, processed and shared at many places and aggregated at different locations. Similarly, the interface is distributed, and its changes affects different processing units. It includes tasks executed on the local or remote machines connected together through TCP/IP network. All of that results in serious challenges in design and processing.

Locating the object of interest in the image

Visual servoing is based on finding a set of visual cues in the image and obtaining their location in the image coordinate frame. Furthermore, for a class of bio-robots, tracking human joints has to be done in real-time and consistently.

Safety systems

Direct motion interaction between human and a robot requires high level of safety considerations. A robot of such type has to be equipped with passive and active systems of safety. In addition to that, proper procedures have to be developed.

Analysis of the gait and acquisition speed for tracking

In order to actively track and record human joint motion the following constraints have to be met: acquisition speed must be adequate to minimize motion blurring, the object of interest has to be in the field of view all the time, the response of the system has to be fast enough to account for joint motion, and all system delays and changes in motion direction must be considered. Additionally, the ability of the system to track is dependent on the maximum gait speed and not the average value.

Agent structure

Human-robot interaction and motion coordination requires many tasks to be accomplished at the same time. This applies to physical or computer system tasks on the same machine environment or different location/computer system. These tasks are executed by several independent agents that communicate with each other and the user, and release the resources when their actions are not needed.

Control interface

Effective and easy to use control interface is important for development of robots. In the case of a bio-robot interacting with humans it is even more important that the operator/supervisor has a real-time report of the system. Additionally, the operator needs tools to perform task specific tune-ups or manual adjustments as well as executing start up steps.

State machine analysis

The complex nature of human-machine interactions combined with remote control and decentralized task execution requires proper planning and development of state machines. This concept is well known in the robotic field but the details are drastically different between different types of robots and applications.

Dual vision feedback mechanism

In order to be able to test and develop a tracking robot using fluoroscopy as the main

sensing source and minimizing the radiation dose at the same time, a second video feedback system was used. Its role, besides testing the development of the robot, was to provide the possibility of providing zero radiation exposure and to train and familiarize the patient with the robot. The premise was to be able to switch between those two modes without any major shutdown delays or hardware/software changes.

Human control of the test stand

TFS is a bio-robot of a new kind and the first built prototype, hence it requires careful operation and planning. The minimum number of people required to turn the robot on is one however, when interacting with living subjects, it was established that it requires two or three people to operate it.

X-ray video display and recording

The images acquired during a fluoroscopy session on the Main Control Unit (MCU) needed to be displayed in real-time and recorded on the Remote Host Unit. The operator has to see the scene in order to start the procedure and respond in a case of unexpected events.

Testing procedure

Performing robotic motion in close proximity to human moving subjects requires strict protocol and procedures. In addition to mechanical interaction, the subjects undergo x-ray radiation. It was very important to plan all the actions and development steps in such a way that the risk of human injury or hardware damage is minimized while development and experiment goals are achieved.

Analysis of the video feedback

In order to provide good quality images and minimize radiation, a tool for offline analyzing and improving tracking methods is necessary. Each joint being imaged is slightly different than others due to age, size, gender, tissue characteristics, presence of implants or surgical operations.

Post analysis of the images.

Real-time acquisition and storage of the fluoroscopy images is one of the key issues in this research. For proper further analysis the images should be clear of any unnecessary graphical elements. However, for framework testing and development it is necessary to have the numerical information about the position of the object being tracked as well as visual cues.

D. Experimental Environment and Constraints

Hardware and Software Overview

Tracking Fluoroscope System

The joint tracking platform was built based on the experience gained from the first TFS prototype [43, 44]. The onboard National Instruments computer allows controlling 8 axes of motion, four of which are being used to control the wheels and steering. The remaining 4 were be used for joint tracking. The onboard computers were equipped with several I/O cards as well as image acquisition devices. They were be used as a motion controller and user interface computer. The user interface was designed to work remotely. LabView Real Time was chosen as an operation system for the platform. All the software was written in LabView code. It was equipped with a custom TCP/IP based interface in order to communicate with the operator's station and a vision computer remotely.

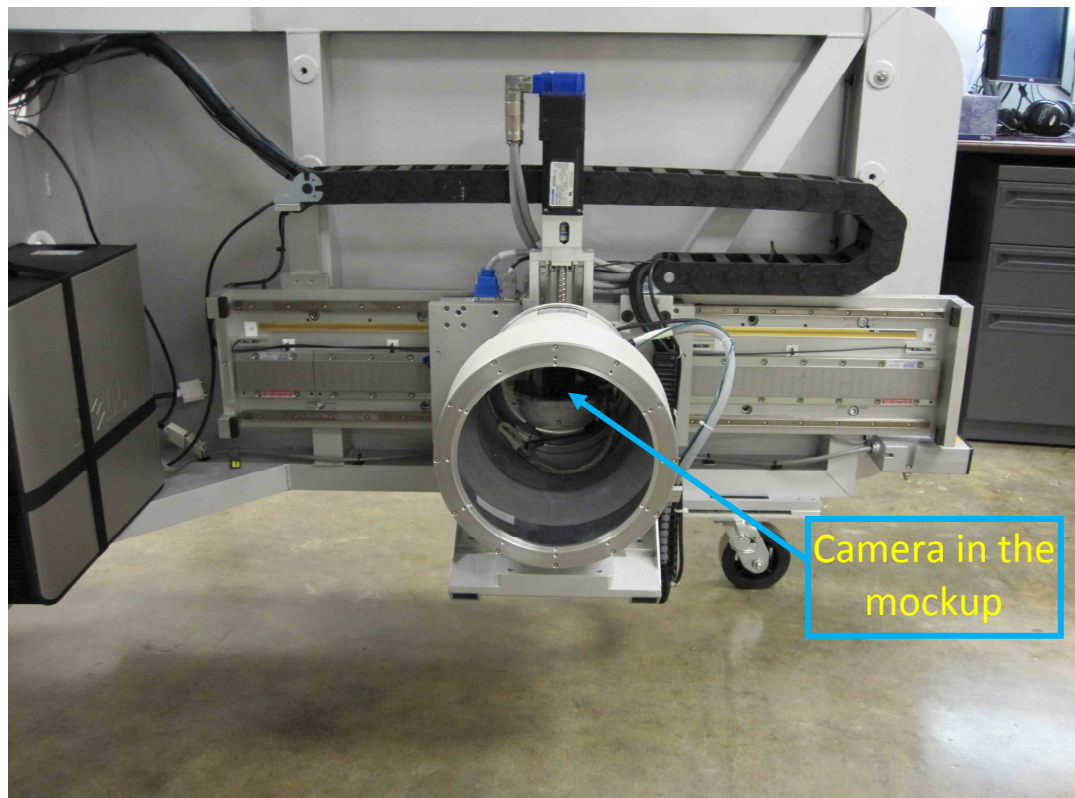


Fig. 9. X-ray intensifier mockup with a camera mounted.

The vision computer is a separate machine also running RT LabView OS and the code was developed be in LabView language. It was used to process the laser data, interface the camera mounted on the X-ray receiver mockup (Fig. 9), process the vision data and send the results to the motion controller. The camera is a digital CCD Fire wire camera capable of acquiring images at 60 fps at 656x491 pixels resolution. Furthermore, the real time video of the motion will be recorder and wirelessly sent to the operator's station.

Teleoperation

The robotics lab In MABE department was equipped with two Titan II arms (Fig. 10). These are six DOF arm with a maximum reach of about 2m when fully extended and a maximum load of 250lbs. Joints are equipped with 14 bit encoders, which allows it to complete majority of teleoperation tasks that do not require extreme precision. The arm is controlled by a PC104 controller that accepts joint commands and runs a control loop at 200Hz. The controller was be interfaced through a TCP/IP protocol and that integrated the Whole Arm Manipulator (WAM) as a master controller [52].



Fig. 10. Dual arm Teleoperation System with a mockup.

Tracking Fluoroscope Limitations

The Tracking Fluoroscope System operates in close proximity to human subjects. Significant measures were taken in providing safety from mechanical or radiation injury. That included hardware, software and training solutions. The following list presents the major functions and requirements.

- Varying walking speeds and styles
- Varying subject heights
- Continuous walk capability
- Mechanical adjustability to different subject heights
- Wireless and tetherless

General operator's and development constraints for the teleoperation

Several aspects of the interaction between the agent and the operation were considered in the research. The major assumptions for the assisted teleoperation tasks include the following actions that will have to be performed by the operator or the agent:

- Showing the object
 - Operator points out the object of interest and make corrections if necessary
 - The vision agent informs the operator about:
 - Level of certainty for the executed tasks
 - Impossibility of the planned task
 - When it needs corrections
 - When the task is done or the task is not progressing.
- The vision agent software was programmed using the native LabView code and its vision tools. If necessary, custom functions were created in LabView or C language.

- The operator's station is able to access the motion controller interface, receive, display and store the image data from the robotic platform.

E. Concept Execution

Visual servoing agents and strategies – Tracking Fluoroscope System

The visual servoing agent in the TFS is tracking the joint of interest. Its role is to find the best estimation of the desired motion output based on the input given by the image processing agent.

The following variables are the input for the agent:

- Position of the four motors
- Current video feedback in use
- Position of the joint in the image frame (2D)
- Time until the next update of the joint position and since last one.

Based on this information the agent controls the desired position of the robot. The motors are controlled by the servodrive in a position loop mode. The motion agent provides the absolute desired position and the servodrive closes the position loop internally.

The agent is running at the rate of 333Hz and in every cycle has to read four encoders and calculate two outputs: horizontal and vertical (Fig. 11).

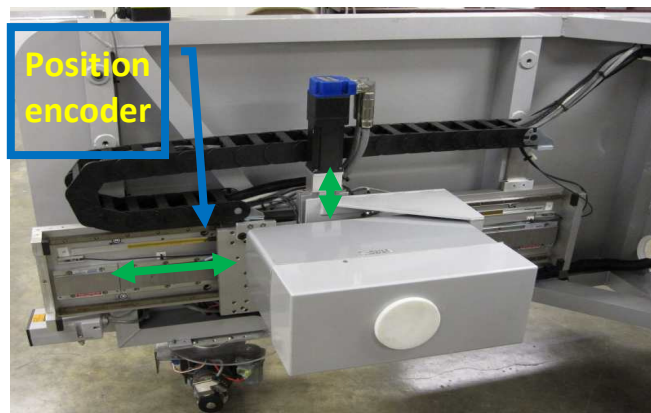


Fig. 11. -Linear drives with the encoder shown.

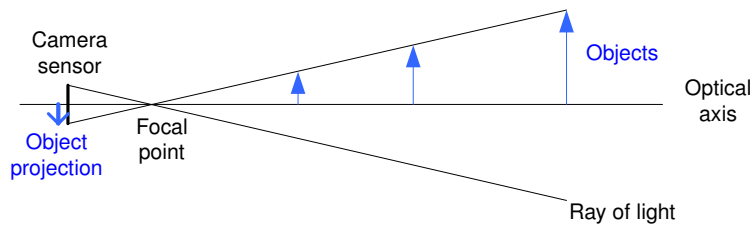


Fig. 12. Relation of image to real world objects

Image to real world coordinates

One of the first issues to solve is to translate the position of the object of interest in the image in pixels unit into real word coordinates (robot frame coordinates). Since the object is not planar, it does not stay at a constant orientation with respect to the camera, its distance to the camera is not constant and its projection is not ideally linear, there are no sensors in the system that can measure the distance, a precise translation from image coordinates is very difficult to achieve. The described situation is presented in Fig. 12 and Fig 13. Considering all of

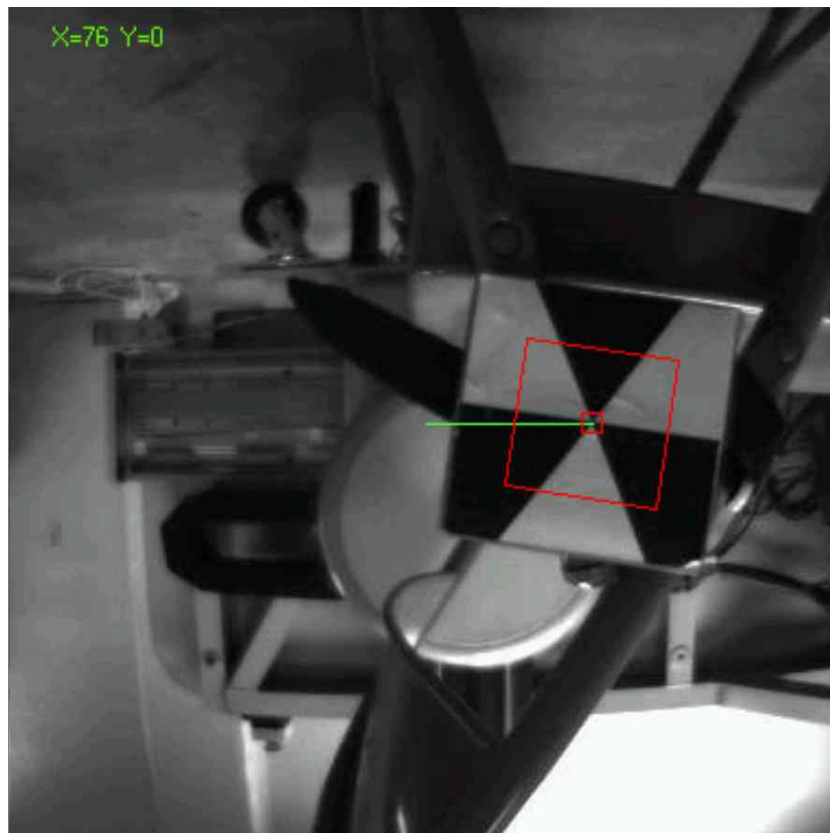


Fig. 13. Projection of the object of interest. Distance to the object difficult to calculate.

that, the following measurements were conducted: length of the object of known size at the closest and farthest location where the real object can be found. The goal of the measurement was to obtain the following factor:

$$\text{Pixel width/real width} = \# \text{pix/mm. (pmm)}$$

For the camera feedback the minimum scaling factors were found to be:

$$Pmm \text{ min} = 0.59 \text{pix/mm for objects closest to the camera}$$

$$Pmm \text{ max} = 1.11 \text{pix/mm for objects furthest from the camera}$$

Since the system does not know the distance to the object, it can use only one value. If it is too big, the video servoing creates a control signal overshoots the object and causes oscillations. If it is too small, the tracking performance is reduced. Since the amount of motion in a joint is not the same for horizontal and vertical movements, the scaling factors can have different values. The following absolute values were found experimentally for the camera feedback system:

$$Pmmx : 0.55 \text{pix/mm}$$

$$Pmmy : 0.7 \text{pix/mm}$$

In a similar way the factors were obtained for the X-ray system:

$$Pmmx : 0.34 \text{pix/mm}$$

$$Pmmy : 0.27 \text{pix/mm}$$

These factors should also hold a proper sign for the direction of motion required to move the objects.

Vision tracking feedback adaptive gains

The image position error can be anywhere from 0 to almost 100 pixels in the current system. A small value suggests the motion happening is of low velocity whereas high values

suggest high object velocity. It comes from the fact that the systems acquires feedback at constant rates hence, even if it moves at a very high speed and brings the last position error to zero, by the time it acquires the next reading the next error will most likely be of similar value. For the case of walking it also can be inferred that small error can be a sign of a different phase of walking as knees move forward and backward toward the trunk. Hence it was decided to apply additional adaptive gains to prevent the system from oscillating at low position error values and fight the error position more aggressively when the values are high. The following two formulas were experimentally developed.

$x^2 \cdot 0.012 + 0.439 \cdot x$ for the X-ray case

$0.012 \cdot x^2 + 0.472 \cdot x$ - camera case. x image error in pixels.

A graphical representation of this equation is presented in Fig. 14.

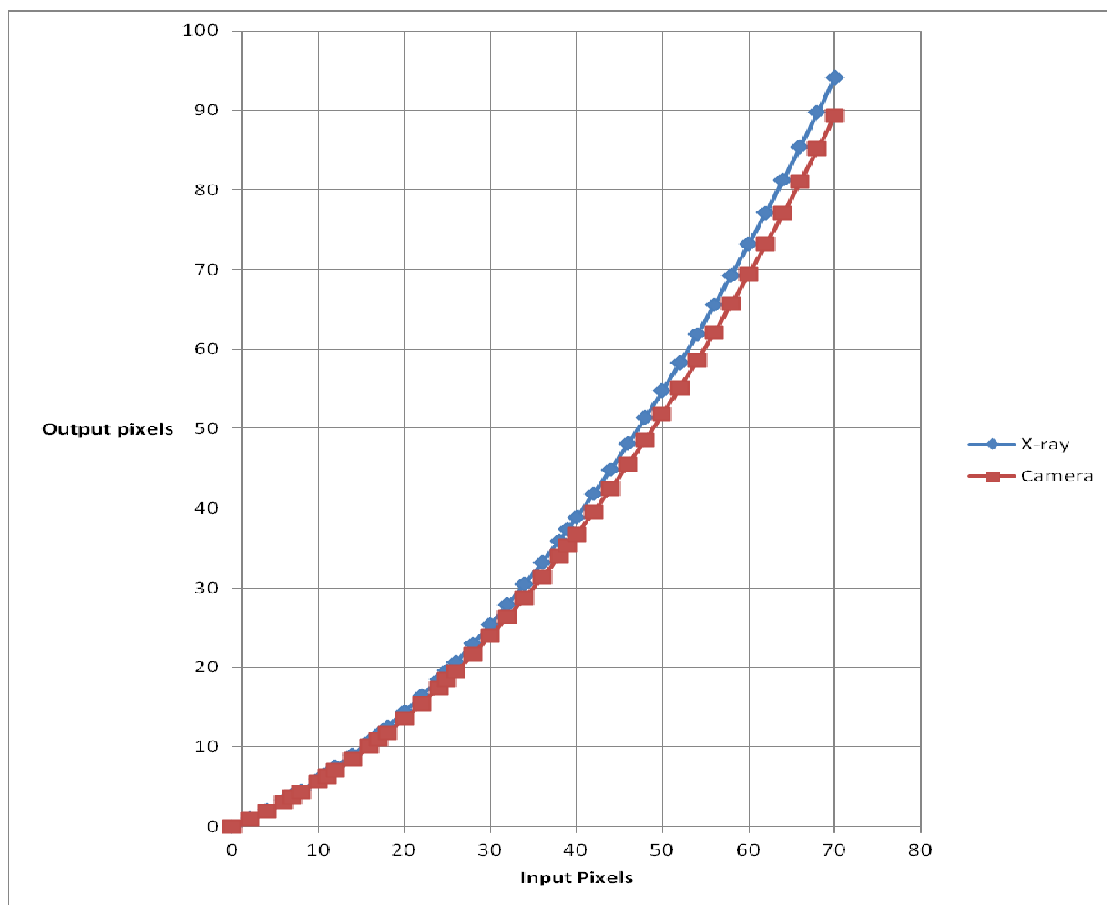


Fig. 14. Relation of measured error and error used for control signal.

Asynchronous mode of execution

The joint tracking agent runs at 333Hz while the video processing agent can run at a varying rate from 30Hz to over 60Hz. The high rate of the motion controller is dictated by the need of interpolating the motion position in between updates. The more interpolated points, the smoother the motion is. Making the motion smooth can affect the performance of tracking. Considering those factors, a custom method of controlling the motion was developed.

The agent receives image position updates at 100Hz + rates, which is faster than the video processing rate. In order to distinguish old values from new ones an extra variable was added to the communication packet. Every time the agent receives an update it knows if it is the most recent update. For example, for every six updates, only one is the most recent. The problem gets more complicated (Fig. 15) by the fact that a recent update can happen in the 60/333Hz (video to motion rate) scenario either per 6 or 5 cycles ($333/60$). The agent must properly address this.

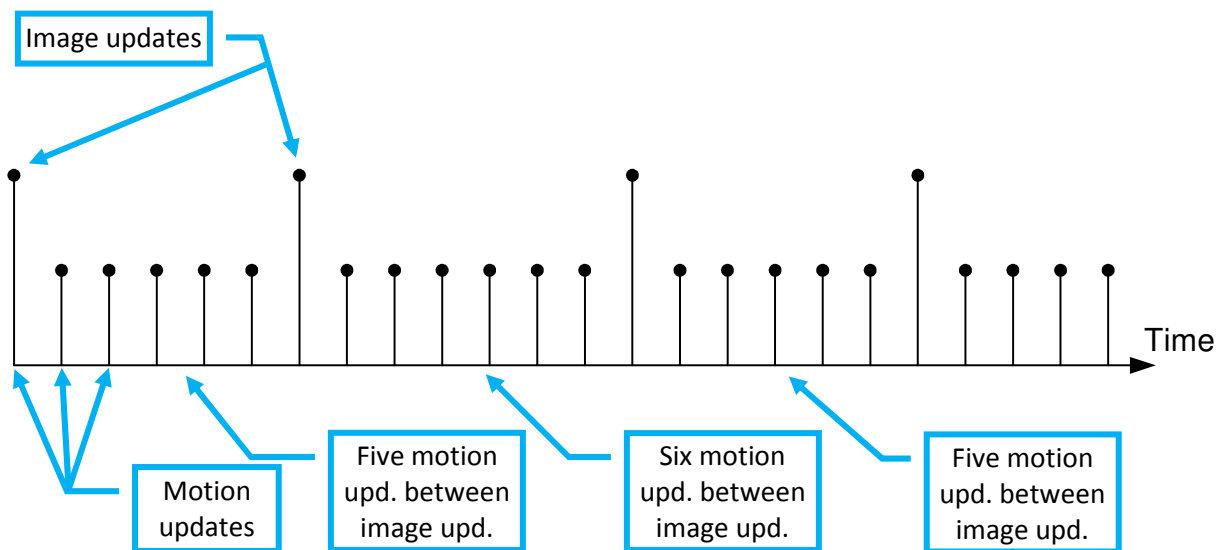


Fig. 15. Image and motion update on the timeline.

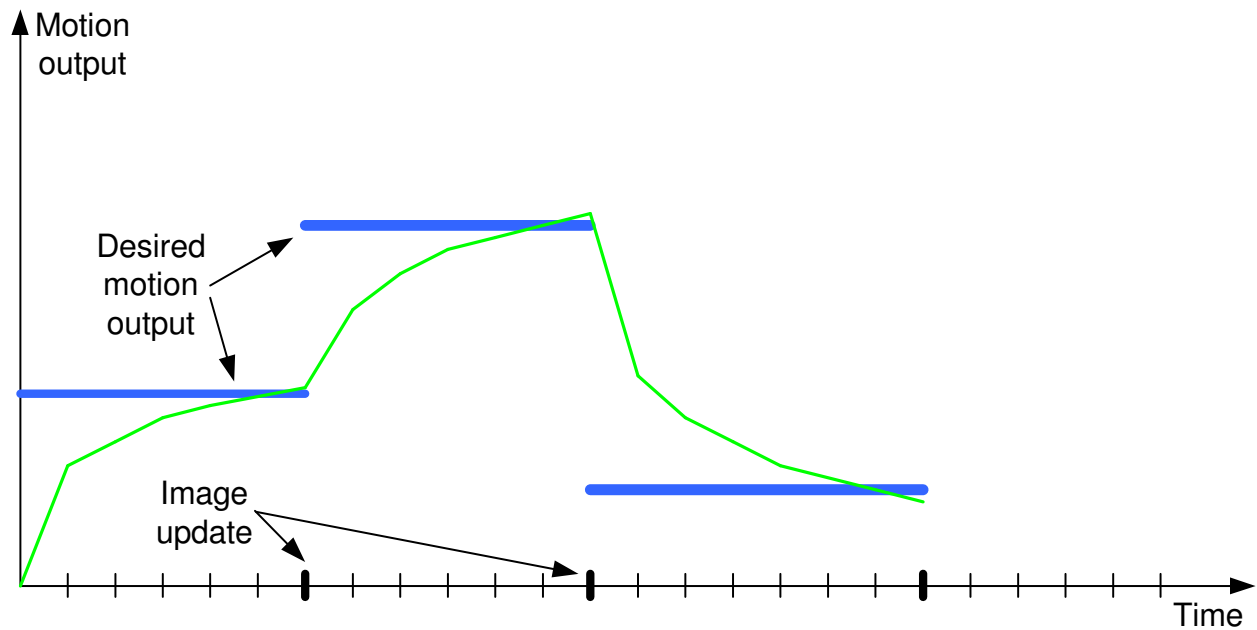


Fig. 16. Motion output ramping.

The agent has to act differently when it is in an update cycle or a regular cycle. The following approach was experimentally developed. During an update the agent sets its control output to about 60% of its desired value and in the following cycles it ramps it up to the full value. An example of that approach is presented in the table below and in Fig 16.

Table 2. Motion output ramping.

Motion cycle number	Motion output %	Remarks
1	60	New value from the video agent
2	70	First value after update
3	80	
4	90	
5	99	Expected last value before a new update

Motion cycle number	Motion output %	Remarks
6	105	Expected last value before a new update. This happens every second update.
7	107	Exception value for unexpected slowdown in the transmission or the operating system
1	60	New value from the video agent
2	70	First value after update
...	...	...

Under normal conditions the agent may receive a new update every five or six cycles but in order to account for an unexpected situation an extra cycle was devoted to create a motion control output. Before a new vision update reaches the motion agent it has to go through several hardware and software layers and each one of them possibly can create an unexpected slowdown. The ramping map like in the Table 2 was created separately for the X-ray and camera mode due to different dynamics, field of view and motion stability.

Position units

In the motion agent, the position of the camera and the object of interest is presented in volts of output. The whole runway length both in vertical and horizontal directions was scaled to be represented from -10V to 10V, which is also the output of the controller. 1V represents exactly 32mm of motion for the horizontal drive and 24.13mm for vertical drive. The agent internally and in the interface is using volts as position units.

Preventing excessive motion

The agent checks during every cycle if the desired motion is within physical limits of the

drive and also checks for a single a maximum desired change per single update. The system currently allows for no more than 2V of motion in any direction in a single cycle (64mm or 48mm per cycle).

An example of full motion sequence

Initial value form the image agent: 37pixels

Mode: camera.

Current position 5V.

Cycle 1.

$0.012 * 37 * 37 + 0.472 * 37 = 34 \text{ pixels}$. Desired motion in pixels. Adaptive gains application.

$34 * 0.34 = 11.56 \text{ mm}$. Desired motion in mm. Change of coordinates (pixels to mm).

$11.56 / 32 = 0.36 \text{ V}$. Desired motion in volts. Change of units.

$\text{Min}(0.3; 2\text{V}) = 0.36$ Desired motion . Checking for increment limits

$\text{Min}(0.36 \text{ V} + 5\text{V}; 10\text{V}) = 5.36 \text{ V}$. Desired final output. Checking for motion limits.

$65\% \text{ of } 0.36 = 0.23 \text{ V}$. $0.23 \text{ V} + 5 \text{ V} = 5.23 \text{ V}$. Motion desired output in the first cycle.

Ramping up the signal from 5V to 5.36.

Cycle 2.

$\text{Output} = 0.7 * 0.36 + 5\text{V} = 5.25 \text{ V}$. Second ramp value.

Cycle 3.

$\text{Output} = 0.8 * 0.36 + 5\text{V} = 5.29 \text{ V}$. Third ramp value.

Cycle 6. Last cycle.

$\text{Output} = 1.05 * 0.36 + 5\text{V} = 5.36 \text{ V}$. Motion output reaches its final value.

Modes of operation

The agent has four different modes: Idle, manual, auto and align. Their detailed descriptions are described in the interface section. In all these modes the agent checks for position limits and restricts the output if necessary.

The align mode is a transition mode between the manual or idle mode and auto-mode. It smoothly brings the position error to zero before the actual motion of the object of interest starts. Every time it checks if the error in pixels is below 10 pixels and if that is not true it allows for 1% of the desired change in position per cycle. When a new update from the image agent is received, the motion agent checks the threshold again and if it dropped below 10 pixels it switches to full speed auto mode of operation. An example is presented below:

Cycle= 1, Image update Initial error = 35pixels

Error= 35pixels, 35pixels >10 pix. Align mode is active

0.01 x 35 =0.35pix. Allowed motion in a cycle (1% per cycle)

Cycle 2

0.02 x 35pixels =0.7 pixels. Allowed motion after cycle 2 (cumulative 2%).

Cycle 6

0.06 x 35pixels =2.1 pixels. Allowed motion: 2.1 pix

Cycle 1, Image update

Error=32.9 pixels > 10 pixels.

0.01 x 32.9 =0.329 pixels. Allowed motion in a cycle (1% per cycle)

Cycle n , after k image updates

Error $n < 10$ pix, mode switched to full automatic.

Considering the rate of the motion at 333 cycles per second, the alignment process is executed within a few seconds. The manual mode is based on the input from the operator. A slider control is present on the screen and its position represents the position of the motor. The change of position done by the operator is filtered using a mean value to prevent abrupt changes. The interface part of the motion agent is presented in Fig. 17.

Visual servoing agents and strategies – Teleoperation System

Vision servoing

The visual servoing agent is significantly different from the fluoroscope system. It receives the following data on its input:

- Image position of the object of interest.
- Current arm position and orientation.
- Location of the camera with respect to the arm.

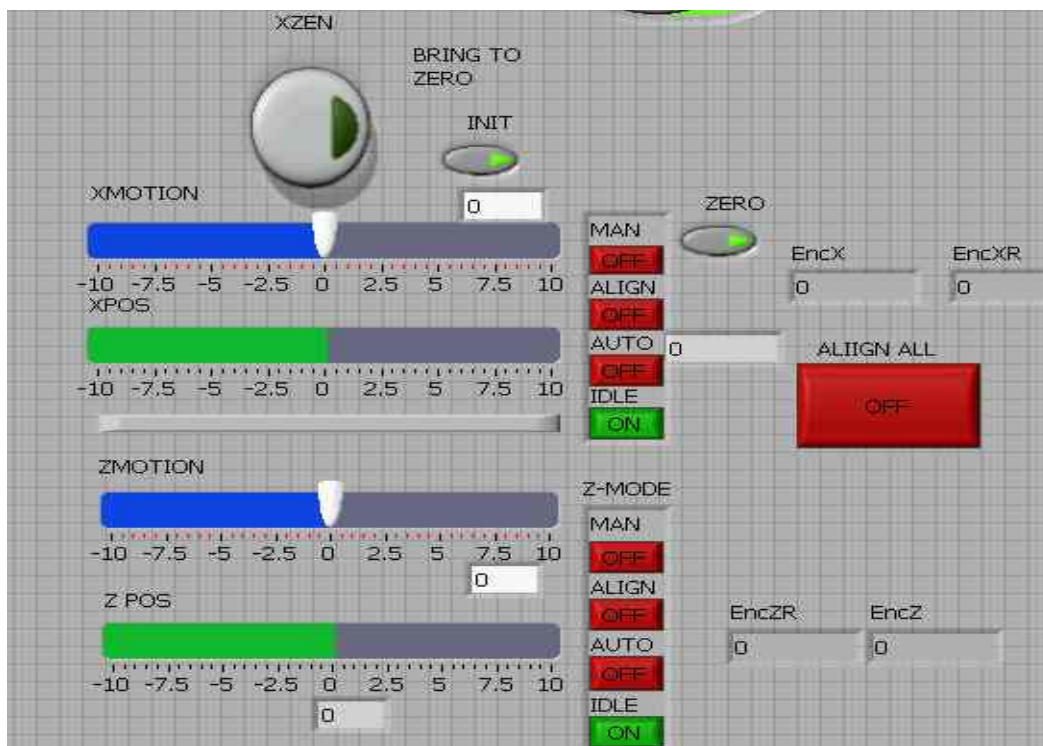


Fig. 17. Motion agent interface.

Based on these data the agent will calculate the direction of motion of the arm. The final joint change calculation is done by the motion control agent of HLC.

The position of the object of interest is not directly translated into the arm position units. Instead, knowing the orientation of the camera relative to the arm, the image position is used to calculate the direction of motion but not necessarily the rate of motion or absolute distance. The decision to stop the motion comes from the desired image feature location. If the requested requirement is met, the motion stops. In the TFS situation the approximate distance to the target was estimated, while in this work it relies mainly on the video feedback error direction.

High level Controller development

The vision servoing agent produces Cartesian control data; hence a High Level Controller is necessary to properly translate Cartesian control into joint control. That was also achieved in [52]. However, the HLC needed a major upgrade to accept a control signal from other sources other than WAM. The HLC also has to deal with non synchronized input output control signals

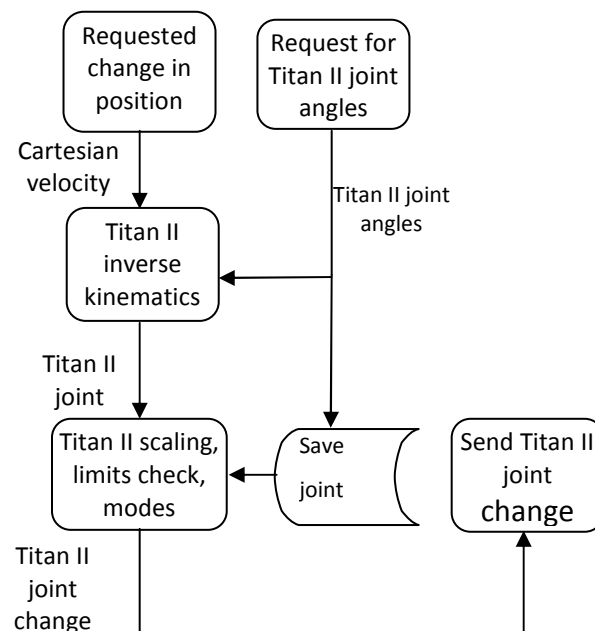


Fig. 18. HLC data processing.

and incompatibility in the form of the control vector between different input devices. It has two kinds of outputs: PC104 and Roboworks simulation with two modes. In the normal real mode PC104 gets the control signal and the Roboworks simulation software servers as an arm visualization platform. In the second purely simulation mode, the operator has a possibility to test his/her approach in a simulation without the real arm being involved. HLC addressed the different frequencies of those two outputs in a proper manner. The HLC is run under a real-time clock and coded in C language. The system diagram is presented in Fig. 18.

In order to directly control the servo valves of the Titan manipulator a Low Level controller was necessary. The Low Level Controller was initially developed by Oak Ridge National Laboratory for Schilling® manipulator control. It originally used serial communication. The communication server was changed to communicate with the Titan High Level Controller through TCP/IP in both directions (Fig. 19). The main controller was also changed to accept different type of input and close the control loop with that input.

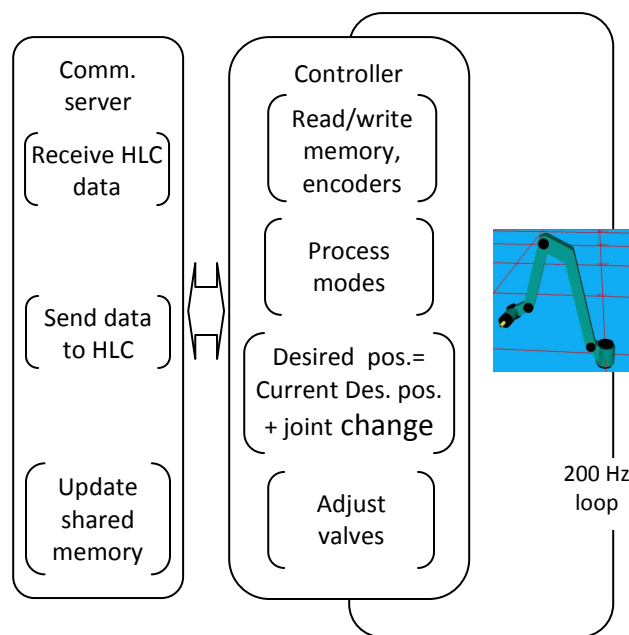


Fig. 19. Titan II Low Level Controller.

Master control

In order to provide direct manual control over the Titan arm the Whole Arm Manipulator was used. It is a new type unique controller. The master and slave arm used in this research are two kinematically dissimilar manipulators. Due to this fact, it is not possible to use joint to joint mapping. Moreover, the master arm is a redundant manipulator, which results in greater dexterity but makes the control more complex. Hence, the High Level Controller uses velocity control in the Cartesian space. This allows mapping from 7 DOF joint space of the WAM into the 6 DOF joint space of the Titan. This approach involves also problems of singular configurations.

Since, the WAM controller is a real time application, a special software setup is needed in order to extract the joint data. Its main loop runs at 500Hz and must have high task priority. Hence, for communication purposes there is another process running at lower priority and rate. The controller every 0.033 s sends the current joint data through a FIFO (*First In, First Out*) pipe mechanism to the communication process, which forms a proper data packet and sends the packet immediately over the network to the Titan High Level Controller. It is shown in Fig. 20.

The High Level Controller receives the data from WAM and asks for a current position of the master arm. Next, it performs the WAM forward kinematics and Titan inverse kinematics.

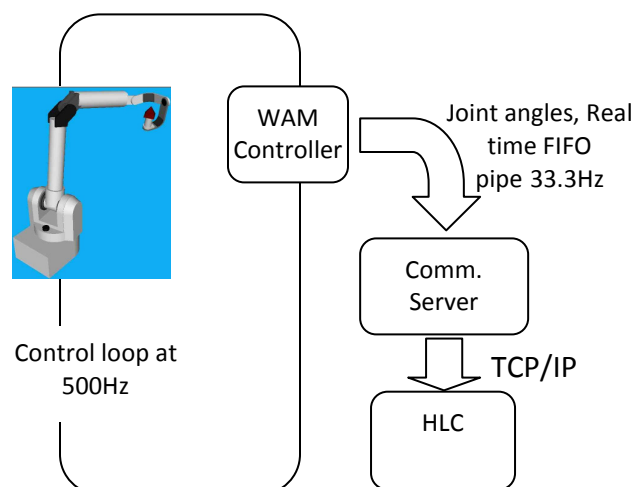


Fig. 20. WAM communication diagram.

As a last step it combines the current Titan position and its limits and sends the data over to the Low level Controller. The mathematical engine is presented below [53, 54].

First, the joint velocity of the WAM is calculated. Then, Cartesian velocity is obtained. In the next step, using a pseudoinverse, joint velocity of the slave arm is calculated and the joint position increments are found

These calculations must be accomplished before the next portion of the data arrives. The algorithm uses singular value decomposition for the Jacobian inverse. In order to find a stable solution, a pseudo inverse is used. Therefore, a threshold for the singular value must be found. Small values decrease system tracking, whereas high values allows for an unstable solution. A suitable threshold must be determined experimentally.

At the same time, the controller must handle its own graphical user interface, keyboard input, communication with simulation workstation and main GUI in the touch screen computer. The controller can work in two modes. In the first one, the operator can see the results of its actions only in the simulation; in the second one the operator controls the real arm and the current state of the arm is shown in the simulation.

$$v_{WAM} = \Delta q_{WAM} / \Delta t$$

$$V_{Car} = J_{WAM} v_{WAM}$$

$$v_{Titan} = J_{Titan}^+ V_{Car}$$

$$\Delta q_{Titan} = v_{Titan} \Delta t$$

v_{WAM} – WAM joint velocity, Δq_{WAM} – WAM

joint angle increase, Δt – time period,

V_{Car} – Cartesian velocity, J_{WAM} – WAM

Jacobian, v_{Titan} – Titan joint velocity, J_{Titan}^+ – Titan

Jacobian pseudoinverse, Δq_{Titan} – Titan

joint angle increase

This helps the operator to get used to interface and exercise his/her actions when needed. The interface is transparent because the operator uses exactly the same technique to control the arm in both modes.

Based on tests and inspirations from industrial arms it was found that certain additional controlling features are useful when working in Cartesian space. These include the ability to disable any of the six linear or rotational axes of movement, resulting in constraining the motion to a line plane or rotation. The relative useful Cartesian range of motion for the operator sitting in the chair and operating WAM is about 1-2 feet whereas for Titan it is about twice this size. Since the workspace of WAM is about the half of the Titan II, an additional option was introduced that allows scaling the velocity such that the operator does not have to do frequent indexing, e.g. making movement of one foot forward, idling, moving back the master, activating and making second movement one foot long.

Additionally, in order to facilitate the use of the system, certain predefined modes were introduced. They include: linear motions (rotations of the end-effector blocked), rotational motions (displacement of the end-effector blocked) and joint to joint control with one joint controlled at a time.

The option to disable certain axes showed up as very useful in operations like cutting, and dismantling where the precision was needed. In such a case, even when the master arm is pointed in the wrong direction, only the intended component of that motion was used to drive the slave arm.

The High level controller was developed to accept four kinds of control inputs: WAM master arm, internal agent (i.e. behavioral agent), vision agent, and graphical interface. The functionality of the HLC is similar for all the input control types.

Network distributed data processing and communication – TFS and teleoperation

The systems described in this work are located on different processing units. They are connected in local networks, both wired and wireless. The tasks on the same machine are executed very often in parallel on several cores of the processor.

Networking

The network structure of TFS is presented in Fig 21. In general, the computer system has to provide three major functions: sensing, motion control and user control. Since it is a wireless, mobile system the user control is performed at the Remote Control Unit. Due to the complexity and variety of jobs, the sensing and motion control parts were split. One process (IAPU) is responsible for acquiring the vision and laser range data, processing it and sending it to the MCU. The MCU controls all the motion systems.

TCP/IP was chosen to be the protocol of communication. The main reasons were: lossless data transfer and performance. In order to avoid unnecessary communication breakdowns, both MCU and IAPU were set on a wired network. Additionally, they run real-time

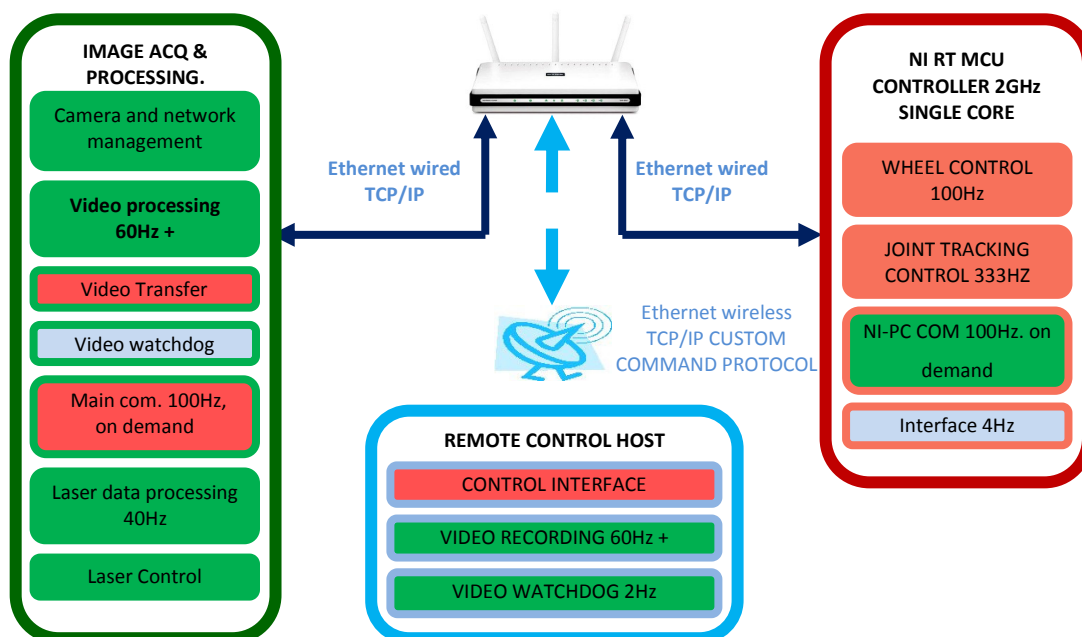


Fig. 21. Network topology.

operating systems for improved performance. Special care was taken to make sure that network communication coming from different operating system is compatible with each other. This comes from the fact that the formatting at which data travels is operating system dependent.

Each system communicates with the others using a custom command and data protocol. The IAPU sends the calculated positions of the objects and images. It receives control commands from the MCU. The MCU receives the processed information and user control; sends system status and controls the behavior of IAPU. The RCH receives the images and status of the system and sends high level control to the MCU.

Control systems are dependent on fast feedback and communication. In order to meet this requirement, the Nagle's algorithm [55] for data packaging was turned off on all channels. This algorithm saves bandwidth at the cost of speed, binding small portions of data together and sending them as one. Data bandwidth in this work was not saturated and it was crucial to avoid any additional delays.

The topology of the Teleoperation system is presented in Fig. 22. As in the case of TFS, the tasks are executed on several processing units. The heart of the system is the High Level Controller performing all kinematic calculations and providing control interface. Two other systems were hardware low level controllers: WAM and Titan Controllers. Also, like in the TFS situation the sensing system was split from the motion control system (HLC) and low level controllers. The sensing system (Vision Control Unit) receives image data, processes it and sends the acquired information to the HLC and the images to the Smart Interface. The Smart Interface is a system interfacing with the operator and showing the status of the system.

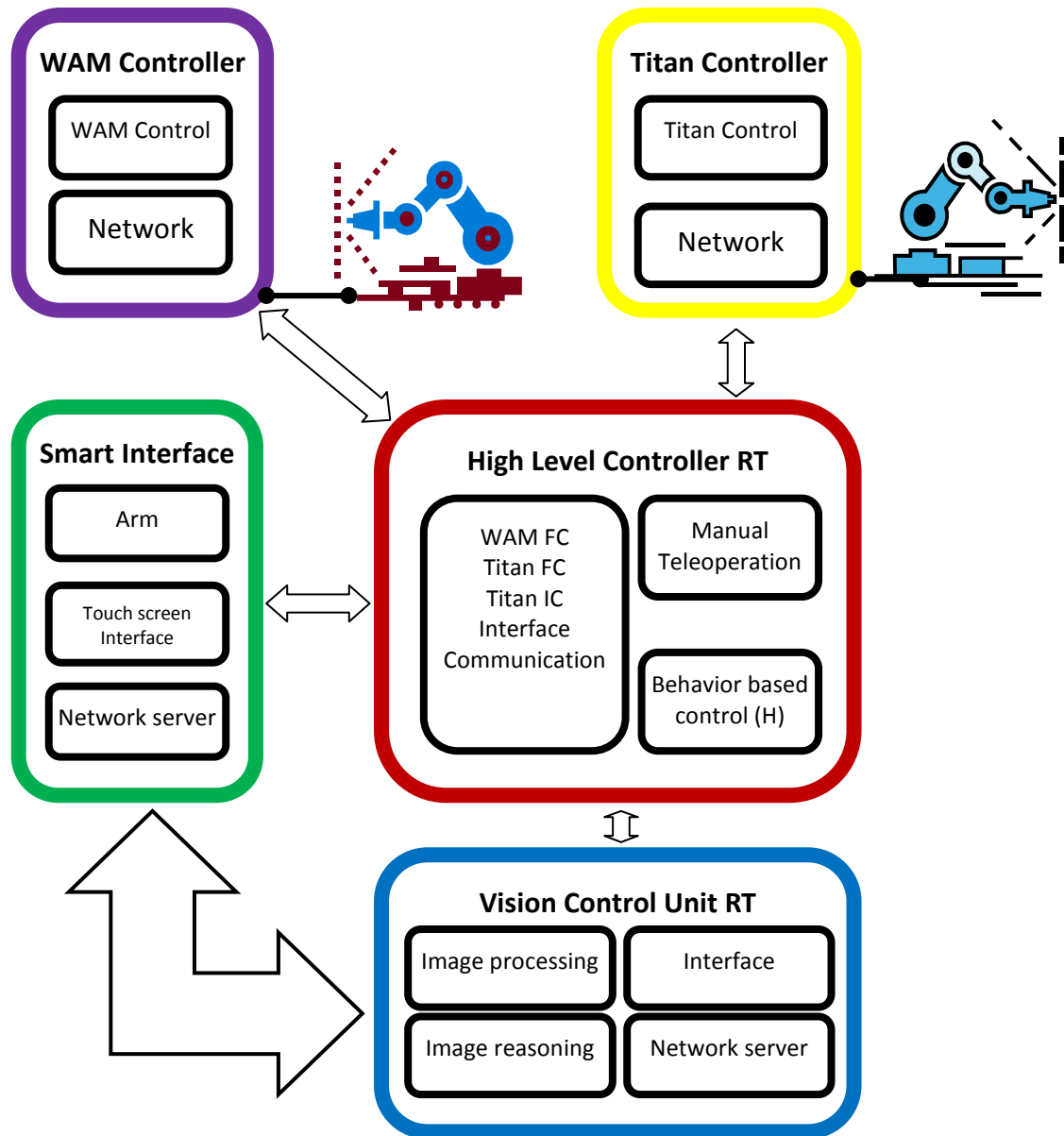


Fig. 22. Teleoperation system topology.

Comparing the teleoperation and bio-tracking system, the similarity in structure is obvious: Smart Interface-Remote Control Host, VSU-IAPU and MCU with HLC and arm controllers.

TCP/IP was also used connect all the systems with the same emphasis for performance (Nagle's algorithm) and reliability.

The dissimilarity between operating system was stronger in this case. Several operating systems were involved: two different distributions of Linux, one UNIX based system (PC104), one Windows system (SI) and one LabView RT system. Proper actions were taken to ensure compatible data exchange.

Inter Process Communication (IPC)

In order to improve performance and take advantage of the multi cores technology many tasks were design to execute in parallel. Furthermore, motion control tasks have to have a higher priority level than other tasks. Hence, the systems need properly designed inter process communication. The following mechanisms were used:

- Global variables
- Shared memory
- FIFO pipes
- Semaphores
- Queues
- Notifiers
- Action Engines

Most of these mechanisms are well known in the IPC literature. The last three are native LabView features. The major difference between Notifiers and Queues is single and multi element data exchange.

An action engine is the most complicated and the most powerful inter process

communication tool. It can perform all the functions of other tools combined. Among them, the most important are:

- No race conditions
- Exclusive access
- Data storage
- Function storage
- Multi client access
- High performance

The main disadvantages are the complexity of design and availability on only LabView system. These tools were used to communicate the processes in MCU (TFS) and SI (teleoperation systems). A more detailed description of usage is presented in the *State machines* section.

Locating the object of interest in the image

Considering the type of images in this research, in order to locate a moving object in the image, it is beneficial if its features of interest meet the following criteria:

- Uniform illumination
- Visible and not occluded
- Unique in the image
- Not changing drastically in shape or size
- Its projection under normal conditions has to be relatively constant

- Having a well defined geometry for finding its location, and preferably symmetric

A uniform and stable illumination is a key in many machine vision applications. The changing light conditions can be very difficult to manage when designing a recognition algorithm. Many current techniques are very robust in dealing with this problem however, proper constant light makes the process easier and more reliable.

The second requirement is simple in its nature but very important. The object of interest has to be present in the image and be easy to identify. For example, in a case of a knee image taken with a regular camera and an X-ray machine, in both cases the leg is present in the image but only the X-ray image shows the actual knee clearly different from the rest. In a case of a moving object or moving camera, it is important to make sure that the object is not getting occluded by other moving parts.

The object has to be unique in the image. Presence of objects of similar shape and intensity can cause false matches and, in effect, failure of tracking. The uniqueness applies also to the background. It should be clearly distinct from the object. Additionally, if the object of interest is not present in the image, there should not be any objects that are similar to the target if under normal conditions they are blocked by it.

If the object drastically changes its size or shape, it makes the recognition process very difficult. It is much harder to create a model or a pattern match that does not have a well defined shape or size. The change of the size or shape of the image feature can be a result of change of location of the objects.

Even if the object is not changing its size or shape but due to motion its projection is changing, the recognition based on that projection can fail. Such situations can happen when the plane of motion is changing, i.e. the person is changing the direction of his/her walk.

An object that has a well defined geometry and is symmetric is also easier to find and

track. Due to motion (translation or rotation), the calculated location of the object can vary and cause oscillations in the control signal. Circles and squares are examples of object of well defined symmetry.

The mentioned criteria can be relaxed if additional measures are taken to prevent the faults. For example, if the object changes its size or shape in the image a second model can be created to account for the change. Another solution can be the use of image processing filters or complex cascade models. Thirdly, the process of finding the location of the object has many levels and if one level fails, the second method is used.

Tracking Fluoroscope System

Two distinct cases have been investigated thoroughly: artificially created visual cues and natural joint image without fiducial markers. In the first case the marker was attached the knee of the subject.

The first case is easier due to the fact that the artificially created cues can be designed to meet the criteria for tracking while the natural image does not give such flexibility. Two different patterns were created for tracking in the camera mode and X-ray mode. They are shown in Fig.23.

The characteristic features of them are: well defined geometry, symmetry and good contrast. In the camera mode the background is built of industrial components hence, the



Fig. 23. Artificial markers.

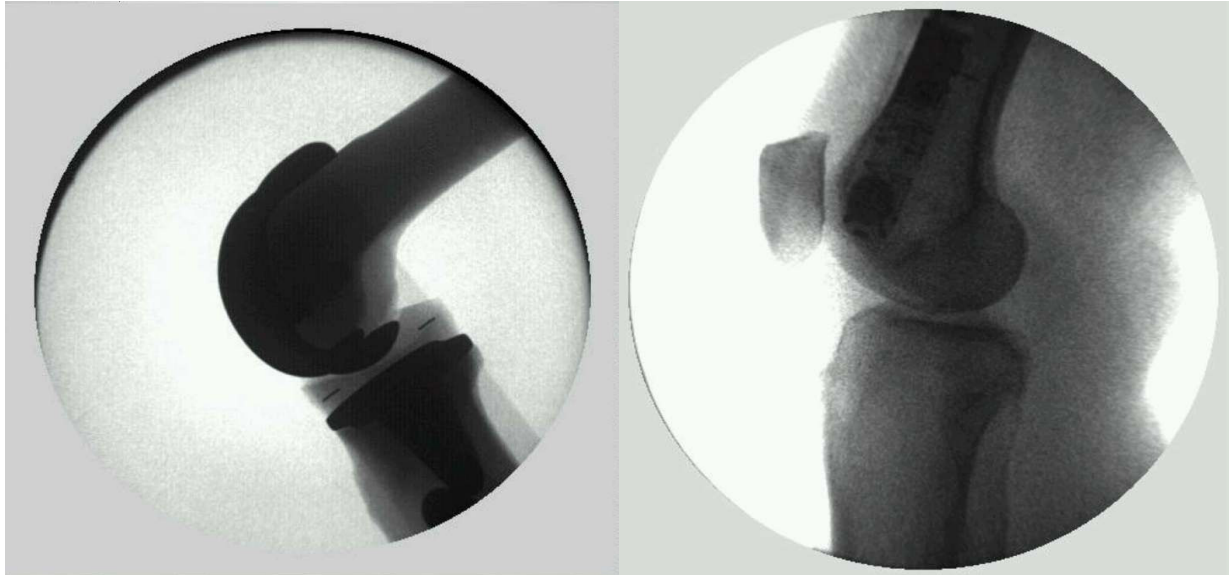


Fig. 24. X-ray image of artificial and natural knee.

pattern was designed to avoid 90 deg edges. This way the number of false positives was limited. The second pattern was built for testing the tracking in the X-ray mode.

A knee in motion can rotate significantly while changing distance from the camera. This, in effect, is changing the size and shape of the marker. Additionally, the background is in motion as well and can be a source of false positives. Hence, after several trials the marker was chosen to be scalable with black and white fields. Such a shape allows it to be found when it gets closer to the camera, since the center portion of the marker looks like the whole.

In the case of human joints it was not possible to create an artificial pattern. The model of the object of interest had to be created based on the available features in the image. As previously it was done in two cases: real and artificial human joint (Fig. 24).

Tracking

The feature tracking problem was investigated in two different cases (Fig. 23). The first case was a marker attached to a real walking person and the second was a manually operated artificial object created for testing purposes. The goal was to be able to find the dynamically moving object in every frame.



Fig. 25. Knee tracking with a marker.

The algorithm used to track the artificial marker in the image was pattern matching. A template of the marker was created and used to track the knee position in the image. Template matching is a good tool when the object of interest has variable intensity or rotation [12]. It is a smart, correlation based technique. A sequence of images with the marker found is presented in Fig. 25.

The case of the artificial knee was different. There is no strictly predefined marker or shape. The knee is a three dimensional object and its projection can vary (Fig. 26). Also, the way X-rays are going through the body of the knee can affect the image. Depending on the distance

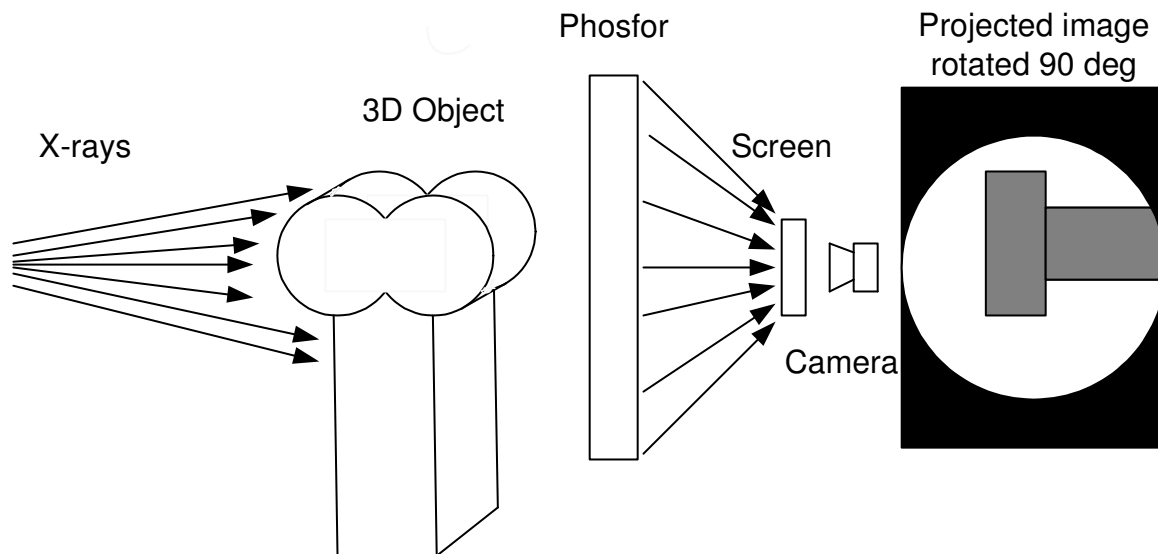


Fig. 26. X-ray projection diagram.

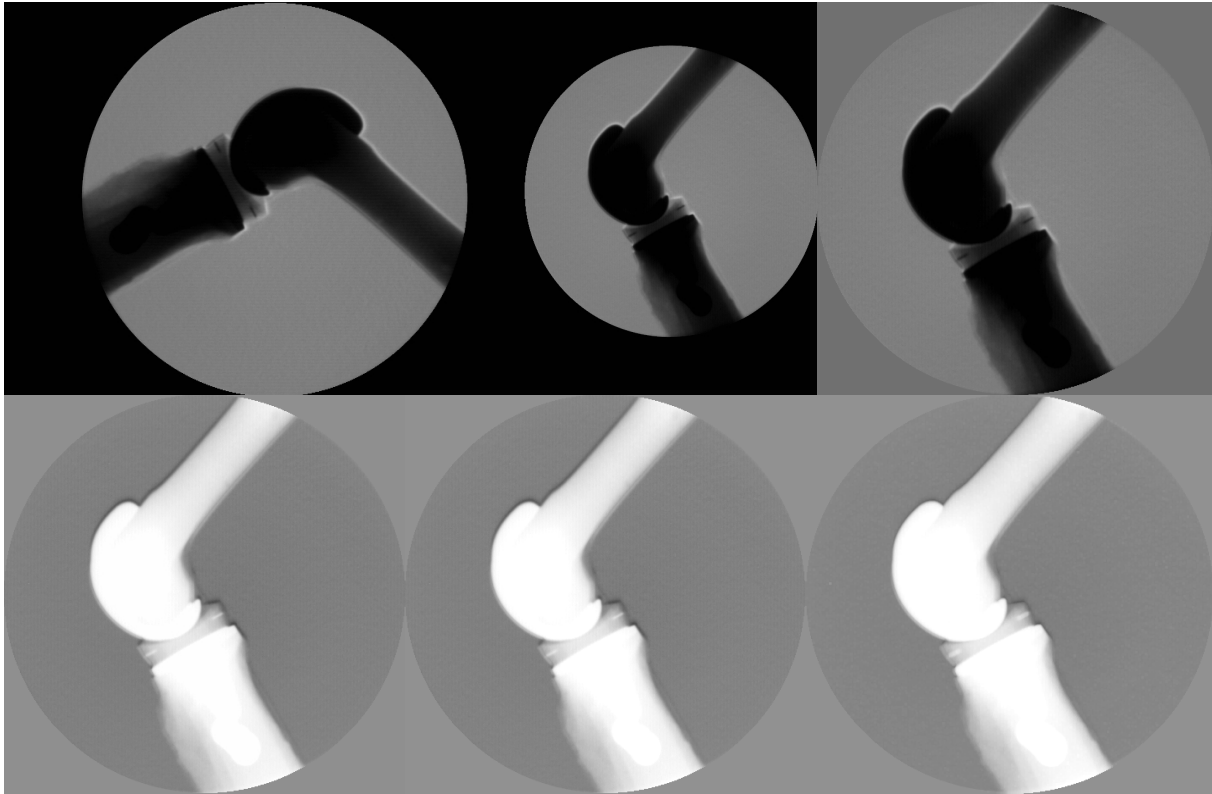


Fig. 27. X-ray image preprocessing sequence.

to the image intensifier the size of the projection can change.

However, compared to the marker case, the background is very uniform and stable. The outer border of the image and the background; both have nearly constant and uniform intensity. The image of the knee is rotated compared to the normal orientation by 90° . Before the process of finding the knee is started the images undergoes a few preprocessing steps presented in Fig. 27. The image is first rotated to its natural orientation. Then the border is exchanged with a constant intensity smaller border so the background and the border are of similar intensity. In the next step, it is inverted so the knee is of high intensity and the background stays at low intensity. Then the contrast and brightness are adjusted, followed by a morphological processing to remove small artifacts. At this point the image can be used to locate the knee.

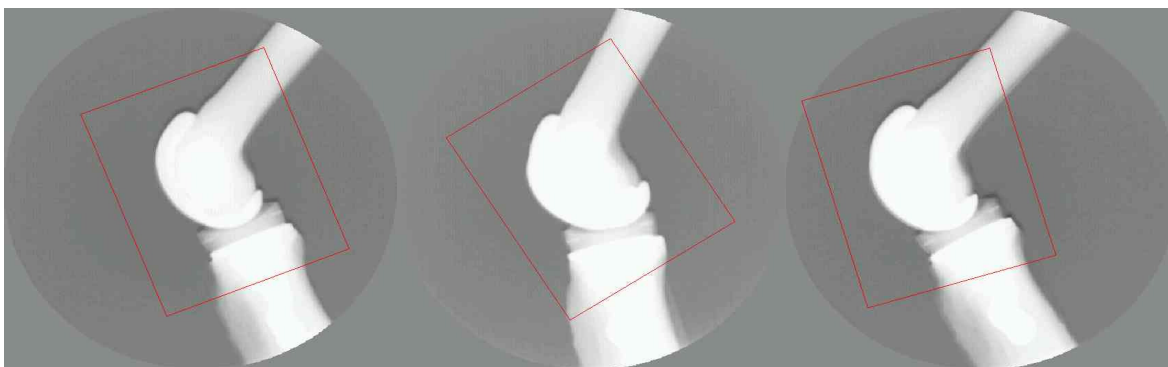


Fig. 28. Artificial knee tracking.

In the last step, the artificial knee was matched with a predefined template. A sequence of matches is presented in Fig. 28. In order to match and track the artificial knee, a template was created (Fig 29). It had to have a proper amount of background in order for the matching function to find it successfully. The red regions in the figure represent the parts of the image to ignore. The field of the view is circular and the assumption was to create a template of similar boundaries.

The image of a natural knee is less regular than of the emulated knee and the soft tissue is also present in the image. The bones are in motion and slightly changint in intensities and projections.

After several test and careful study it was concluded that one of the most consistent, yet unique image sections is the region of the knee at the femur/tibia interface (Fig. 30 and 31).



Fig. 29. Knee template.

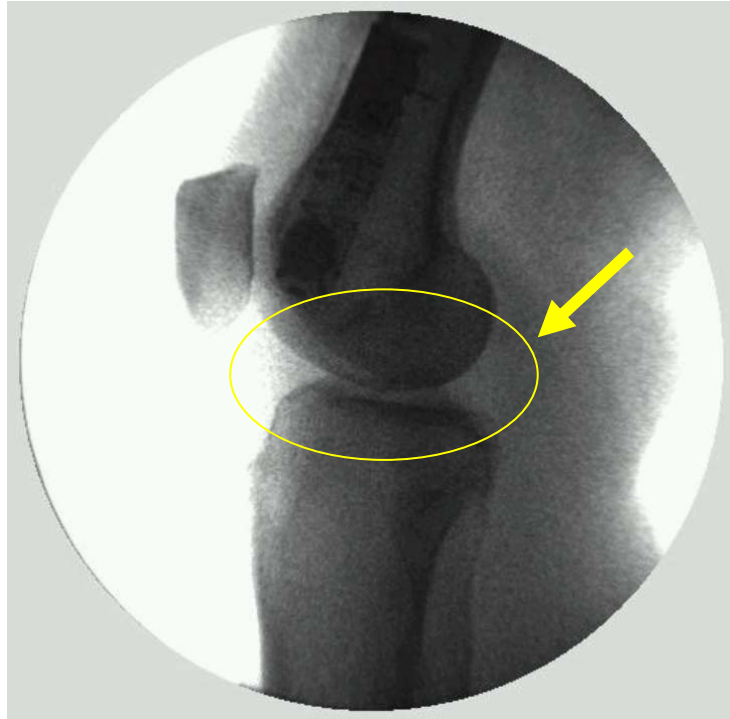


Fig. 30. Knee with the region of interest marked.

The center of it is also the favorable center of the X-ray cone. Having this part of the joint in center of the image guarantees proper perspective and full view of the joint of interest.



Fig. 31. Region of interest extracted from the image.

Teleoperation

In this framework, a different approach was used to locate the objects of interest. Considering the nature of unstructured environment it very hard to declare or define the object of interest ahead of time. The target object can be different during each task and change in its perception to the vision sensor drastically. The same applies to image filtering and preprocessing.

It was decided that the object has to be pointed out by the operator; hence appropriate tools were introduced in the control panel. In the first step the operator either with the use of his/her finger or a pointing device draws a region of interest (Fig. 32). Next the operator has to

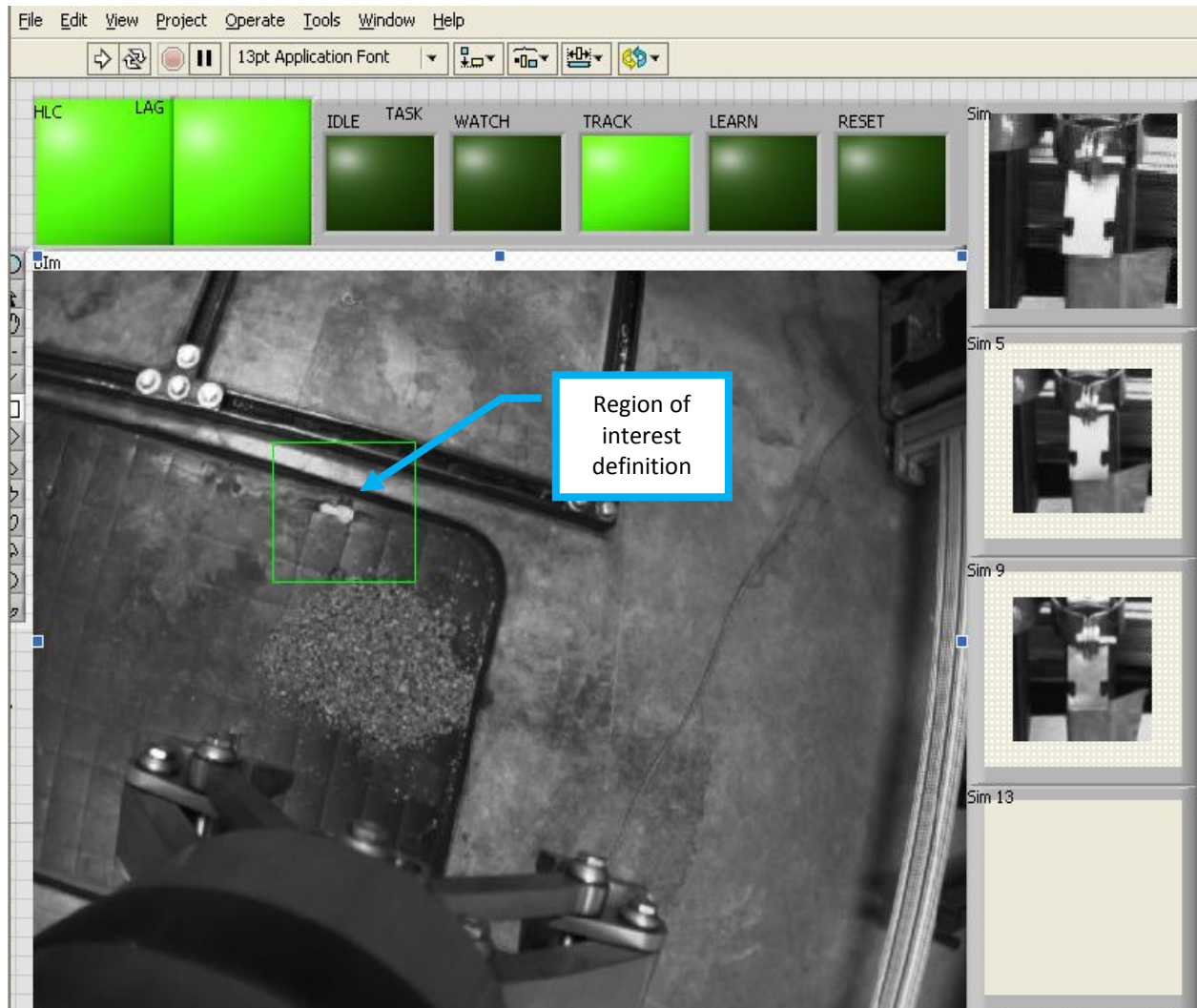


Fig. 32. Manual definition of the region of interest.

tell the agent to learn the features of the drawn region. If both the vision sensor and the object stay still, the vision agent can easily find the target. The system can begin this action if it is directed by the operator to track the object. During every cycle of finding the object a measure of similarity is calculated. If the value falls below a certain predefined threshold, the agent continues to track but also learn features of the new acquired and found regions. Then it searches for both set of features representing the same object. The operation can continue until the system meets the criteria of successful tracking or the target is lost or the memory for allocating new features is used and the target is lost. The system was designed to learn and track up to 16 sets of features. The maximum number of sets of features is limited only by the performance of the system. Large number of features increases searching time but conversely improves the probability of successful execution of the task. This approach allows for changes in illumination and can help with objects with high reflectivity.

In this concept the cycle of tracking can be also a cycle of learning if it is necessary. A set of automatically learnt objects is presented in Fig. 33.

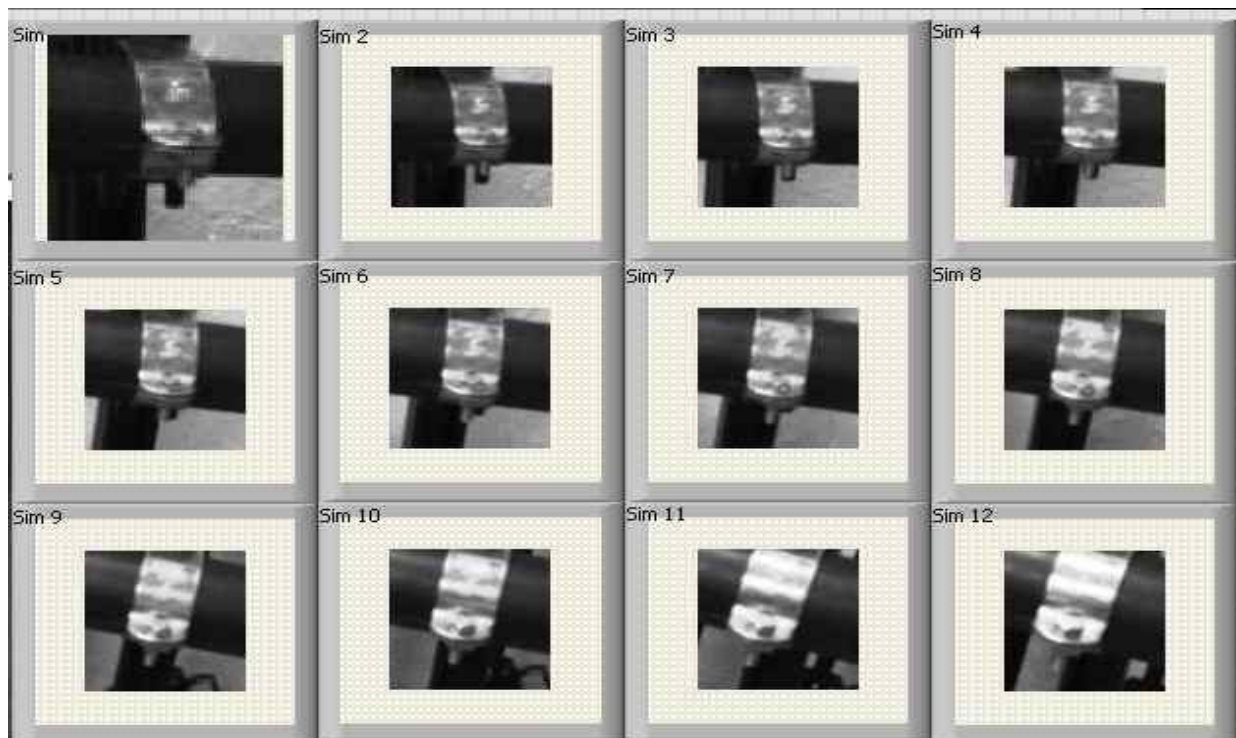


Fig. 33. Series of images feature learned automatically.

Template matching

In all cases the technique used for locating the object of interest was pattern matching. It is a well established method for finding objects in the image. It measures similarity between the saved template and a feature present in the image. As a result, it provides a score - a measure of similarity, and a location of the object with the presented similarity. This tool is designed to work under various conditions: changing orientation, small scale changes or light changes.

The underlying engine behind the chosen pattern matching is normalized cross correlation. The basic cross correlation is described by the following equation:

$$C(i, j) = \sum_{x=0}^{L-1} \sum_{y=0}^{K-1} w(x, y) f(x+i, y+j)$$

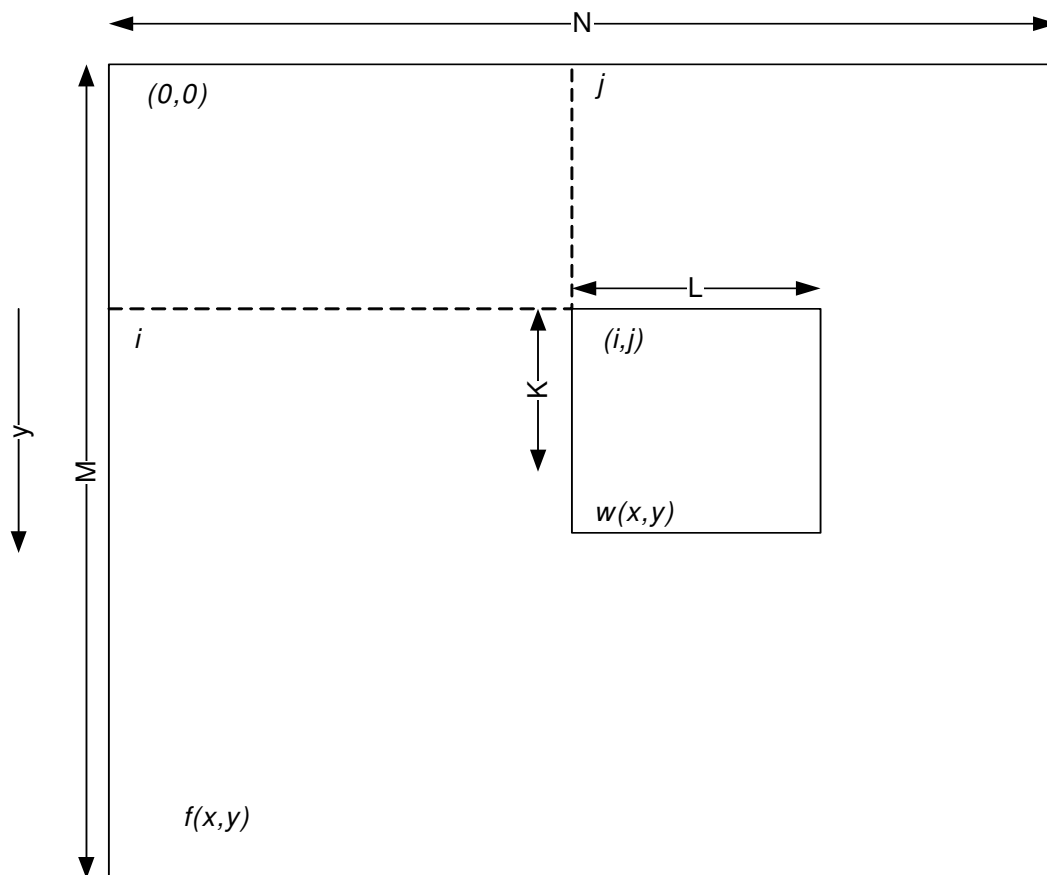


Fig. 34. Correlation process of the template and the image.

A template is defined as $w(x,y)$ of size $K \times L$ and the full image is $f(x,y)$ of size $M \times N$, where $K \leq M$ and $L \leq N$ and $C(i,j)$ is the correlation at point (i,j) . The maximum value of C points out the region of the best match. Fig. 34 presents the correlation procedure. The template window is moved horizontally and vertically to calculate the similarity value. The Equation and Fig. 34 originates from [12].

Matching executed this way is very computationally expensive. In order to provide better performance the template is analyzed to extract only the characteristic information. Additionally, the geometry information can be used to find the objects faster as well as pyramidal matching. The latter approach is based on reducing the size of the image and performing a coarse match first.

The template is analyzed and sampled before the match phase. Once the process is done, it is not necessary to repeat it again and the extracted information is saved. The template image has to be acquired first, and if necessary, regions of the template to ignore for the analysis can be defined. Such process is presented in Fig.35. The corners were excluded from

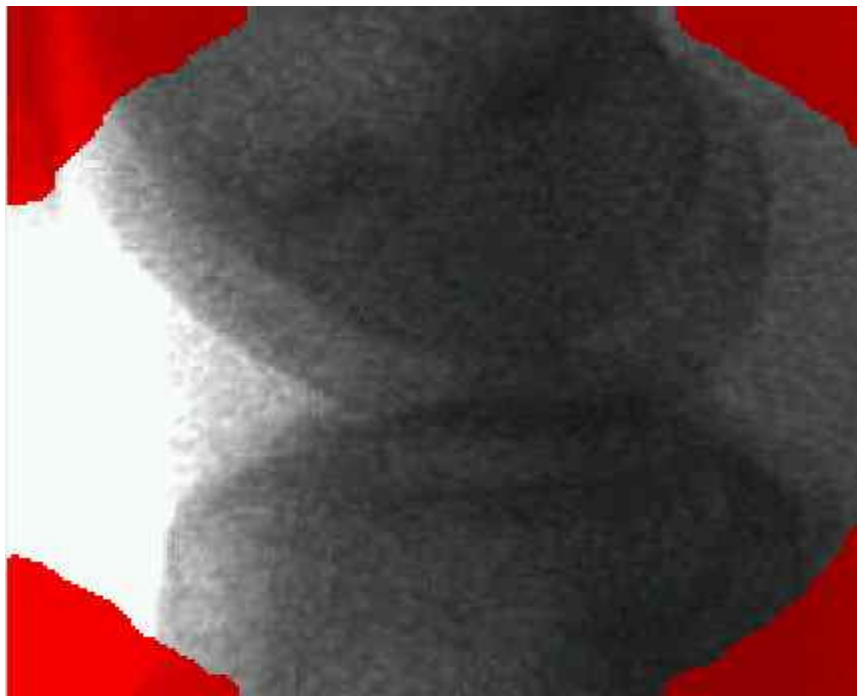


Fig. 35. Creation of the matching template. Parts in red are ignored.

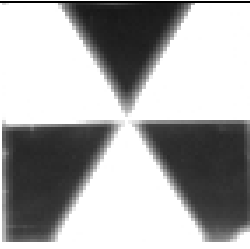
the image due to the fact the useful part of the image is of circular shape. This way, when the object of interest is at on the edge of the image, the template has a higher chance of finding it.

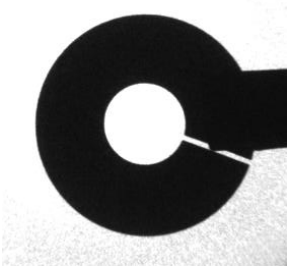
Defining threshold score

The matching function returns in the output the location and orientation of the found image as well as the score of the correlation. In order to classify the object as found, a score threshold has to be defined. Even objects without any apparent similarity can have a positive score. The threshold score is the border line dividing true objects from false matches and for the most part it has be established experimentally. The better the object is defined, the larger the gap between false matches and easier to establish the threshold. The lower the threshold, the higher the probability of finding the object when it is rotated or projected at a steep angle. However, it also increases the chance of finding object that are not the target. Raising the threshold too high will eliminate false matches but can also eliminate true matches. The matching tool used provides scores in the range of 0 to 1000, where 1000 is an ideal match and 0 is no match at all.

The following table presents the score thresholds found during the process of developing the robot.

Table 3. List of types of matching templates and their properties.

Pattern	Size	Role	Threshold value
	71x69	Camera tracking	580

Pattern	Size	Role	Threshold value
	215x184	Human joint knee tracking	880
	281x275	Manual X-ray non-human testing	550
	239x282	X-ray phantom testing	550

Tracking Fluoroscope Safety Systems

The hardware safety system can be divided into three groups: motion systems and its sensors, active safety sensors, and emergency systems. The structure of the safety systems is presented in Fig. 36.

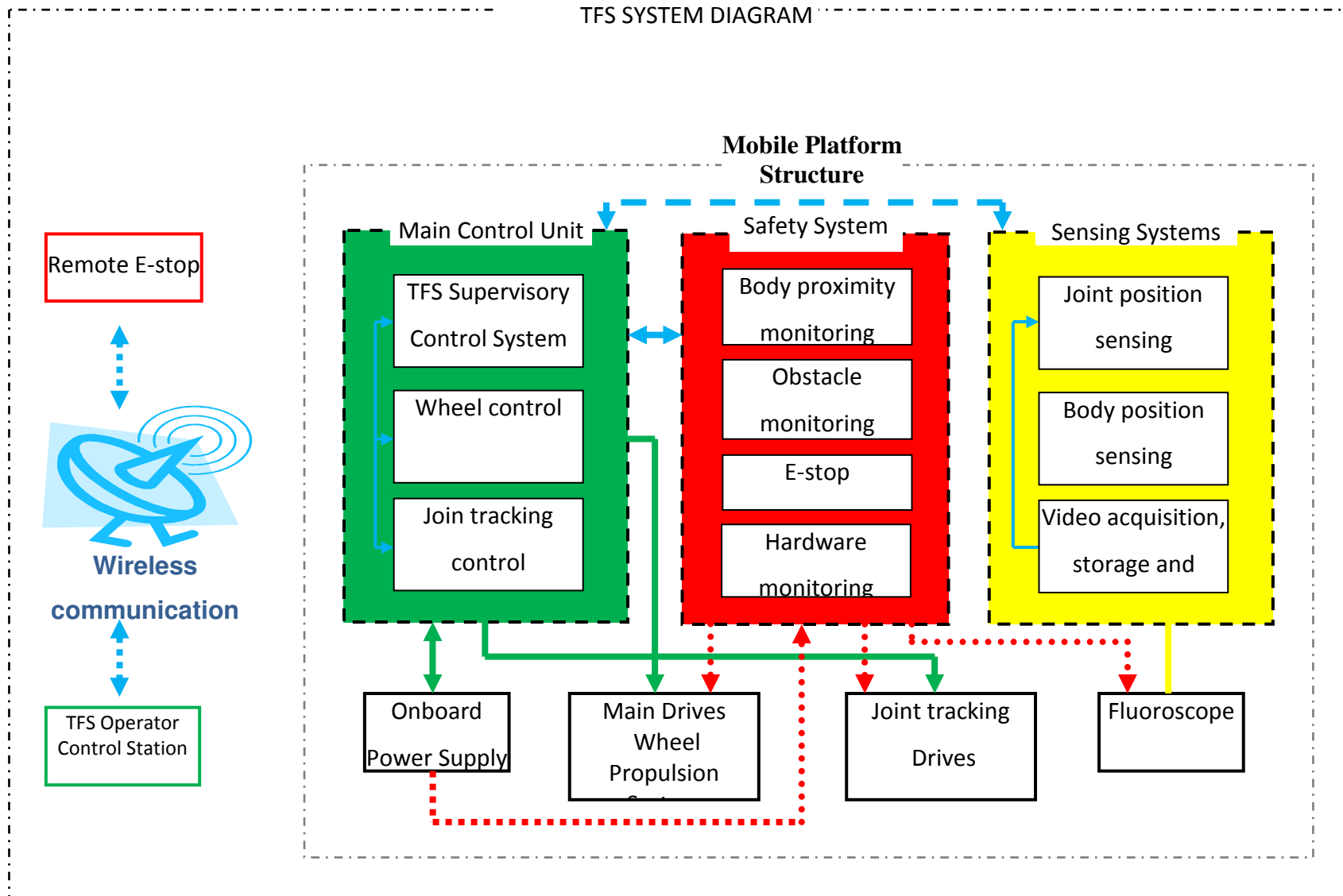


Fig. 36. Tracking Fluoroscope System Diagram.

The motion systems are equipped with software and hardware limit sensors (over travel sensors). These protect the equipment from mechanical damage.

The second group is represented by sensors and devices working purely as an active automatic safety system. It includes: ultrasonic range sensors and bumper contact switches. The platform is equipped with six sonar sensors (in Fig. 37) that monitor its perimeter for unexpected objects or persons. The bumper switches are mounted in the cavity area to detect a contact between the patient and the mobile platform (shown in the picture).

The emergency system includes: a wireless remote control/E-stop, hardware E-stops, Programmable Logic Controller (PLC) and Main Control Unit (MCU). The remote control is an industrial grade radio frequency based system, and includes a remote control and receiver mounted on the TFS. The remote control is held by the operator during patient testing and if released stops the motion of the robot at any moment. The hardware E-stops are positioned to allow laboratory personnel and the patient to easily reach them from any position (In Fig. 38).

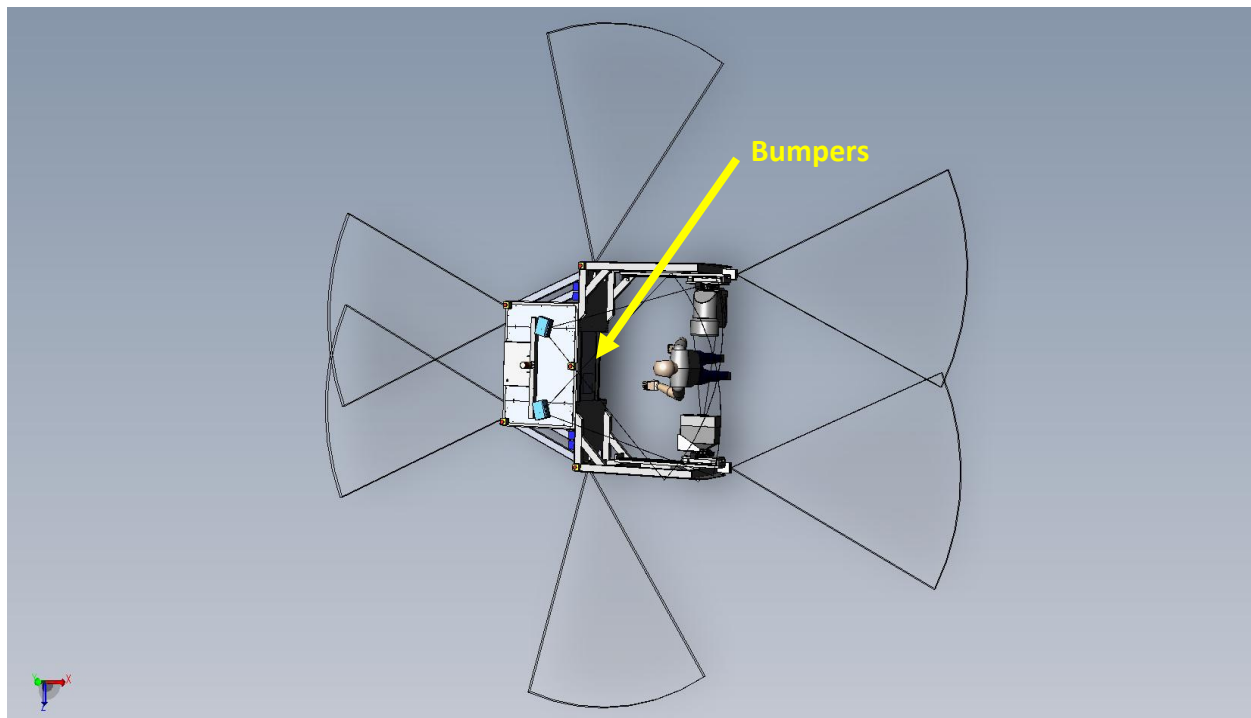


Fig. 37. Sonar and bumper setup.

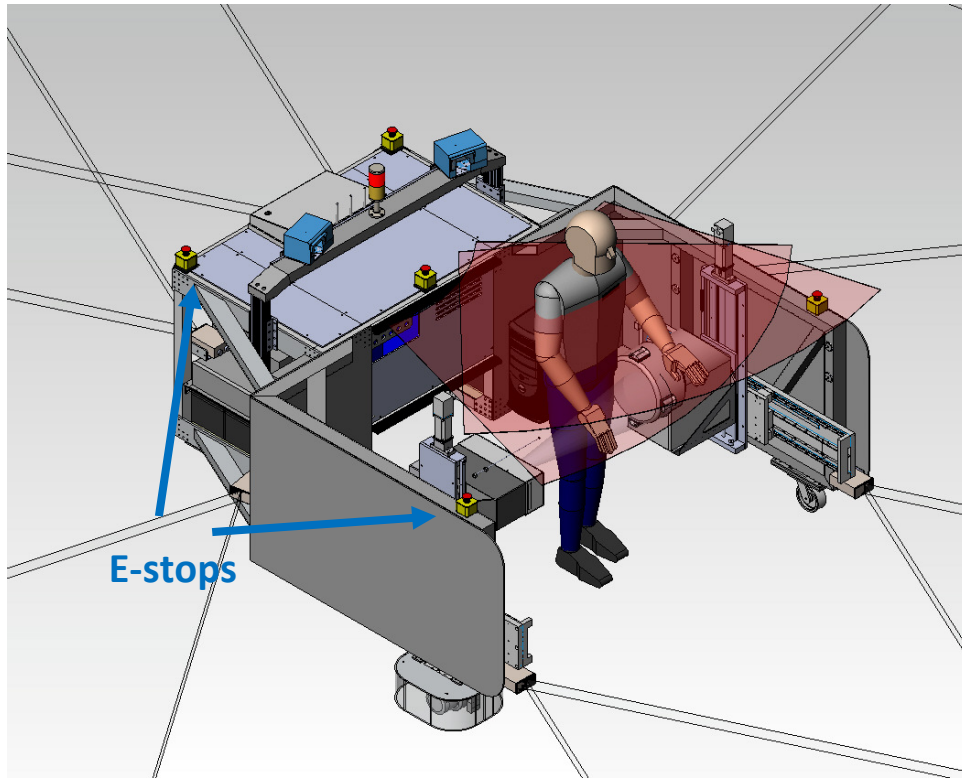


Fig. 38. Emergency stop button configuration.

The safety sensors and servo drives are also monitored by the independent PLC. The remote control and hardware stops are directly wired whereas the bumper switches and sonars are monitored by the PLC, which is also connected to the MCU.

In addition to these safety features, several cognitive features were added. The communication agent senses connection with remote computers. If the computer responsible for motion feedback fails, the MCU shuts down the motion and radiation. If the subject escapes the sensing area the motion controller stops the wheel drives.

Strong safety measures are required for equipment involving nuclear radiation. The receiver and the source of X-rays have to hold the same orientation (axis) with respect to each other. This is accomplished through the servodrives and tested in real-time. The encoders on both sides are read every cycle of the motion loop (333Hz currently) and if the difference

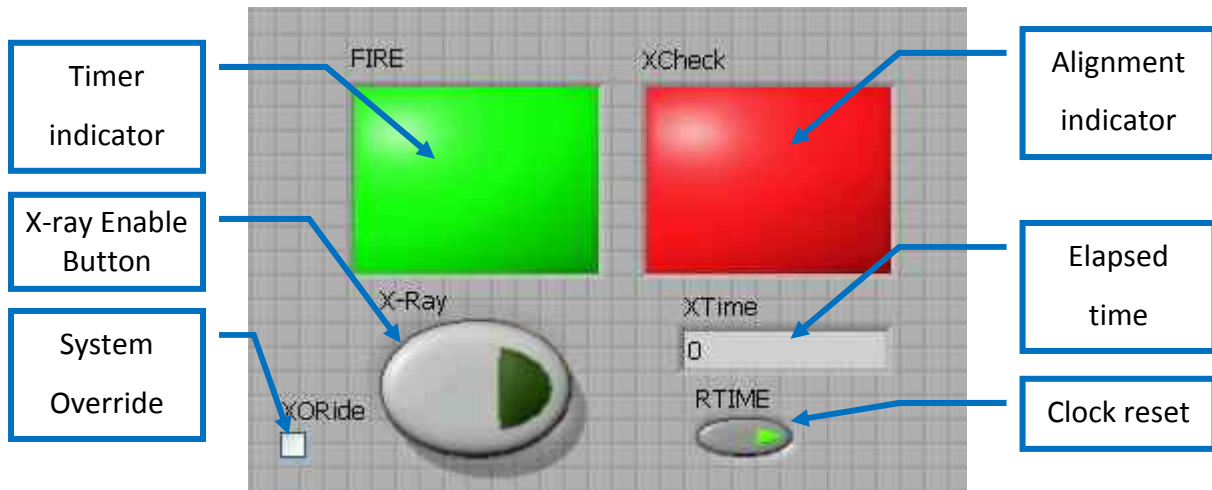


Fig. 39. X-ray control interface.

between the two is above the predefined amount the X-ray system is shut down. This check is executed for both vertical and horizontal axis. In addition a hardware radiation leak detection system was added. If the cone of radiation hits the outer of the image intensifier the X-ray system and motion system are suspended.

The total radiation exposure is also important for upholding safety standards. An independent integrating clock was developed onboard to prevent exceeding radiation exposure limitations. Both the alignment status and timer values are available for the operator to see (Fig 39).

Analysis of the gait and acquisition speed for the TFS

The following numerical example presents challenges that the motion control system has to overcome in order to achieve the required level of performance.

Knee displacement calculation

The maximum relative displacement estimation acquired based on the cycle analysis (Fig. 40 The picture originates from [45]) is roughly 36cm (17cm +19cm) and it happens in 42% of the cycle. The displacement per full step is 86cm (36/0.42). One cycle is considered one step.

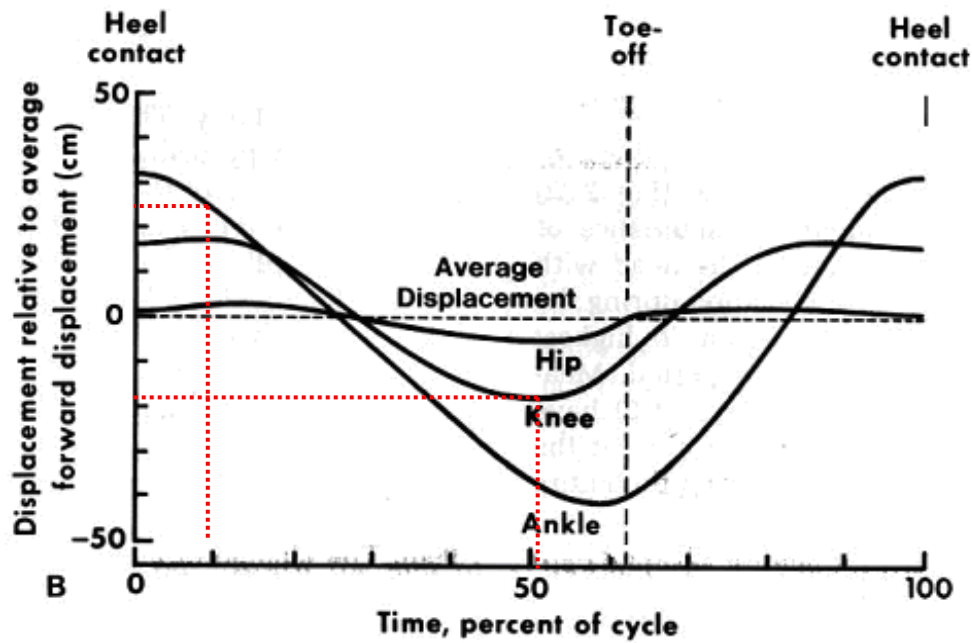


Fig. 40. Human gait cycle.

Additional data

Normal walking speed 2.56steps/s (normal but highest in the age range)

Step size 0.786m

Sampling rate 30Hz

$86\text{cm/step} \times 2.56\text{steps/s} = 220.16\text{cm/s}$ or 2.2016m/s – maximum (worst case scenario) speed of the knee relative to the trunk.

The displacement per image frame is calculated as follows:

$2.20\text{m/s} / 30\text{s} = 0.073\text{m}$ or $2.87''$ per video cycle. That is the worst case scenario but does not include the trunk motion.

Assuming the walk happens at 2.56step/s at the length of 0.786 m, the speed of the trunk is: $2.56\text{steps/s} \times 0.786\text{m/step} = 2.01\text{m/s}$.

Hence, the total amount of motion in the worst case scenario is 0.14cm per frame at 30fps.

The visible diameter is about 15.87cm (6.25") (at the distance where the knee normally is) and from the image point of view the useful diameter is about **10cm (4")** maximum (**1" each way**) (Fig. 41). The object has to be visible entirely in the image. The knee is not a point object. These estimations were based on the following assumptions and observations obtained during the experiments and development:

The total image size is 492 pixels in diameter.

The object being tracked (knee) is about 200 pixels wide.

The allowed motion of the object in the image is $(492 - 200) * 1.1 = 321$ pixels.

The extra five percent in each direction (hence the "1.1" factor above) of motion is the assumed algorithm ability of finding a partially occluded object. The template matching can properly recognize an object that is five to ten percent scaled down or up or partially occluded.

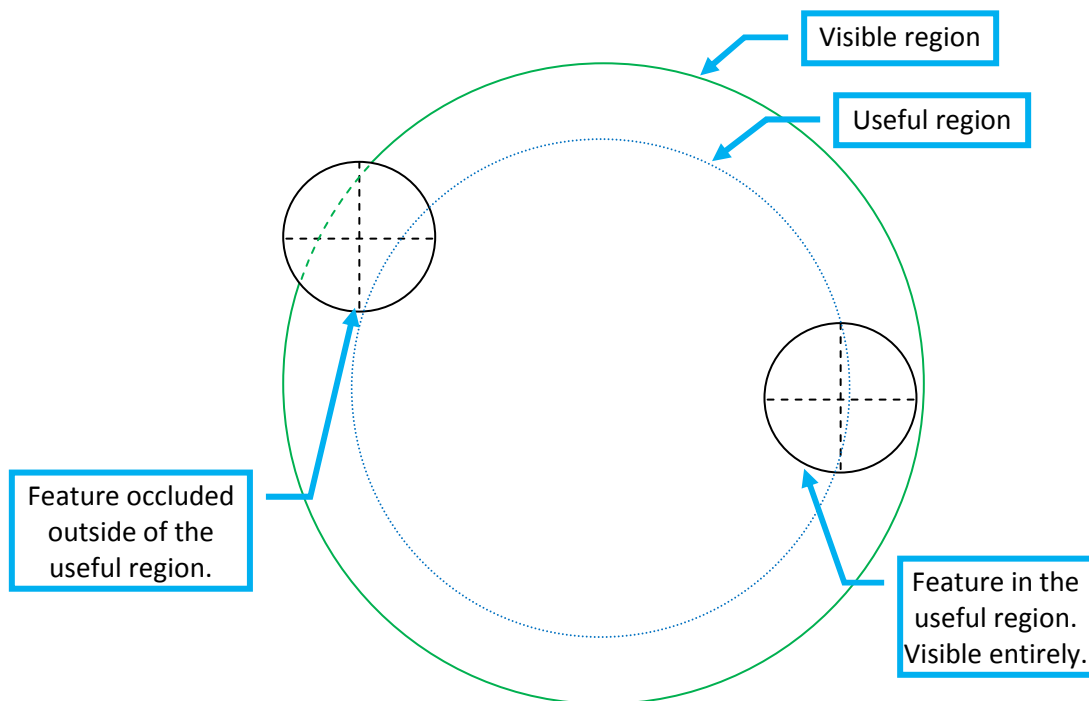


Fig. 41. Geometry of features of interest in the X-ray image.

The amount of motion is:

$$321/492 * 15.87 = 10.35\text{cm} (4.07'').$$

Because the motion can happen in either direction and in order to keep the object in the field of view, the maximum amount of motion can be 50% of this value:

$$0.5 * 10.35\text{cm} = \mathbf{5.16\text{cm}} (2.03'').$$

Forty two percent of the cycle period translates into:

$$1\text{step}/2.56\text{steps/s} * 0.42 = \mathbf{164\text{ms}}.$$

This is about 5 frames long (33ms per frame). This means that for 5 consecutive frames the system has to track the objects under these harsh conditions.

The speed of the object is a sum of the speed of the trunk and relative speed of the object. The relative displacement of the trunk is balanced by the propulsion motion of the frame. However, due to the very high mass and inertia of the robot, this is not fully accomplished. It is expected that the wheel propulsion system can account for 50% to 90% of the necessary motion to keep the trunk at constant distance. It can never be 100% due to the system response time and if it is below 50% it means it practically does not work.

The presented numbers consider only the acquisition delays which are 33ms for 30fps acquisition rate. The total response of the system includes: acquisition time, image processing time, transmission delay, motion control delay, actuation delay and response of the servodrive. When taking these into account, the required speed of acquisition and update rates have to be higher.

The following calculations are an attempt to estimate the required speed of acquisition necessary to track the knee using a standard 9" image intensifier.

The amount of trunk displacement actuation executed by the wheel propulsion system:

Considered : 50%, 70%, 90%.

The amount of motion left to be executed by the joint tracking system:

Considered : 1ms, 0.6ms, and 0.2ms (50 %, 30% and 10 %).

The amount of motion from the relative knee displacement:

2.2m/s (gait analysis)

Expected maximum delay of the system (from image processing to creation of new control signal, does not include motor response time):

10ms. (Assumption worst case scenario)

The maximum allowed motion per single image frame:

5.16cm.(calculated above)

Maximum total speed under consideration:

3.2m/s, 2.8m/s, 2.4m/s (2.2m/s + the trunk motion).

Additional displacement of the object in 10ms system delay:

3.2cm, 2.8 cm, 2.4cm. (0.01s x total speed)

Maximum allowed motion per single frame considering the delay of the system:

1.96 cm, 2.36cm, 2.76cm (5.16cm – displacement during 10ms).

The frame rate required to keep the object in the field of view considering the above constraints:

$320/1.96 = 163fps$, $280/2.36 = 119fps$, $240/2.76 = 87fps$ (respectively).

The same calculations run for the **5ms** system delay gives the following results: 90fps, 74.5fps, and 61fps.

The high image acquisition rates require the camera to have very short exposure times. This, in effect, makes the amount of light reaching the image chip to decrease. For the object to be properly visible in the image in this case, the X-ray intensity has to be increased.

It is apparent that based on the presented calculations, the main drawback of the system is the size of the screen. The bigger the screen the lower the acquisition rates have to be, the lower the dose of the radiation and the maximum load on the motors.

Agent structure - TFS

Human-robot interaction and motion coordination requires many tasks to be accomplished at the same time. This applies to physical or computer system tasks on the same machine, environment or different location/computer system. Those tasks are executed by several independent agents that communicate with each other and the user and release the resources when their actions are not needed.

The main system components of the TFS are presented in Fig. 42. All the TFS software was written in LabView™ code. The components running onboard the TFS were under real-time operating system control. They were developed to take advantage of the determinism of the system and the parallel execution. They run at different frequencies, but as much as possible it was ensured that the exchange of data between threads be efficient and without interrupts. Every subsystem in Fig. 42 that is enclosed in its own block runs independently of the others.

Camera and network management

The first agent under the Image processing unit is the camera and network block. Its primary role is to establish the TCP/IP communication between the remote control host and system devices. It includes: Main Control Unit, two laser scanners and the Remote Control Unit. In the case of Remote Control Host, it creates two connections, one for video transmission agent and one for the watchdog agent. The agent is also responsible for establishing communication with two digital cameras and setting its acquisition and control parameters as

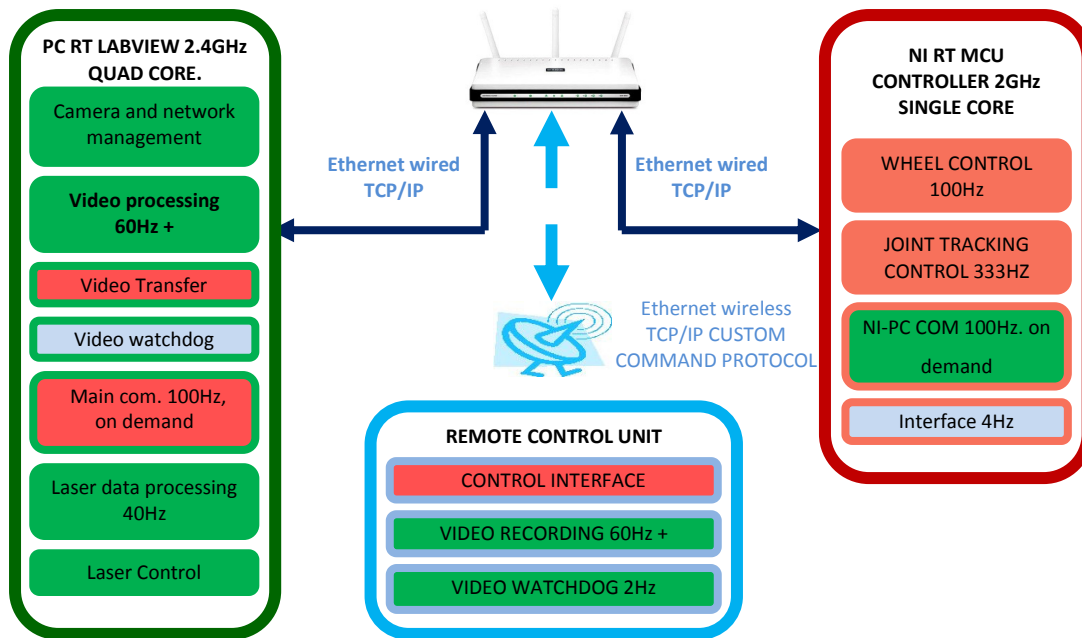


Fig. 42. TFS system topology. Colors are representing the data exchange and source.

well as switching between them.

Once it has achieved communication, the agent's execution is suspended. If the system senses that one of the connections was broken, it automatically tries to reestablish it until successful. Since the system is able to track with either the X-ray system or the camera, it is necessary to change, for example, pattern templates or coordinate frame details. This occurs online, without shutting down the system or restarting the software. The request to change may only come from the operator's console.

Video processing agent

The video processing agent can work with two different kinds of input. The first one is the external camera tracking the joint marker and the second one is the X-ray imaging system. The video agent is activated from the remote console interface. Depending on the image source the agent loads proper match patterns and their settings. Settings include: a predefined pattern file, matching threshold and image features like resolution, shape or masks. The X-ray settings can be additionally customized through a configuration file. Three different features can be set

there: type of X-ray pattern, its match pattern files and recognition thresholds. In order to change those settings, the image processing computer has to be rebooted, but the software does not need to be recompiled. A sample configuration file is presented below in the Table 4.

Table 4. An example of the vision configuration file.

1	Score:580
2	XScore:870
3	XRobScore:550
4	Xtaks:Human
5	MRFile1:C:\ni-rt\CirPattBig.png
6	MRFile2:C:\ni-rt\CirPatt.png
7	MPFile1:C:\ni-rt\Pat1.png
8	MPFile2:C:\ni-rt\Pat2.png
9	example:
10	"FScore:15" Reading field is 7
11	"XScore:27" Reading field is 7
12	"XRobScore:690" Reading field is 10
13	"Xtaks:Human" Reading field is 6 no length specified.
14	"Xtaks:Robot" another example
15	"MRFile1:d:\TFS\sick\markers\XRmarker2.png" Reading field is 8

Once all the features have been loaded and the operator starts the acquisition process, the agent starts executing its normal functions. It acquires and preprocesses the image. In the next step it performs pattern matching based on two different references and adds the numerical position of the found pattern to the image. In the next step it passes the position data to the communication agent and the image to the video transfer agent. If there is a problem found with the camera it switches into idle state and informs the camera and network agent. It can be set into the idle state any time by the operator as well. This agent is completely independent of motion control, which means that it can function without the motion control agent turned on. This way it allows for testing the quality of the recognition process without risking abrupt motion. It also allows for displaying the feedback without motion being involved.

Video transfer agent

The video transfer agent is responsible for sending the image to the remote control console. Once it gets the signal from the processing agent it packs the image into a proper form for transferring over the Ethernet using TCP/IP and waits for another image. When there is no image to send it stays in the idle mode. If the transmission was not successful it informs the connection agent about the problem. If any of the transmission setting are changed, the order of updating is as follows: connection agent, video processing agent, transmission agent. The video processing agent forwards the transmission details along with the data.

Video watchdog agent

The video watchdog is an additional agent running in the background at a low frequency to ensure the video connection is alive. If the two computers are not exchanging data, it is hard to find out if the both of them are still online or responding. The watchdog sends a single byte packet twice a second to make sure the other side is responding. If not, it reports to the connection agent to reestablish the communication. This allows for fast, reliable communication without unnecessary delays.

Main communication loop agent

The main communication loop is responsible for exchanging the control data with the operator's console and reporting the position of the object with respect to the robot. This includes image features and distance to the trunk of the patient's body. It sends the data based on two triggers. The first one is a constant 100Hz communication loop. The second one is on demand request from the image processing agent or laser processing agent. These send communication requests when they have a new set of data. This way the motion agent receives new tracking data as soon as it is available. The list below shows all the current data exchanged between the Image processing computer and the remote interface.

Image processing computer to remote host

- Laser presence

- Laser reading status
- Body position algorithm
- Body angle
- Body position (2D)
- Laser error status
- Vision system status
- Video error status
- Vision communication status
- Joint position (2D)

Remote host to the image processing computer

- Laser ON/OFF
- Body position algorithm
- Vision ON/OFF
- Vision source.

If the communication fails, the agent reports it to the network agent and goes into an idle state until the connection is reestablished.

Laser data processing agent

The laser data processing agent is an integral part of the body tracking subsystem. Its role is to receive the data from the scanners, process it and provide body position information.

A detailed description of this agent is included in the *Human subject body tracking* section.

Laser control agent

The laser control agent is responsible for turning the laser range finder on and off. It requires sending a specific command. The agent checks if the communication is established and if there is a request from the operator to turn the laser on or off. This agent only sends data one way and checks if the transmission was successful. If it fails, the agent reports it to the communication agent.

Video recording agent

This module performs two main tasks: displays the real-time video feedback of the current source and saves the feedback to a file. In the first step it establishes two connections: one for receiving video stream and another for the watchdog agent. Once this part is successful, it waits for video data. Since it depends on the operator to turn the video transmission on, the agent is set to wait indefinitely for the first data packet. Under these conditions, it is not possible to confirm connection status hence, the watchdog agent is used to monitor it in the background.

When the first frame arrives it is considered as a signal of continuous video acquisition and the agent switches into a displaying and /or recording state. In this state the operator can do any of the following options: record the images, permanently add the overlay to the image or manually reset the connection. At this point the image is always displayed on the screen. Once the frame is received the overlay data is extracted from the image. If the operator chooses to record the images, the image is stored in the AVI format without compression. The overlay information by default is attached to every frame in the additional data location specified in the AVI format. If the operator chooses to, it can be permanently attached to the image (becomes visible).

The agent monitors three types of occurrences. Since the images are stored without compression, it can quickly reach the file size limit, hence before that happens the agent directs

the data stream to a different file that is already open and ready to be written to. The maximum size of the file depends on the operating system. The second and third type is related to the connection status. It can be broken or the operator may choose to stop the transmission. Both of these situations are sensed and the agent can switch either to the first initial state or the idle state. The agent can be forced to switch to the initial state also by the operator or the watchdog agent.

Watchdog agent

The purpose of this module is to monitor communication between the remote host and the video processing computer. The system uses TCP/IP sockets for data exchange. In order to sense a broken or closed connection one of two kinds of events must happen: the remote part closes the connection or one of the packet transmissions fails. When one side closes the connection it sends a special message to the counterpart. If there is no such message or ongoing communication, the system has no way of telling if the connection is alive or not. Hence, the only way to monitor it is to have an agent exchanging messages and thus testing the connection.

Wheel control agent

The wheel control agent's role is to provide proper wheel propulsion to the robot. It has two modes of operation: manual and automatic. Under the manual mode it outputs a voltage signal to the servodrives running the motors at the torque proportional to the signal.

In the automatic mode, the execution is far more complex. It has to take into account the following parameters: last known body position, desired distance setting, update in the body position and interaction with the joint tracking system. Last known body position is received from the laser processing agent through the communication agents. The body position is additionally filtered to provide smooth motion. Along with the desired distance, these are the two main inputs for the PID control. The subsystem is asynchronous with its feedback but it can sense when new feedback arrives and react properly. This means it receives data that has been

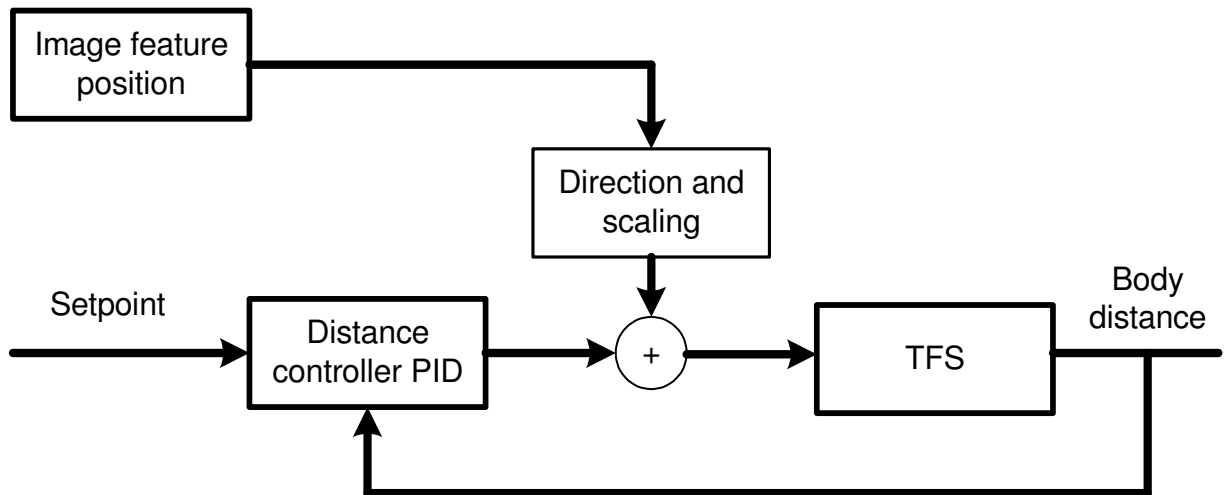


Fig. 43. Wheel control system.

processed and not processed at different rates than the rate of the system. This makes the system robust and independent of unexpected changes in performance of other components.

Due to known dynamical interaction between the wheel propulsion system and the joint tracking system, a change to standard the PID system was applied. It uses feed forward (Fig. 43) technique to counteract the inertia forces from the joint tracking motion system. If the direction of the PID output is same as the direction of the joint control system, the PID output is increased, if the direction of the PID output is opposite to the direction of the joint control the output is decreased. The amount of feed forward output is proportional to the expected joint tracking travel in horizontal direction. This way the interaction is minimized and the tracking error decreases.

The specific constants were obtained experimentally using the Ziegler-Nichols method. The system was expected to bring the error to zero in a half cycle (critically damped) and the normal steady state error was expected to be within 10cm.

Joint tracking control agent

The role of the joint tracking agent is to find the best estimation of the desired motion

output based on the input given by the image processing agent. A detailed description of this agent is given in the *Visual servoing agents and strategies* section.

MCU communication agent.

The MCU communication agent is the counterpart of the image processing agent. It establishes the communication and exchanges the data received with its local agents. As with other agents, it is fully autonomous and reestablishes the connection automatically. If it senses a fault it shuts down the wheel propulsion in order to avoid uncontrolled events.

Interface agent

The MCU has a module responsible for updating and reading the interface as well as monitoring non performance features of the robot. It reads the power level, current usage remote control status, enables and disables the brakes, X-ray servomotors, senses the total time of radiation and exchanges these data with motion agents. It executes at the rate of 4 Hz.

Agent structure – Teleoperation

The agent structure of the teleoperation framework is presented in Fig 44. The communication and connection management agent in the teleoperation system are executing the same kind of tasks as the corresponding agents in the Fluoroscope System. Significant differences are in the image processing agents, interface and the arm control.

The most significant difference is in the image processing and displaying. The operator has a direct view from the camera. When a fresh, new object of interest is defined with the use of the interface, the vision processing agent receives only the image coordinates and extracts the features from its own copy of the image. Then the extracted mini image is send back to the interface for displaying. The vision processing agent can send back three types messages: system status update, status update with the full camera view and system update with the full view and with the new object of interest (manual or automatic).

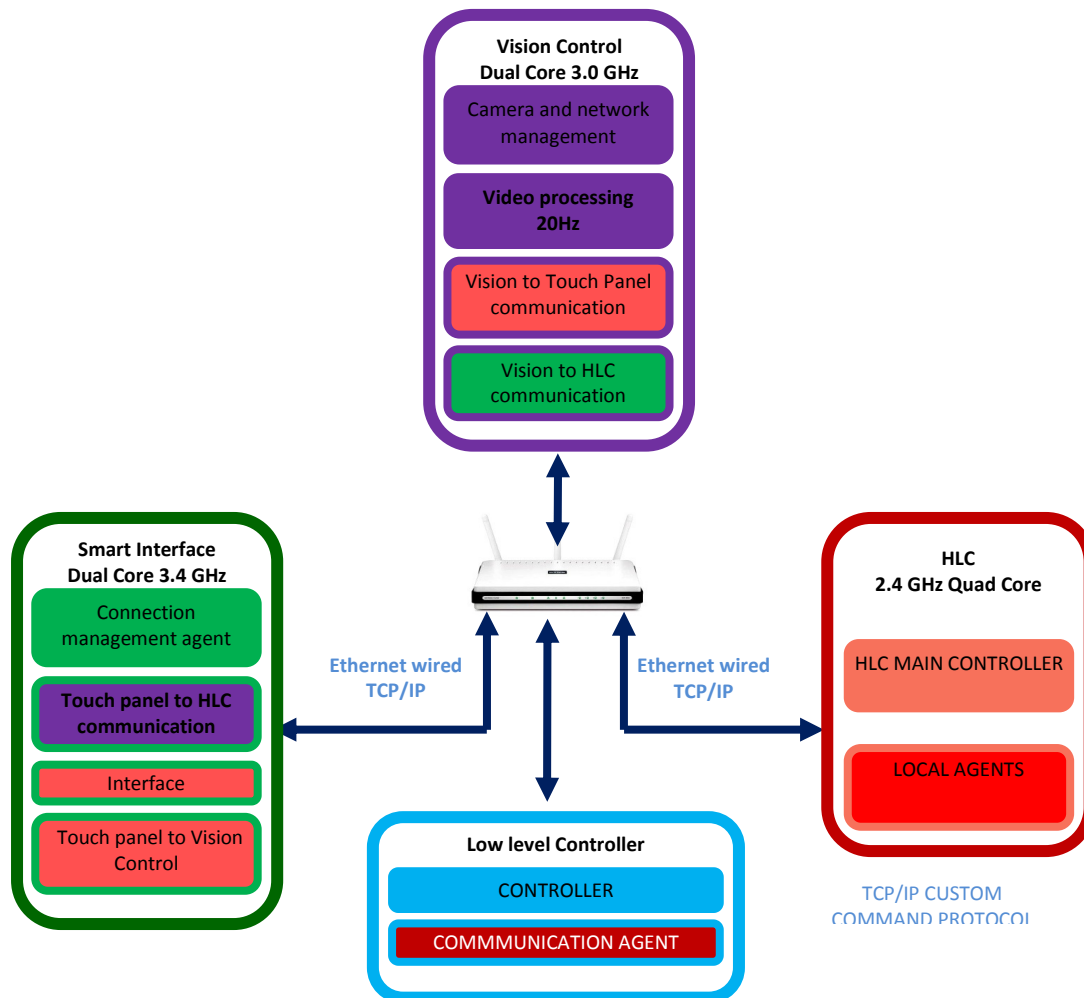


Fig. 44. Teleoperation system topology. Colors represents the data exchange and source.

Control interface

Tracking Fluoroscope System

The control interface is broadcasted from the MCU to the operator's console as a native LabView virtual instrument (Fig. 45) or in the form of a website through a browser. If necessary, the control interface can be run from another device/computer with proper network settings and data processing power.

The operator can observe and/or change the following aspect of the system: laser and video tracking, system power status, X-ray status and motion control and networking.

The following parameters of the laser tracking can be observed: connection to the laser range finders, laser measurement status and its error status. The operator can turn the measurements on and choose the algorithm type used to find the position of the object.

The interface shows two kinds of information about the video system: general system status and display status. General system status tells the operator whether all the video devices are connected and the video agents active. The display status reports that vision acquisition is on and the data transmission is functioning properly. The only action the operator can take is to turn the acquisition/transmission and display on/off and choose between the external imaging device or X-ray system imaging.

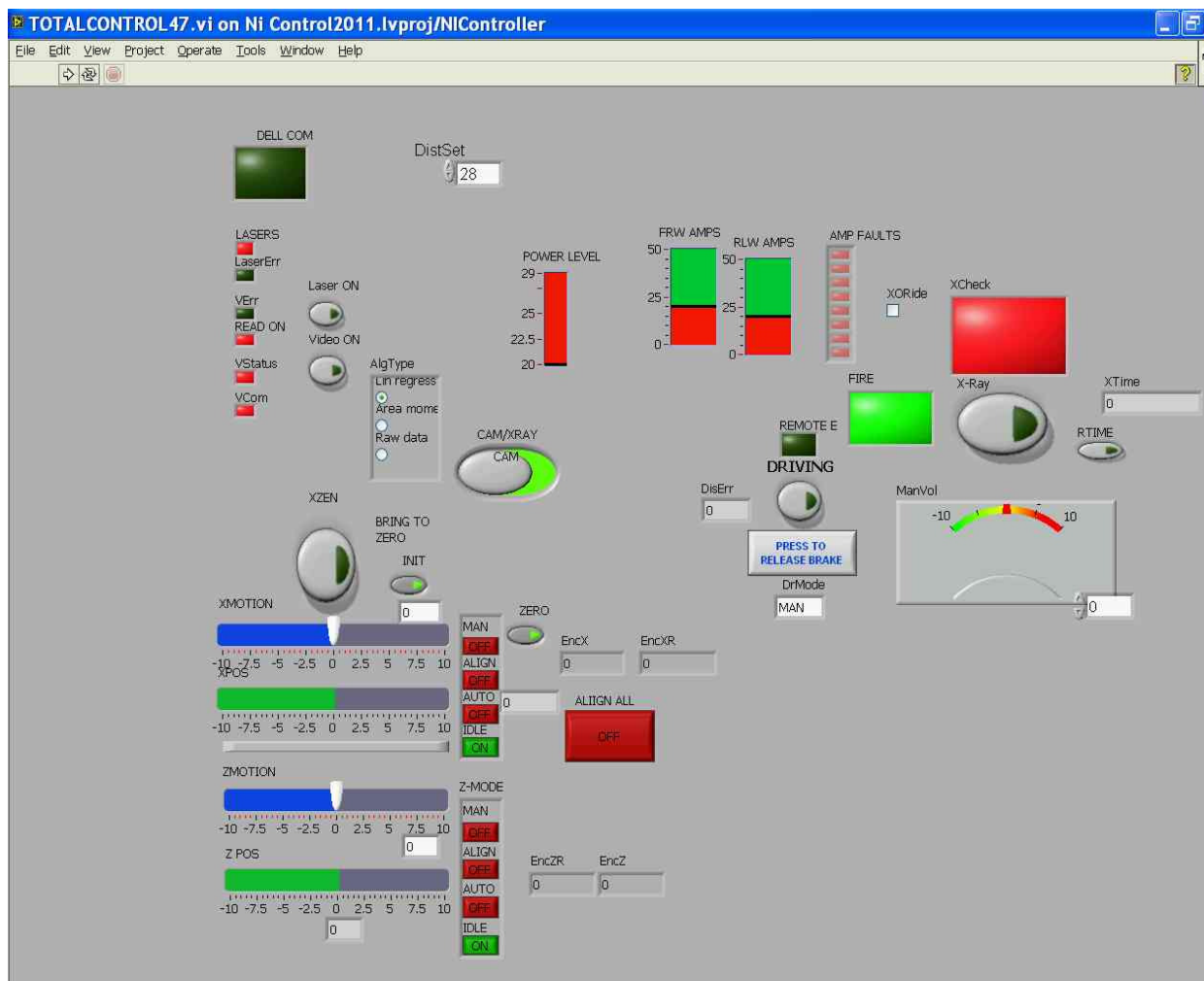


Fig. 45. TFS control interface.

The interface also shows current battery pack power level as well as wheel motors current usage and enabled servodrives.

In the case of the X-ray system, the following parameters can be observed: X-ray alignment status, X-ray radiation time elapsed, and over time status. The first parameter tells if the x-ray source and X-ray receiver are properly aligned, the radiation clock is showing the total accumulated radiation time since last reset and the overtime status shows if the maximum time was used. The operator can turn the X-ray switch on/ off, reset the timer and override the alignment and maximum time settings. The override feature can be used for testing and troubleshooting situations.

The motion control system includes the wheel propulsion control and the joint tracking control. The wheel control part allows for enabling and disabling the brakes and enabling and disabling the servodrives. The operator can control the wheel propulsion in two ways: automatic and manual. Under automatic control, the operator can set the desired distance of the subject and robot and monitor current position error. The only required action is to enable the servodrives and choose the proper mode of operation. Under manual mode, it is possible to drive the robot forward or backward using the provided manual selector of output signal. It is proportional to the wheel torque.

The joint tracking control has four main modes: idle, manual, auto and align. If the servodrives are enabled the robot always starts in the idle mode. In this mode the positions of the motors are held constant. In the manual mode the operator can move the device to a desired position independently for the horizontal and vertical axes. In the auto mode the robot control signal is created based on image feedback and human actions are not necessary. The last mode is a transition or alignment mode. In order to smoothly transition from a random position to the first desired position without an abrupt motion, the image feedback is used at a every motion rate. When the position error gets between the desired limit, the mode is automatically switched to auto control. The interface is equipped in two scales: one showing

the desired position and one showing the actual position. Additionally, the operator can observe the current encoder readings.

The second part of the control interface is composed of the interface of the video recording agent and is described in details in the recording agent section. It allows for viewing current image feedback and recording the images.

Teleoperation system

The teleoperation system can be controlled using two graphical interfaces. The first one is a part of the HLC with limited functionality. Two graphical interfaces are: High Level Controller GUI and main system GUI. High Level Controller needed its interface for testing, developing purposes and in a case when it is desired to run the system with minimum configuration (without the touchscreen or gripper).

The High Level Controller GUI view is presented in Fig. 46. It shows the modes of the controller, the status of the arm, communication statuses of: Low Level Controller, WAM controller, Roboworks PC and touchscreen. Furthermore it displays operation modes, scaling factors, blocked axes, condition number and four recent system messages.

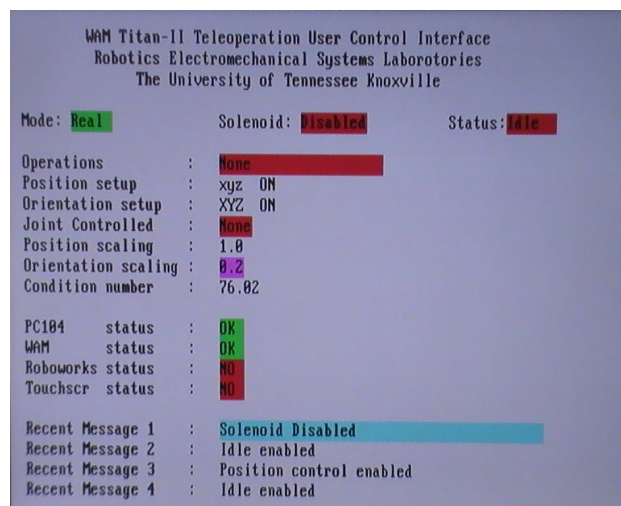


Fig. 46. High Level Controller GUI.

The main system interface was equipped with significantly extended functionality (Fig. 47). In addition to all the functionality of the HLC it allows for changing the input device: WAM, external/local agent, vision or keyboard/manual control. Space and functionality was allocated also for a mini replica of the slave arm (minimaster). This interface also allows for joint and Cartesian control of the arm with custom adjustable constraint. In addition a gripper control was added.

The vision control part allows for remote management of the Vision computer and its agents. The major modes of operation are: idle, watching, learning, tracking and reset.

In the idle mode the vision system is inactive and there is no vision feedback. The watch

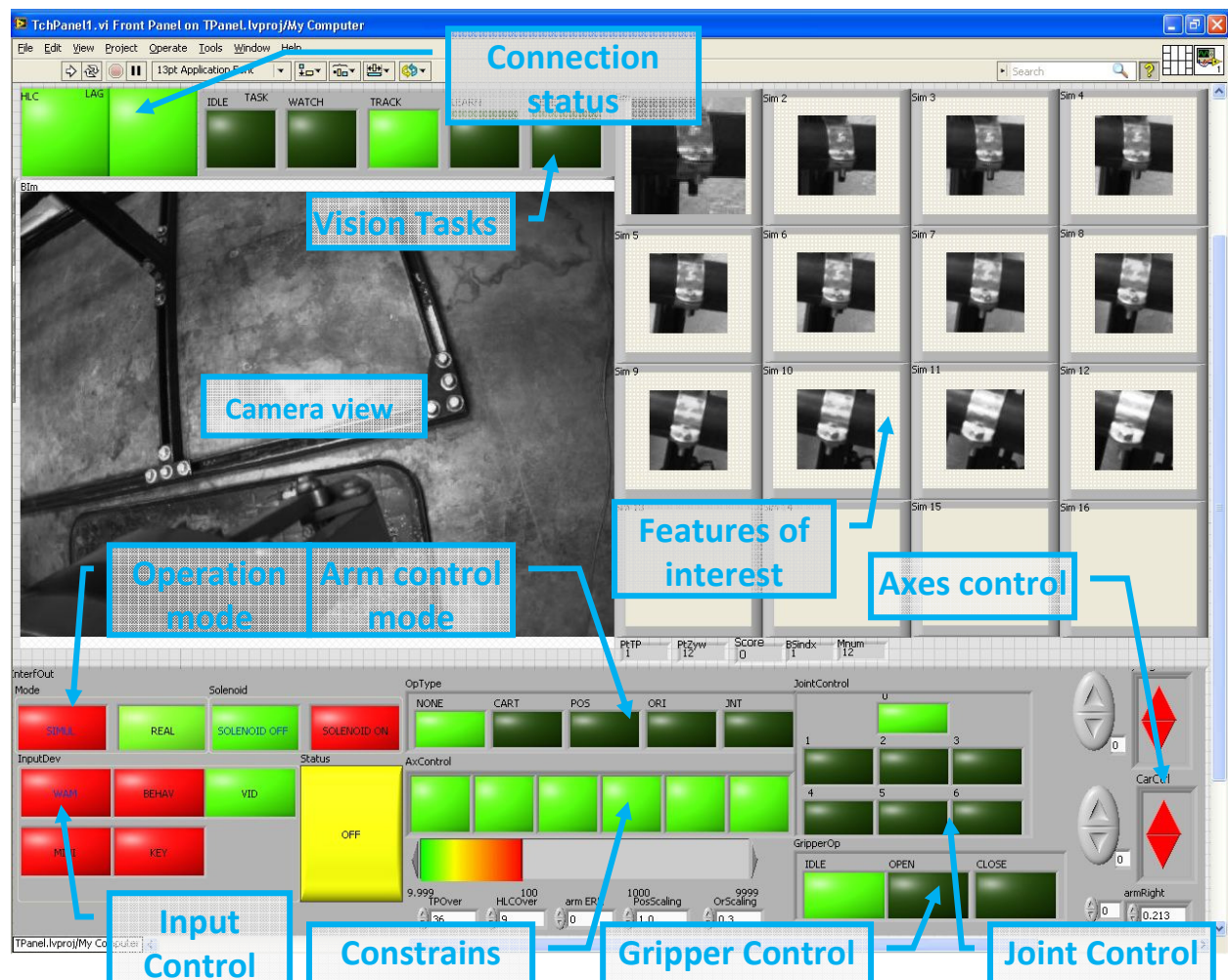


Fig. 47. Main teleoperation control interface.

mode is used to view the camera. The learn mode is used to define a region of features of interest. The region can be pointed with the use of mounting device like mouse or finger (touch screen). The track mode is used to track the requested object in the image. When this mode is in use the field showing the features of interest fills out as the system automatically adds new templates. The last mode resets the memory of the system and brings it back into initial stage.

Additional information is presented in the interface like the current recognition score or the number of the templates used to find the object.

The interface was designed with the intension to be used on a touch panel for improved functionality and performance. The active controls were sized allow finger touch control.

The interface is part of the Compact remote console used for remote control and management of robotic arms (Fig 48.)



Fig. 48. Remote Console Station.

Fluoroscope state machine analysis

The complex nature of human-machine interactions combined with the remote control and decentralized task execution requires proper planning and development of state machines. This concept is well known in the robotic field but the details are drastically different between different types of robots.

It was decided to design a unique and complex state machine for the Main Control Unit. The reason for that was the fact that it controls and monitors a number of devices at a very high rate, as well as performs communication and provides interface to the user. Other systems were using far less centralized architecture.

The MCU has five distinct state machines:

- Network vision data
- Network vision state
- Laser/vision control
- Joint motion control
- Wheel propulsion control

Each state machine has the following properties:

- Client-server architecture
- The server is a call-back routine
- Multi client, single server
- Only one client can access the state machine at a time

- Presents a complete state of the variables of interest
- Is the only unique copy of the system
- The state machine is a combination of values and methods (functions)

The client can be any part of the code or loop that has a set of new values and functions to read or write to the server: The server is only present when it is called by the client and after that it is shut down preserving its memory. There can be any number of clients calling the same server but only one at a time. If the server is executing a request from a client while another one is calling, the new client will be given access as soon as the server is available. This way, there is no possibility of race conditions. Proper design of the server prevents variable overwriting before the client can read them. That is especially useful when one client has to process each portion of data before it gets overwritten by the agent gathering the data.

Each server has the following structure: one action input and one or more data inputs and one or more data outputs, and a set of action cases. A typical state machine used the research is presented in Fig. 49.

Every time a client calls the server it must specify an action, and depending on the action, it might have to provide a proper input. The output of the server is always provided. The actions can be divided into three distinct categories: initialization actions read actions and write actions. The initialization action is executed only once based on the set of initial data saved in the server. The read and write type actions are executed when a client updates or queries for the state of the system. The server can store any kind of data or functions related to the system, for example history of the requests.

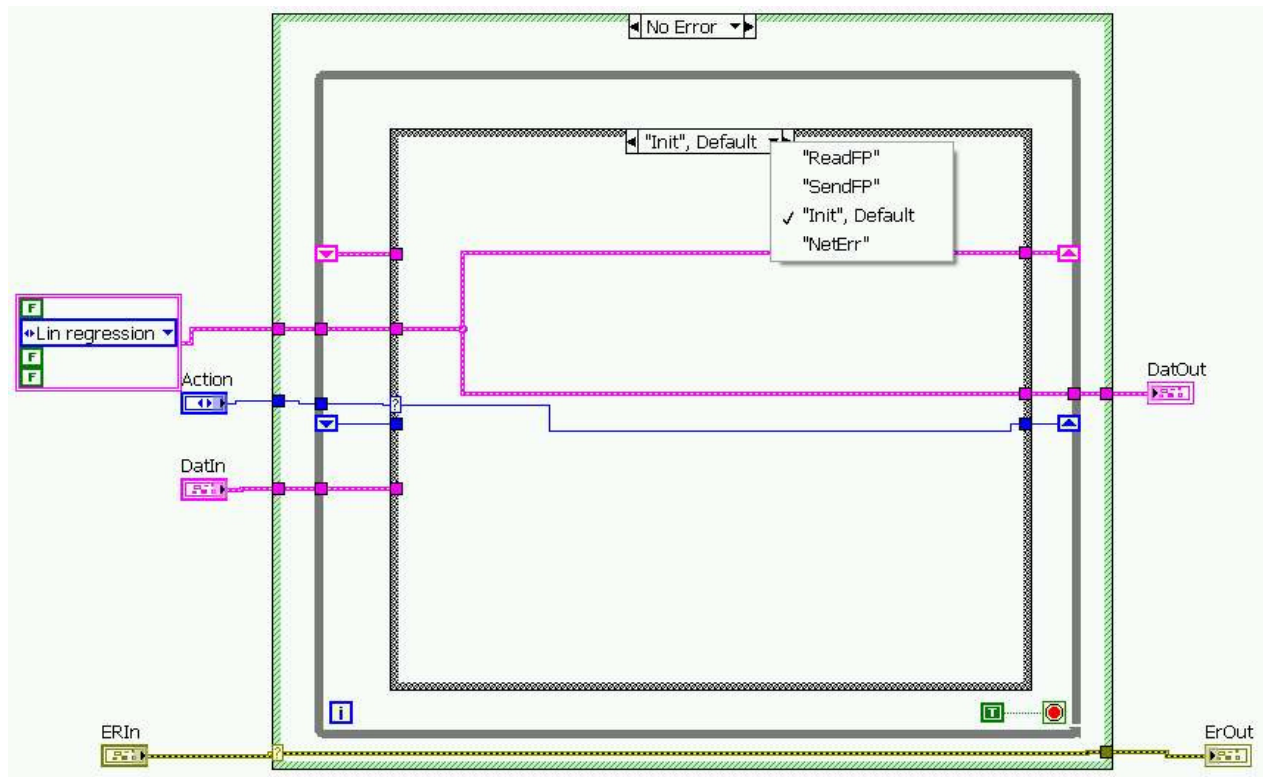


Fig. 49. Typical state machine structure.

Network vision data state machine

This state machine is accessed by three agents: communication agent, tracking agent and wheel control agent. Beside the initialization action there are three other types of action: storing data from the network, loading data for the tracking agent and loading data for the wheel control agent. The agent stores and processes the following features: distance of the object to the robot measured by the robot, position of the object of interest in the camera and the read status of those variables. The read status distinguishes new (updated data) from old/processed.

Network vision state machine

This state machine holds a functional status for the following devices: lasers, cameras image processing computer and video storing agent. It is accessed by two agents: communication agent and interface agent. It can execute four different actions: initialization,

network data storing, reading the overall state and updating the network status. A special mode just for network update was developed in order to store the update on the network faults and new connections as soon as possible.

Teleoperation system state machines

Similar concept of state machines was applied to in the development of the teleoperation framework. The agents are exchanging data and saving its state using server client architecture. The Smart Interface uses it to keep and share the status of network connections, interface changes, and image control. The Vision Control Unit uses them for camera and network management, image acquisition control and motion management.

The standard actions of the state machines are similar like I the case of the tracking fluoroscope: initialization, data reading, data writing and data processing.

Dual vision feedback mechanism in the fluoroscope system

In order to be able to test and develop a tracking robot using fluoroscopy as the main sensing source and in the same time minimizing the radiation dose a second video feedback system was used.

It required having two full sets of acquisition hardware and flexible software as well as

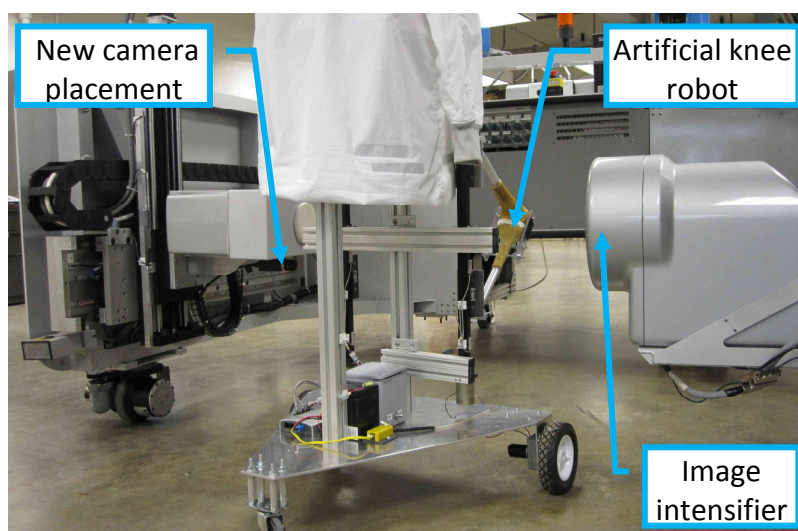


Fig. 50. TFS with the knee robot.

the interface that can switch the feedback online. Due to the X-ray image intensifier geometry, the location of both hardware sets was different (Fig. 50).

The system was equipped with an industrial grade image acquisition and an X-ray image intensifier feed. The industrial camera was located on the right side of the robot under the X-ray source. At this location the camera vertical axis is aligned with the X-ray image vertical axis. The vertical axis is shifted by roughly four inches. This allows for testing the robot under geometrical conditions that are the closest to the real ones. It includes the interaction between the robot and human, relative position and motion limits as well as safety concerns. The X-ray moving components should be directly on both sides on the subject body not shifted forward/backward or up and down. Either of those situations could introduce a risk of the patient being hit by the robot.

Different locations of the image sensors required a use of different position sensors. The X-ray system uses the position sensors (vertical and horizontal) on the left side and the training video sensor side is using the right side position sensors. The computer system was designed to be able to acquire data from all the sensors at the same time, however only one set is used for servoing at a time. This includes two camera interfaces and four positioning interfaces.

Another aspect is being able to calibrate and establish dual communication with both sets of sensors. In the case of position sensors the geometrical calibration happens on both sides simultaneously. Once the both side are aligned the operator calibrates both sensors at the same time. The camera communication is done by an independent communication agent. It continuously monitors for a lost connection and, if necessary, attempts to establish it back without interruptions to other parts of the system. In a case of both feedback sensors presence, the communication is created simultaneously.

The modular nature of the system allows for online change of the tracking sensor. Assuming that both sensors are present and functioning, the system can switch between them simply by using a GUI button. The servoing agents were designed to accommodate this change

by using proper position feedback and hardware constants. The image acquisition and processing agent was developed to switch online as well. The video transfer agent was designed in such a way that it can function the same way no matter what input it receives.

The acquisition, image processing and servoing parameters were saved in the code while the status of the feedback choice is shared by the agents to choose proper actions. If it is necessary to change certain parameters like match pattern threshold or the pattern itself, it can be done remotely by changing a configuration file. This requires a full restart of the image processing software but does not need a system recompilation. The operator interface and the Main Control Unit are not affected by this and report back when the remote agents are ready to function.

Human control of the TFS test stand

TFS is a bio-robot of a new kind, the first prototype, and requires careful operation and planning. The minimum number of people required to turn the robot on is one, however, when interacting with living subjects it was established that it requires 2 or three people to operate it.

The first person is the main console operator. His/hers role is to properly bring the robot into its idle state with all the functionality online. Once the robot is ready there are several steps required to begin the tracking. The operator turns on body tracking: lasers measurements acquisition and processing and motor servoing. After this step the patient's body/chest distance to the robot is controlled by the computer. The next step is to turn on video acquisition, processing, display and recording. It might be necessary to correct the patient's position at this point. In order to be able to track the joint of interest, it has to first be place in the field of view. Once that is achieved the operator can turn the knee tracking motor on and the subject can begin walking. At this point the role of the operator is to monitor the system and execute emergency shutdown if necessary or a regular shutdown procedure if the task is complete.

The second person is responsible for guiding the subject and observing the robot. His/her role is to position the patient in its default location. Once the robot is turned on, the person watches the human-robot interaction and gives instructions to the patient. In a case of direct danger to the subject or robot an emergency procedure is followed. The important part of the guide personnel is to instruct the patient before and the exercise as well as interact with the console operator.

The third person is the X-ray technician. His/her role is to set proper imaging techniques and monitor the X-ray device. The scenario in which the guide and the X-ray technician are the same person is possible.

The order of operations is as follows: instruct the patient (guide), bring the machine into idle state (console operator), take positions (all), turn body tracking on (operator), turn X-ray on (X-ray technician), turn full tracking on (operator), observe and instruct the patient walking (guide), monitor the functioning (all), stop the patient (guide), turn off the robot (X-ray technician and operator), release the patient (guide).

X-ray video display and recording

The images acquired during fluoroscopy session needed to be displayed in real-time and recorded on the Remote Host Unit. The operator had to see the scene in order to start the procedure and respond in a case of unexpected events.

Before the image could be displayed or recorded, it is processed first. In the next step it is transferred to the image transfer agent. The transfer agent repacks the image into a string form and then sends it to the Remote Host Unit and then waits for the next image. The image is sent with extra overlay information showing the position of the object of interest and its distance to the center of the image (Fig. 51).

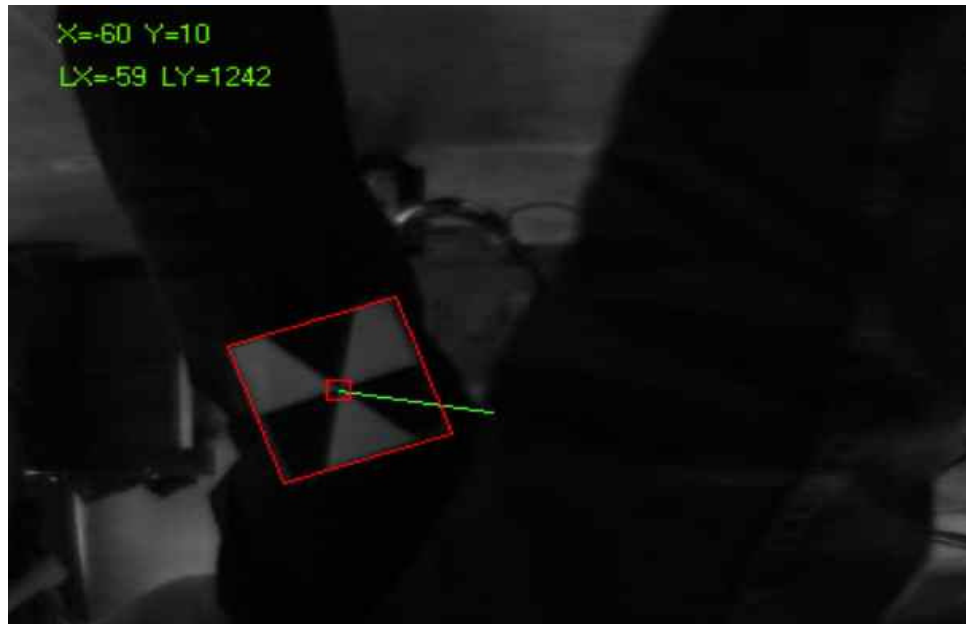


Fig. 51. Image with full overlay: Object of interest and body position.

On the host machine the image is received, unpacked and displayed with the overlay information in it. The operator can turn the recording on if required (Fig. 52). In this case the, overlay information is stored in the data part of the AVI frame and the image is recorder on the hard drive.



Fig. 52. Display and recording interface.

TFS testing procedure

Four types of test were performed considering human-machine interaction:

- Indirect human-machine interaction without radiation
- Direct human-machine interaction without radiation
- Indirect (remote) human-machine interaction with radiation
- Direct human-machine interaction with radiation

The first type involves vision servoing tests where the object of tracking was not directly attached to the human or the robot had a limited safe range of motion. These tests were performed to check the functionality of motion-image processing feedback loop and first pattern matching markers.

The second type involves direct human-machine interaction with all the motion system online, however without the use of radiation. In these experiments, overall tracking was evaluated and motion gains and necessary settings updated.

Third kind of procedures included experiments with the use of radiation but without human exposure. The second feedback was developed and tested; the system was expanded to include two different kinds of visual feedback, on the hardware, software and algorithm level.

In the fourth type of test human subjects were used to further improve the performance of the interaction with the use of radiation. Due to legal and safety limits the total radiation was restricted to two minutes per subject. This required a very precise routine to be planned for each subject depending on the motor skills, presence of implants and the current stage of development.

The following general steps in the routine of every subject were present:

- Familiarization with the robot, its characteristics and safety features
- Explanation of the experiment
- Initial acquisition of the match pattern
- Simple test of the match pattern
- The main planned exercise
- Optional exercise/data collecting
- Stopping the tests at 120 s (or before) of exposure regardless of advancement or results.

The goal of the first step was always to explain the features of the robotic platform, its design and expected behavior. This way the subject was more relaxed during the main exercises, focused on the task and providing more valuable data. Also the patient was familiarized with the main safety features as to how the system behaves in a case of malfunction.

In the next step the plan of the test was presented to the subject. Each step of the procedure was additionally repeated at the beginning of routine. At this point the patient was asked if he/she can perform those exercises. If necessary, the procedure was modified to accommodate the subject's specific condition. This additionally builds confidence in the patient while directly interacting with the robot.

During the initial image acquisition the motion servoing was not present, however the joint of interest was moving within the limits of the field of view. The recorded image sequence was used to build a set of matching patterns that were used in the further steps. This step was

short in order to minimize the used time of radiation. In a case when the matching pattern was already present in the system, this step allowed for evaluating it and making a decision of using the current or a new pattern.

In the following step, the subject was performing a simple task, such as making a step forward or bending the joint. This allowed for estimating the quality of the created match pattern and updating it if necessary. This step was executed with the robot running in automatic mode with the vision feedback loop online.

In the main task, the patient was performing the major exercise that was providing the most data and giving direction for next tests and improvements. Depending on the stage of the research, this consisted of a full step, stair rise or several continuous steps.

Additionally, an optional task was always planned if the allotted exposure time allowed. This could anything from a simple static motion test to advanced motion.

All tests were stopped at the 120s mark or before whether they were completed successfully or not.

Analysis of the X-ray video feedback

Each joint being imaged is slightly different than others due to age, size, gender, tissue characteristics, presence of implants or surgical operations. It is preferred that all possible steps are taken in order to minimize the radiation exposure.

Two kinds of analysis are needed to be performed: one at the beginning of the test and after the test. In the first kind a fluoroscopy video sample is taken and analyzed to determine if current system settings are proper for the joints of interest. In the second kind, the entire recorded test is analyzed one more time in order to improve future performance and/or trace possible reasons of poor test quality.

The developed tools have these particular features:

- Ability to load recorded X-ray images
- Ability to display presented images at its native speed and frame by frame.
- Ability to load the same match pattern and image processing tools that are used online for tracking
- Ability to manually adjust image parameters used for tracking and present its results
- Ability to estimate the expected matching performance
- Ability to facilitate creation of new image pattern tracking tools.

The fluoroscopy images are recorded in AVI format without compression and the analyzing tool is able to load the video file in this format. It is important to be able to analyze the raw images as well as the processed images at its natural speed and frame by frame with the choice of frame numbers. The tool (Fig. 53) shows the total number of frames in the video clip, current frame and a control used to switch from auto to manual mode.

It is critical to be able to load and display the same tools that the real-time mechanism is using. This way, it is possible to find out reasons for artifacts in matching test and improve future matching techniques.

In a similar way, factors like matching minimum scores, regions of interests or image lookup tables can be adjusted and its effect can be analyzed without the use of radiation.

All the mentioned functionality allows for estimation of the matching performance of the used tools. In a case when this is unsatisfactory, different algorithms can be pursued without the need of making real tests. Also, along with improved performance, the creation of



Fig. 53. Image matching analysis tool.

matching patterns is facilitated. By knowing which images score high and which low, a new pattern can be introduced to algorithm, thereby improving the overall system performance.

Image tracking analysis

Real-time acquisition and storage of the fluoroscopy images is one of the key issues in this research. For proper further analysis, the images should be clear of any unnecessary graphical elements. However, for framework testing and development it is necessary to have the information about the position of the object being tracked (Fig. 54). Initially this information was being added to the image permanently but because of the mentioned reasons that had to be abandoned.

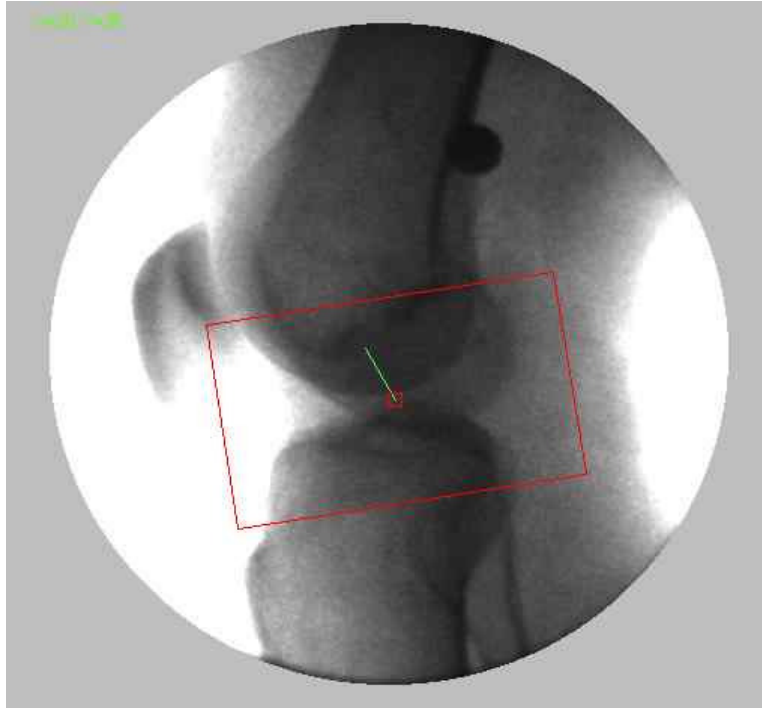


Fig. 54. Knee image with overlay information added.

One of the options under consideration was to create another file with the geometrical information and keep with the video file. Because this creates the risk of losing the track of the file to file assignment this idea was abandoned.

It was decided that the best solution would be to include the information permanently in the video file without affecting the image. The format chosen for video storage was “avi” It allows for storing extra data information for each frame of the video. The data can be of any kind as long as it is with size limits.

The video server used for storing the images is receiving the images and geometric information in a form that does not affect the image. In the next step the server packs the data in the video file, storing the geometrical information in the data segment. Then the video is saved, and once played it does not show the geometrical information. When the analysis of the tracking is necessary additional software is used to create a copy of the video with tracking data. This application reads the data segment of the video file and creates a geometrical

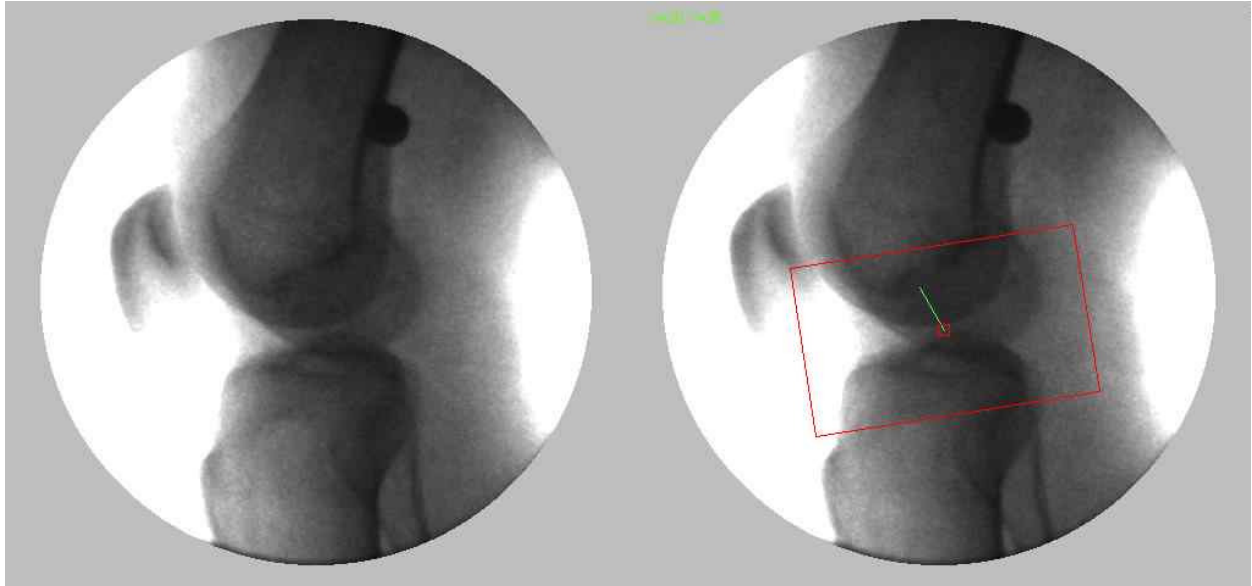


Fig. 55. X-ray image with and without the overlay information.

pattern in every image. Hence, the recorded video can be turned into videos with full tracking information (Fig 55).

Human subject body tracking in the TFS

This method was developed to keep constant position of the body with respect to the robot during a natural walk.



Fig. 56. Tracking Fluoroscope System.

The patient body tracking problem was approached using two 2D laser scanners. That allowed for calculating the center of the body position and orientation with respect to the vehicle (Fig. 56). It was decided to use two scanners to improve the resolution and quality of measurements with the person changing his/her orientation with respect to the vehicle. This way, the horizontal profile of the body is always oriented optimally toward the sensors.

Choice of measurements

Due to the fact of the required minimum distance between the object and scanner it was necessary to place it above the top of the platform - 1m above the ground. This means that for the majority of the population the lasers have to be pointed at the region between the waist and top of the head. During normal walking, a person's arms are in motion. The degree of motion may vary drastically from person to person. The top projection of the neck and head does not provide enough information to conclude the orientation. That leaves only the top part of the region between elbows and shoulders as in Fig. 57. It also requires height adjustability of the scanner for different patients. Another way to compensate for different height could be to provide adjustable vertical angle - tilt. However, during walking, the distance to the object may vary due to nonlinearities and hard-to-predict body curve. During motion, the angle at which the body is scanned is changing, the vertical curve of the body is not covered evenly. The remission is not constant and the number of scanned points can vary as well. Hence, it was decided to use two scanners symmetrically positioned with respect to the object at the same height. It is assumed that the maximum change of the orientation angle of the person is relatively small $\pm 15^\circ$. In a case of a higher value the robot should perform an emergency shutdown due to large tracking error. This also puts a limit on the dynamics of the person.

Scanner setup

Both scanners point at the person when standing in the middle of the joint tracking workspace as in Fig. 57. It was tested that the minimum data update frequency required for smooth operation of the platform is 20Hz [43]. The scanner can work at a minimum rate of 200Hz and the data can be acquired in real time only when using an Ethernet TCP/IP

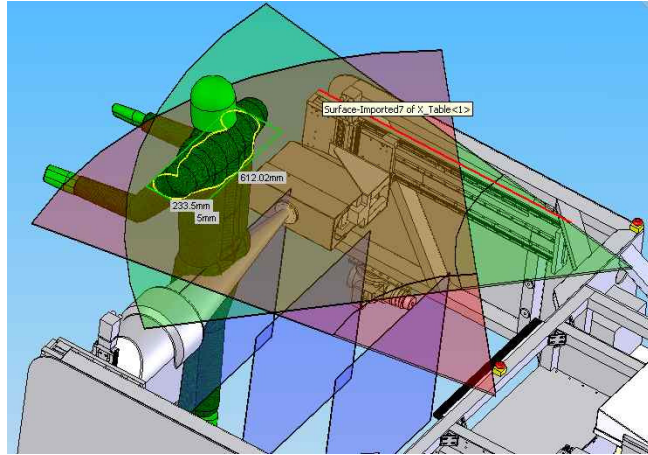


Fig. 57. An example of scanners pointing at model's shoulders, maximum distance.

connection. The other available port, RS232, cannot provide data transfer at this rate. The processing PC has to work in real time in order to receive the data in a timely manner. It was decided to perform all processing using the real-time LabView environment. Currently, the data after processing (filtering and averaging) are updated at 40Hz rate.

In order to obtain the position of the body, the data has to be filtered first and then processed. Based on the acquired data it was observed that the artifacts may have remission value up to 12-13%. The threshold was correspondingly set to be 15%. All data below this threshold are removed from the set, effectively decreasing the number of body profile points. In order to provide uniform repeatable results, each person is required to wear a white lab coat or other white top. This allows for easier separation of real data from artifacts. The internal scanner algorithms allow for filtering using the median filter and the average value. The average value calculation effectively slows down the data transfer. The median value does not cause this problem. Both algorithms are turned on and the data coming to the processing PC is already pre-filtered. After receiving the measurements, applying the threshold and calculating the Cartesian position, the curve of the body looks as in Fig. 58.

Method of finding the position and orientation

The position of the person is found by calculating the center of the area for all data

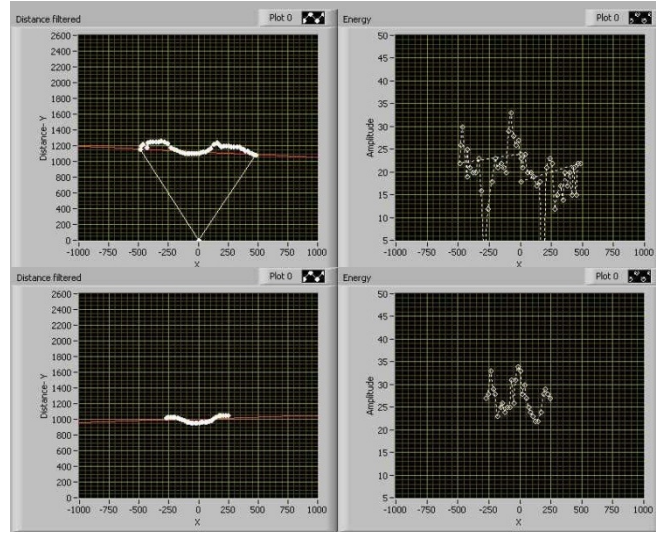


Fig. 58. Data before and after filtering. Top - unfiltered, bottom – filtered.

points. The most difficult problem is to obtain the change in rotation. Several algorithms were considered to obtain the orientation for tracking. The first one was linear regression. The curve of the body for the same person, seen by the scanner, can be a part of an ellipse or an arc. Even if considering a small motion of the arms just below the shoulders, the point path does not change significantly. Hence, the linear approximation would not either. The linear regression could be used to obtain orientation of the body [56].

The body curve can then be approximated using the equation:

$$y = ax + b$$

$$a = \tan(\alpha)$$

The data available for processing is x- y values of the point of the curve. Hence, a system of equations can be formed as follows:

$$\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_i & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{bmatrix}$$

where N - number of points.

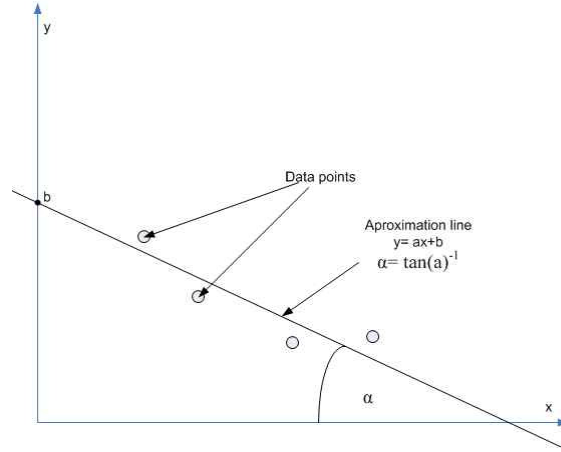


Fig. 59. Linear regression line.

Therefore, the solution can be obtained as:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_i & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix}^+ \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{bmatrix},$$

the plus sign denotes a pseudoinverse. The angle is easily obtained:

$$\alpha = \tan(a)^{-1}.$$

The results of the above calculations are shown in the Fig. 59 and 60. The continuous red line represents the approximation from eq. 1. Since the real rank of the inverted matrix is 2, no degree of freedom can be dropped. This means that in the case of outliers in the data set, the solution will be distorted. The pseudoinverse was found using singular value decomposition. An example of the resulting line is shown in Fig. 59.

Another way to obtain the orientation is to find the principal axis [57]. For a symmetrical object it coincides with the symmetry axis and its position and orientation change accordingly. In order to do that, the center of the area has to be found first, and it is done as follows:

$$x_c = \frac{\sum_{i=1,N} x_i}{N}$$

$$y_c = \frac{\sum_{i=1,N} y_i}{N}$$

where, x_c , y_c are the coordinates of the center of area. The points are treated as small unit areas for simplicity. Next all data points are expressed in a coordinate frame attached to the center of the gravity.

$$x_{inew} = x_i - x_c$$

$$y_{inew} = y_i - y_c$$

Knowing that, all moments of area are found:

$$I_x = \sum_{i=1,N} y_{inew}^2$$

$$I_y = \sum_{i=1,N} x_{inew}^2$$

$$I_{xy} = \sum_{i=1,N} x_{inew} y_{inew}$$

I_x, I_y, I_{xy} - second moments of area about axis x, y and product moment of area.

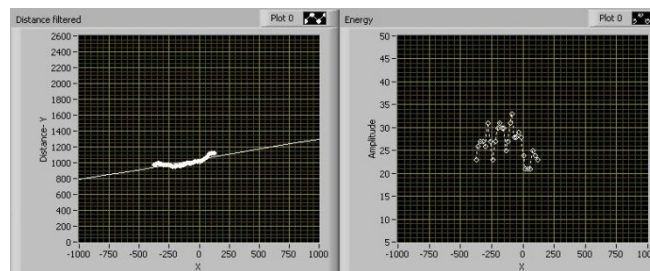


Fig. 60. Linear regression applied to data points.

The coordinate system can be rotated in order to align the coordinate axis with the principal axis. When this happens, the product moment is zero, hence the angle can be calculated as:

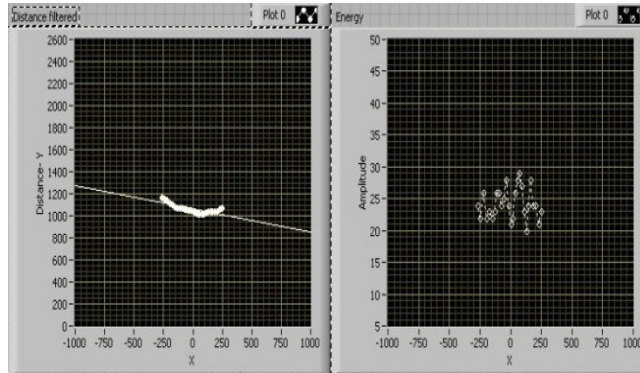


Fig. 61. Orientation line obtained through second area moment method.

$$\alpha = -\frac{1}{2} \arctan \frac{2I_{xy}}{I_x - I_y}.$$

A situation in which moments of area are equal results in dividing by zero. This can only happen in a case of a malfunction and is considered as a reason for an emergency shutdown. An example of the resulting orientation line is shown in Fig. 61.

F. Experimental Results

The following experimental results are presented in the context of the declared fundamental contributions.

Evaluation of the vision based approach of acquiring fluoroscopy images

Body Tracking

Body angle estimation

The body angle measurement was experimentally tested using a pan/tilt unit on a tripod with 360 ° motion range. After each adjustment the angular value from the algorithm and the tilt unit was recorded. Sample results are presented in the Table 5 and Figure 62. The method gives consistent and fairly accurate results in the range of – 15 to +15 °. It depends on the type of fabric and color and well the shape of the surface. The accuracy of the pan/tilt unit settings/reading was $\pm 0.5^\circ$.

Since its role was mainly to sense a coarse change of direction instead of precise measurement it can be used to adjust the orientation of the robot in the future. Currently the robot does not possess the ability to actively control its orientation.

Table 5. Orientation angle test. Values measured and calculated.

<i>Measurement</i>	<i>Linear regression calculation</i>	<i>Measurement</i>	<i>Linear regression calculation</i>	<i>Measurement</i>	<i>Linear regression calculation</i>
0	0	19	17.2	0	0
1	1	20	18	-1	-0.7
2	1.9	21	18.8	-2	-1.5
3	2.8	22	19.7	-3	-2.2
4	3.6	23	20.7	-4	-2.9
5	4.6	24	21.6	-5	-3.9
6	5.5	25	unstable	-6	-4.9
7	6.3			-7	-5.8
8	7.4			-8	-6.8
9	8.2			-9	-7.8

<i>Measurement</i>	<i>Linear regression calculation</i>	<i>Measurement</i>	<i>Linear regression calculation</i>	<i>Measurement</i>	<i>Linear regression calculation</i>
10	9.2			-10	-8.8
11	10			-11	-9.7
12	11			-12	-10.7
13	12			-13	-11.7
14	13.1			-14	12.7
15	13.9			-15	13.6
16	14.9			-16	14.5
17	15.8			-17	15.2
18	16.5			-18	unstable

Two methods were tested to obtain the orientation angle: moment of area and linear regression. Both methods were tested using the same the subject and the results were consistent. In order to do that, an interface panel was developed using LabView environment. It allowed for instant change between algorithms and visual presentation of the results. The sample change in angle is presented in the following Table 6. It shows the relative angle of the person being scanned with respect to the scanner coordinate frame obtained using two different methods. The results were obtained for different body positions. Each row shows two angle values obtained for the same body orientation.

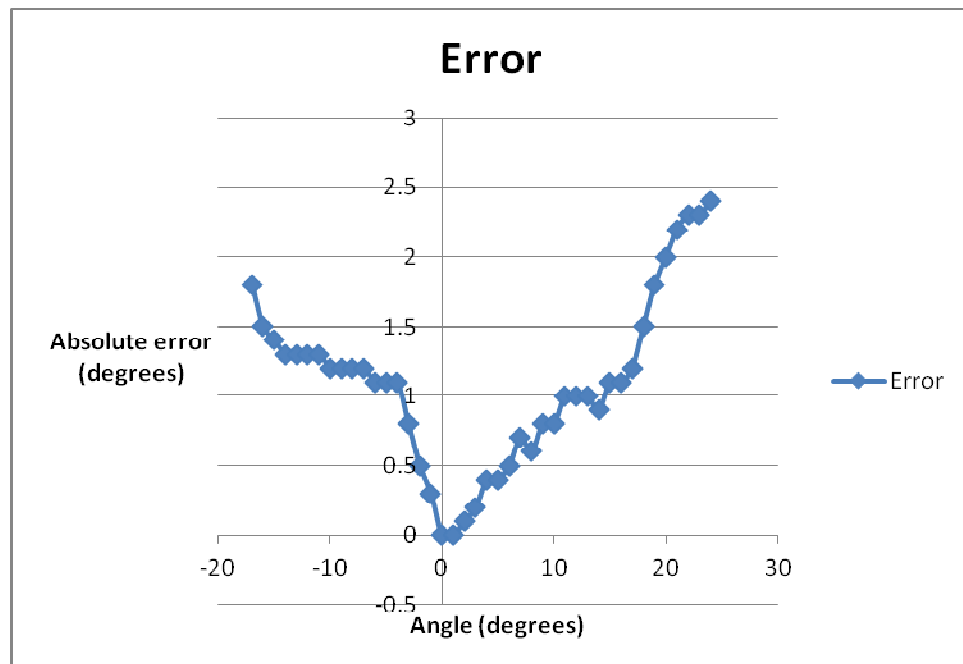


Fig. 62. Orientation estimation error.

All above algorithms were tested using a double scanner and a person naturally standing. The data were brought into a common coordinate frame and aggregated, thus giving a double resolution.

Table 6. Orientation angle obtained using two different methods. Units in degrees.

<i>Linear Regression</i>	<i>Moment of area</i>
0.5	0.6
-8	-9
-12	-12.5
13	14
23	22
-22	21

Body distance estimation

The dual laser system used in the research measures the distance values to the body oval of the chest. The actual distance is estimated based on the geometrical center of those points. Depending on the shape of the chest part of the trunk, the obtained value may vary from person to person. The important part of the algorithm is the ability to sense the change of distance. The optimal distance to the frame of the robot can be adjusted for every subject in the control interface. The measurement of the change of distance is consistent with the laser accuracy which is 3-12mm depending on the fabric, and orientation of the object. It is more than adequate to be able to actively control the distance of the robot to the walking subject. The difference between measurements made with an external device and the robot are presented in Table 7. The Obtained values had very low amplitude of oscillations. Usually it was below 1mm.

Table 7. Body distance estimation. Units in cm

<i>Offset from the setpoint</i>	<i>Offset measured by the robot</i>
0.0	0.0
3.5	3.3
17	16.9
28	27.5
36.5	36.3
-8.3	-7.4
-15	14.8
-21	-21.1
-27.6	-27.5
26	25.6
13	12.6

Active distance control

Active distance control was tested with a subject walking and the robot following. The robot measured the distance and moved its mobile platform in order to keep constant distance to the person. Under normal steady speed the robot was able to keep the person distance within 100mm of the defined set point. The distance versus time is displayed in the Fig 63. When the walking speed is slow to normal the error stays within 50mm. Each full oscillation cycle represents a one full step. The variations in subject to robot distance are directly caused by the variations of the chest to frame distance. When the subject stands still, the robot keeps its distance within 10 mm of the predefined setpoint.

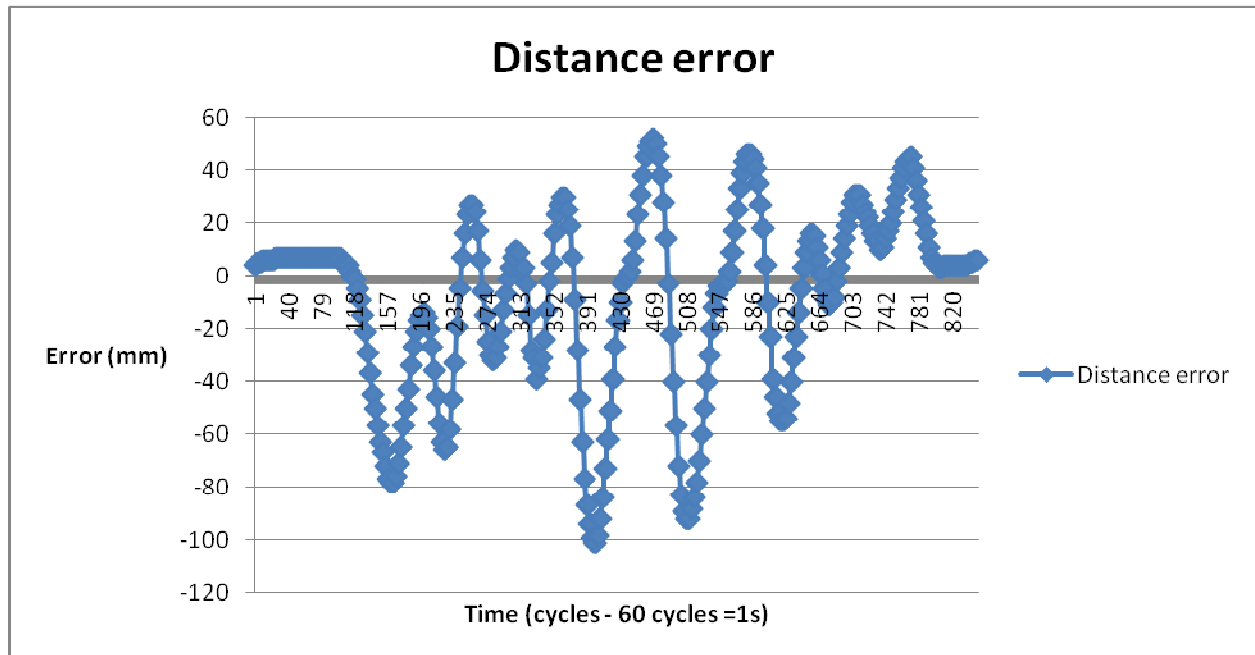


Fig. 63. Laser tracking distance error.

Knee tracking with a camera

The goal of knee tracking is to keep its defined features in the center of the view of the imaging device. This task is independent from the body tracking.

Stationary tracking

When the object of interest stays still, the servoing algorithm keeps it within 1 pixel of position error. The error value of 1 pixel translated in to position error of up to 1.2mm. The test

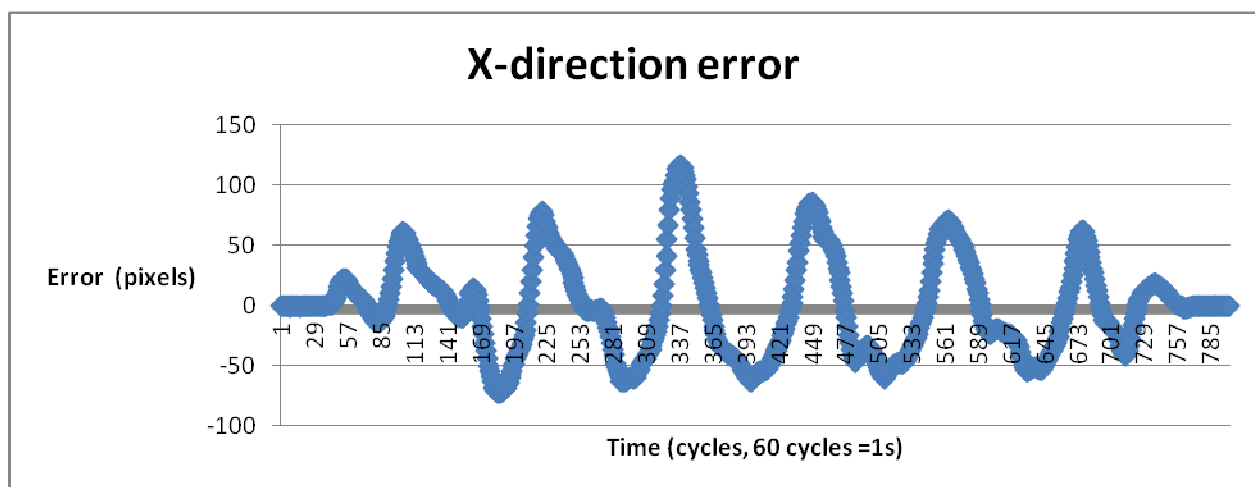


Fig. 64. Image positioning error in the direction of walking. Fast walk.

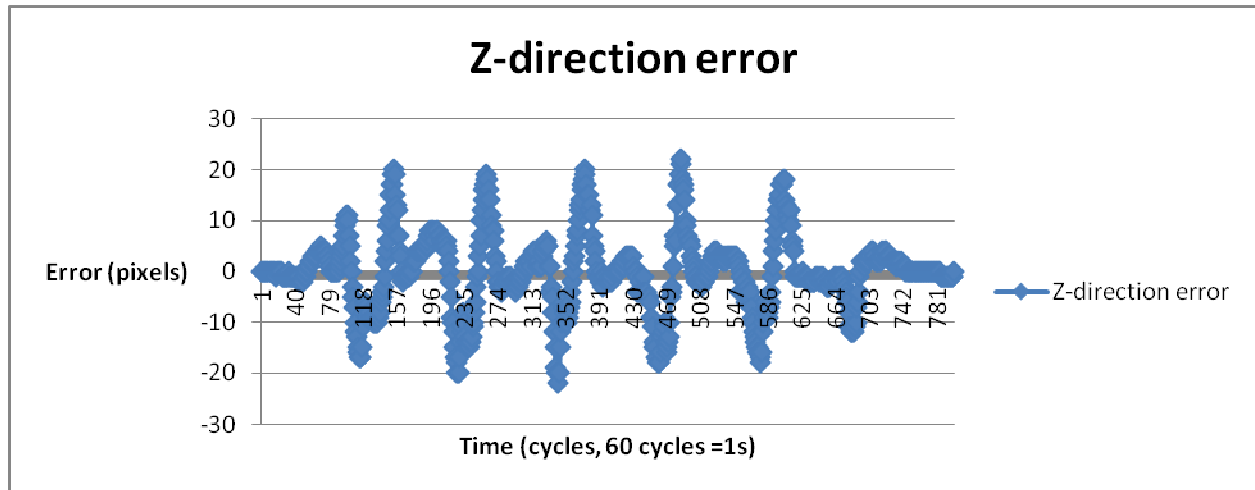


Fig. 65. Image positioning error in the vertical direction. Fast walk.

was executed by putting the object of interest within the field of view and initiating servoing. The minimum image processing value was one pixel. The sub-pixel accuracy was not used due to the computational cost and lack of advantage in improving the accuracy.

Dynamic tracking

During this test the object walked naturally and all the servoing systems were online. The overall error is a result of combined motion of the trunk, knees and the measurements of the robot. The results are presented in Fig. 63 and 64. The dynamic maximum error depends directly on the walking on the walking speed. The error stays below 100 pixels for normal to fast walking speed and below 80 pixels for slow to normal speed. The vertical error is presented in Figure 65. Due to much smaller knee displacement this error stays below 25 pixels.

Knee tracking with the use of X-ray

The following results were obtained using the X-ray unit mounted on the robot and a human subject.

Stationary tracking

When the object of interest rests, the servoing algorithm keeps it within 5 pixels off the center of the image. The resolution of the X-ray image is greater than the one of the camera

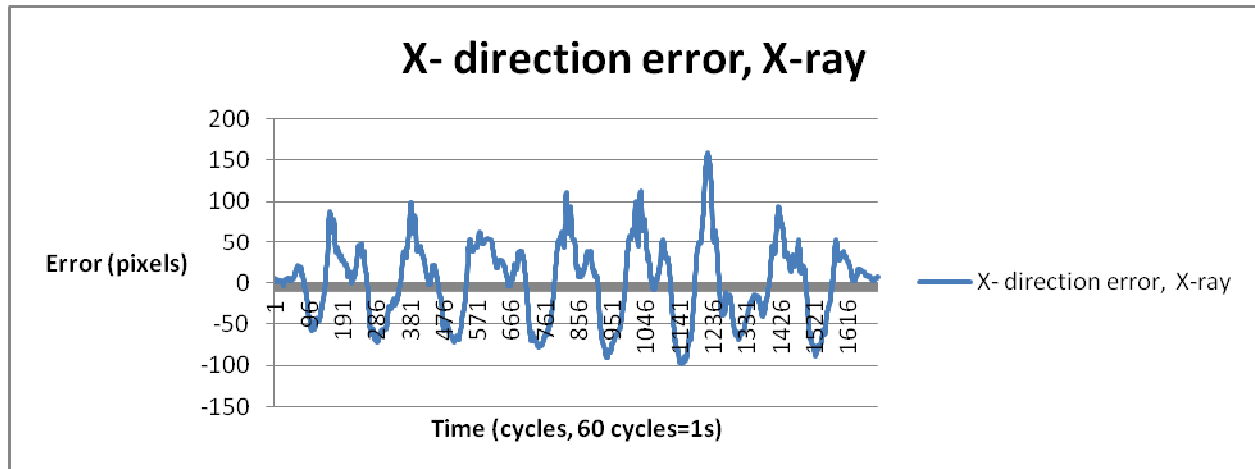


Fig. 66. Horizontal tracking error with the use of X-ray.

and the field of view is much narrower. Thus, the pixel error in the camera mode and X-ray mode cannot be compared directly. 1 pixel of error in the X-ray mode is about 0.35mm.

Dynamic tracking

Under dynamic tracking all the robot systems were online including the X-ray and wheel propulsion. The subject was walking naturally keeping his/her right knee at the front all the time. The reason for that was to avoid knee crossing and related image complications. Under normal condition for slow walking the error stays below 100 pixels in the horizontal direction and about 70 pixels for the vertical direction. This may depend on the walking style. If the person could walk with full knee crossing, the vertical error is expected to be smaller. The positioning error is presented in Fig. 66 and 67.

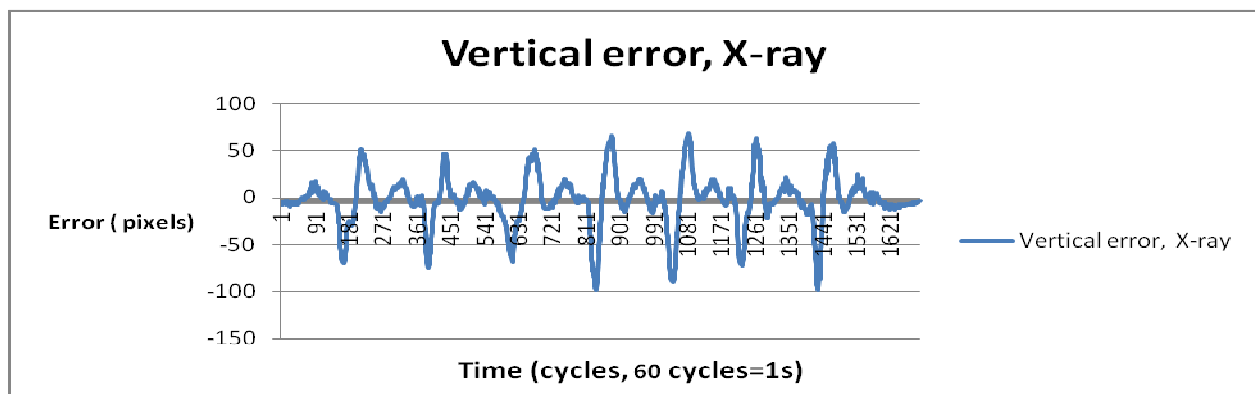


Fig. 67. Vertical error with the use of X-ray.

Framework results for tracking features of interest in human fluoroscopic images

During the experimental phase, close to 30 people took part in the X-ray evaluation. Larger number of participants were experimenting the functionality of the device in the camera mode. Hundreds of hours were devoted to test tracking of various mechanical elements.

The Tracking Fluoroscope System was developed to maximize the following criteria:

- Range of motion
- Range of activity
- Range of application
- Range of image features

Range of motion

The system allows for unlimited forward motion. The only limitation is the length of the room. The motion of the patient in the lateral direction is possible but not currently implemented. The patient may change his/her lateral position within the physical limits of the system (about 26"). The system is able to sense changes in the direction. The distance sensors are able to cover the entire cavity area and provide precise position data. The camera system is able to see up to two feet horizontally and about 18" vertically. The X-ray sensing range system is hardware limited by the 9" image intensifier. The system allows for climbing an extra 16 inches vertically in addition to normal walking height variation.

Range of activity

During the tests the following activities were successfully executed: forward and backward walking, stair climbing, deep knee bending and chair rising. The system allows for other activities as long as the features of interest are not occluded and within the sensing range. The framework was capable of working with subjects of ranging height, weight, age, gender and natural or implanted joints.

Range of applications

The current framework was intended for knee tracking however it could be extended without extensive changes to track ankles or hips. The knee tracking was executed with both the camera and X-ray system. In addition, it was possible to track a moving robot with an artificial knee. Other mechanical objects were tested as well. The major limitation is the physical size and sensing range.

Range of image features

Several kinds of objects were tracked. They included: knees, artificial markers, metal parts and artificial knees. The framework allows for tracking both still and moving objects. The features of interest can be quickly changed. The change from predefined X-ray feature and camera feature is instantaneous. In addition, due to the modular system nature different image processing techniques can be applied in a short time without affecting other parts of the system. Currently the main workhorse of the image system is template matching, however the framework can use other techniques like morphology, shape matching or complex custom geometry feature extraction.

Additional framework functionality

The system allows for dual visual feedback and instant switching between the inputs. Both inputs, the camera and X-ray systems are fully integrated with the motion control.

The framework allows for real-time image display and storing for further analysis. Additional tools for image analysis and post processing were developed in order to improve the functionality of the framework.

Data flow

The proposed framework resulted in the following paradigm of image and tracking data management.

The acquisition of the image takes place continuously whether the processing is performed or not (Fig. 68). Once the operator turns the video tracking on, the processing agent waits for the current frame to be transferred to the memory, then it receives its own copy. The next step is to process the image and find the pattern or feature. When the object of interest is located in the image the coordinate transformation is executed in order to provide the position with respect to the center of the image. The orientation of the object is acquired, as well, but it has no use for tracking since its goal is to keep the object in the center of the image regardless of the orientation. If for any reason the object is not found, the assumed position is the center of the image. It tells the motion controller to go to the last requested position and stop there. Once the feature is found again, the motion continues without any delays. After the position of the feature is established, it is sent to the MCU for servoing. All these steps are executed in just a few milliseconds so the next cycle starts right after the very next image frame is ready.

The vision-related motion control flow diagram is shown in Fig. 69. The cycle starts from receiving the feature position. In the next step, the associated motor position is read. The

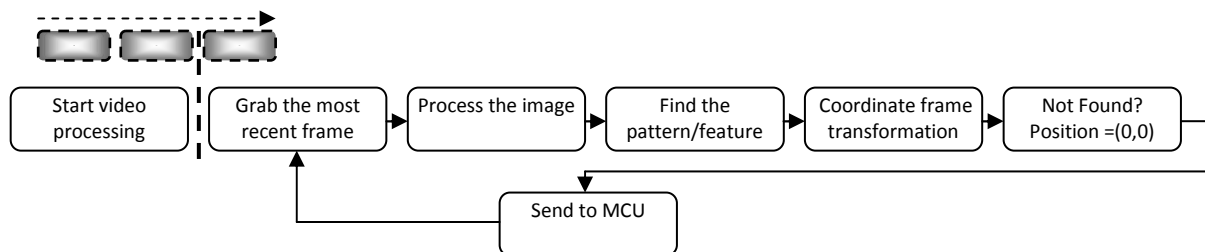


Fig. 68. Feature tracking data flow diagram

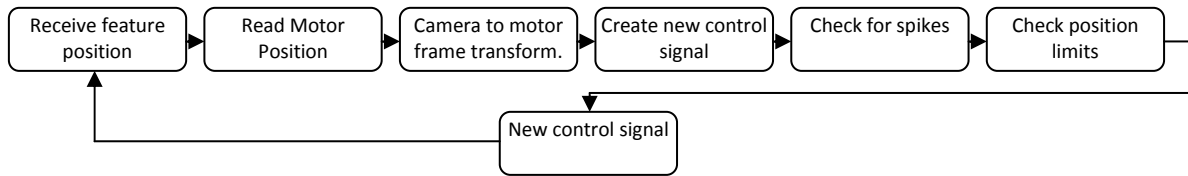


Fig. 69. Vision servoing data flow diagram.

position of the feature comes as a distance from the center of the image in pixel units. This has to be properly translated into common units and compared with the current position. Assuming that the patient's walking is a continuous sequential motion, it is not expected that one position should differ drastically from another, so maximum desired change of position is limited to a predefined value. In the following step, the new desired position is checked to be within the motion limits of the motor and a new control signal is sent to the servodrive.

The transfer of the image to the operator's console takes place in parallel to other threads (Fig. 69). Once the new image is ready and processed, extra information is added to the image. This extra information presents the position of the feature with respect to the center of the image as a line and numbers. It helps the operator to quickly quantify the tracking and diagnose problems (Fig. 70). The extra information can be removed from the image before saving. In the next step, the image data is packed and sent to the operator's console. The execution of the thread is suspended until a new image is ready.

At the operator's console, the data is being received, unpacked and displayed on the screen. If needed, it can be saved as a video clip. Once the cycle is completed it starts from the beginning by receiving the data.

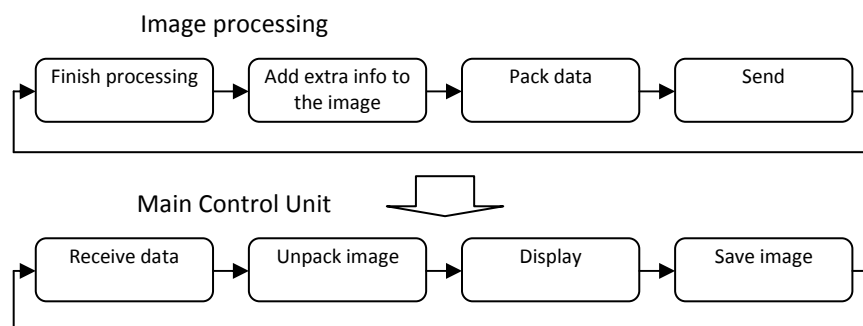


Fig. 70. Image and data storage diagram.

Teleoperation framework results

The teleoperation framework was successfully tested with the following input devices:

- Whole Arm Manipulator
- External behavior based agent
- Vision based agent
- Manual interface (keyboard/touch screen/visual)

Whole Arm Manipulator

The experimental tests included the following types of operations.

- Bolting/unbolting
- Cutting with a reciprocating saw
- Cutting with a use of an external agent
- *Picking up objects of different shapes like concrete blocks and pipes.*

The tooling operations were executed with an experimental three finger end-effector. Operations were executed manually by the operator using the WAM.

Behavioral agent

In this part The High Level Controller (part of the framework) was used as the workhorse for the behavioral agent [58]. Several automated tooling actions were executed with the use of HLC. The creation of the behavioral agent was not part of this research, but the created framework served as the test field for its performance. An arbitrary kind of external agent with

its own set of sensing devices can be used to control the arm ,and the HLC can be provide the interface to control the arm.

Vision based agent

Development of this agent required introduction of a vision sensor to the system and creating proper processing and control strategies. The agent allowed for display, sensing and tracking of various objects in the field of view. The goal was to be able to automatically track and position the object of interest in the center of the view. Once the object is placed in the desired region, further actions are possible depending on the task and tool availability.

In the unstructured environment it is very hard to create futures of interest a priori. Very often this can be the first failure of a conventional approach. In the next step the arm, the object or both can move. In a difficult environment this can dramatically change the visual cues. The illumination may change, so can the reflection/glare of the object. Without redefining the features the object can be temporary or permanently lost. This method gives a chance to take the task all the way to the tooling phase. Depending on the system and the task, other techniques can be used to finish the task. In certain situations it still can be a vision module designed specifically for handling tooling tasks, in other cases it can be for example a force feedback head. This method does not solve all the problems of the task but certainly can be useful in certain aspects of it. It is believed that it can be further explored for certain cases .of tooling tasks.

The major successful features of this mode were:

- Ability to function without a predefined set of features (library)
- The operator shows the object once
- The agent learns the object online, and, if it changes (i.e. due to reflections or changing illumination), it learns new features automatically.

- The operator has the ability to see the process of learning in real-time and react if it is not valid.
- The agent automatically positions the arm in a desired configuration with respect to the object of interest.
- The object of interest are not artificially created, positioned or illuminated.

Creation of a library of predefined features is a long and tedious task, however very useful; it was decided to study the case without a predefined library. Before every new task the system starts with blank memory.

The operator was responsible for pointing the object of interest and at this point his/her role was limited to observing and securing the positioning part.

Once the object was shown and extracted the agent was responsible for its tracking and learning. The agent automatically decided if the set of features was good enough for continuous tracking. When the measure of similarity of the features was below a predefined threshold, the agent was learned additional features and tracked a wider set of features. The learning and tracking process was simultaneous and transparent to the operator. The system was design to track up to four set of features in parallel at the same time and up to 16 features sets sequentially total. This was achieved without perceivable loss of performance to the operator. Introduction of state of the art computing hardware can boost up the performance even higher. In addition, the learning process was displayed to the operator and under his/her control. The operator is aware of the level of certainty the agent is using to track the object.

In this research, the set goal was to bring the object of interest into the center of the image, but in a general case this easily can be expanded to perform more complicated task as mentioned in the other modes.

The main advantage of this approach was that the environment was not artificially

modified or altered. This was tried on a set of standard industrial construction elements like pipes, brackets pieces of wood or aluminum. The same applies to the illumination. For example while the arm was moving the reflection of the object was changing but system was able to learn new features before it lost the track of the object.

Manual control interface

The system was equipped with a double control interface. The HLC is equipped with one, with limited number of features. The main interface was designed to be controlled from a touchscreen as well as with mouse control.

The major features allowed control of the arm in Cartesian coordinate frame and joint control. The speed of the arm in Cartesian mode can be controlled as well as Cartesian axis constraints, limitations. The arm can be forced to move for example only along certain axis or plane. In addition it also allows for controlling the parallel jaw gripper.

Hybrid modes

The hybrid modes were not tested in this research but the framework after small modifications can allow for mixed control. For example, the operator can control the orientation and the vision agent controls the position of the feature of interest.. With the stable vision agent numerous scenarios of control strategies are now open.

G. Summary and Future Work

Summary

This research focused on two general cases of vision based servoing: a bio-robot interacting with human subjects and vision based teleportation techniques.

Two complex framework systems were developed and extensive experimental tests were conducted. The bio robot system was tested with close to 30 human volunteers of different height, weight, age and gender. The teleoperation system was tested on numerous random objects of unconstructed environment.

In both cases, specialized techniques of control were developed and the feasibility of the proposed solutions was established.

Review of Contributions

Three major contributions were introduced in this research. The concept of acquiring fluoroscopy images through the creation of a generalized framework for tracking features of interest in real-time images during natural human motion was explored and evaluated through a broad range of experimental tests. It was established that it is feasible to acquire X-ray images, track the joint of interest and display and store data (for off-line clinical studies) in real-time using a novel, tetherless, fully automatic and wireless bio-robot of a new kind. This capability will allow for more detailed analysis of human joint dynamics including those involving total joint arthroplasty.

A framework for tracking features of interest in human fluoroscopic images was developed. Its generality with regard to the input and output devices, types of tracking features and range of applications was successfully tested. These results can be used to further explore and expand the area of bio-robotic fluoro-based analysis of human joints during natural

movement.

An experimental framework for vision-based teleoperation was created. A multi input control environment was developed to allow for execution of a variety of tasks. Experimental strategies of vision-based control were proposed and tested. The experiments established that is it feasible to use vision input in unstructured environments for robot control.

Future work

There are several areas that can be considered in future development on both systems described in the research.

Currently, the orientation of the subject is estimated based on a set of measurement points. Knowing that the shape of the body does not change dramatically, the data points can mapped to the acquired shape. The orientation can be estimated based on the shape, not the points. This way, the number of acquired data points or their distribution would have limited effect on the final result. In addition, the range of operation could be improved as well.

Another area to consider is improving sensing through introduction of pose estimating devices. Such device would provide valuable information on the coarse position of the joint of interest as well as the rest of the body. This in effect could improve robustness of the system as well as safety. By knowing the pose of the subject an emergency situation can be detected faster and more reliably.

The current techniques of locating the features of interest can be further expanded. More work should be done in the direction of developing a unified, general algorithm for locating features of the same kind. Improved and more general techniques to track the joints could result in decreasing the total time required to perform the study with a patient. This can benefit in the overall perception of the performance and usefulness of the Tracking Fluoroscope System.

Another aspect of feature tracking that can be explored deeper is automatic cue learning. A similar idea to the one presented for teleoperation can be introduced for fluoroscopy tracking.

The tracking system performance could be improved further by gaining a detailed analysis of the computer and motion system dynamics. Precise knowledge of this type can lead to development of motion prediction algorithms. This, in effect, can further improve tracking and clarity of the acquired images.

Positive results of this work can also lead to the development of methods for ankles and hip tracking. The experience gained can be used for further expansion of the range of applications in fluoroscopy of shoulders and arms.

The Tracking Fluoroscope System could significantly improve its performance with the use of a flat panel X-ray imaging. The system would benefit in two areas: increased field of view and lower, better distributed weight. The first part would allow for greater margin of error in the tracking while obtained perfectly good images. Greater margin of error allows for improved dynamics of the subject. Smaller and better distributed mass would also improve system dynamics and, in addition to that, could increase the total execution time on a single battery charge.

The teleoperation vision framework should be further expanded in the direction of adding layers of intelligence and automatic decision making. It seem reasonable to introduce hybrid execution where the vision agent is responsible for a certain part of the arm motion while the rest is under control of the operator or other kind of sensing agent. Additional performance gain can be achieved with the use of external pointing devices like laser beams. They could be used as support for tooling actions.

E. References

- [1] J. H. Krantz, *Experiencing Sensation and Perception*: Prentice Hall 2009.
- [2] D. Coon and J. O. Mitterer, *Introduction to Psychology: Gateways to Mind and Behavior* 10th edition ed.: Wadsworth Publishing, 2003.
- [3] J. N. Herndon and W. R. Hamel, *Telerobotic technology for nuclear and space applications*, 1987.
- [4] W. R. Hamel, S. Douglass, S. Kim, P. Murray, Y. Shou, S. Sridharan, G. Zhang, S. Thayer, and R. V. Dubey, "Human Machine Cooperative Telerobotics," 2003.
- [5] H. R. Everett, E. B. Pacis, G. Kogut, N. Farrington, and S. Khurana, "Towards a warfighter's associate: Eliminating the operator control unit," in *Proceedings of SPIE - The International Society for Optical Engineering*, 2004, p. 267.
- [6] "Dual arm work platform teleoperated robotics system," Argonne National Laboratory-East, IL DOE/EM-0389, 1998.
- [7] H. Humphrey, W. R. Hamel, N. Andrzej, and J. Park, "Large-Scale Dexterous End-Effector Manipulation," in *ANS International Joint Topical Meeting: Sharing Solutions for Emergencies and Hazardous Environments* Albuquerque, NM, 2008.
- [8] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *Ieee Transactions on Robotics and Automation*, vol. 12, pp. 651-670, Oct 1996.
- [9] G. Allibert, E. Courtial, and F. Chaumette, "Predictive Control for Constrained Image-Based Visual Servoing," *Robotics, IEEE Transactions on*, vol. 26, pp. 933-939.
- [10] A. La Tuan and S. Jae-Bok, "Object tracking and visual servoing using features computed from local feature descriptor," in *Control Automation and Systems (ICCAS), 2010 International Conference on*, pp. 1044-1048.

- [11] C. Demant, G. Schmidt, B. Streicher-Abel, M. Strick, and P. Waszkewitz, *Industrial image processing : visual quality control in manufacturing*. Berlin [u.a.]: Springer, 1999.
- [12] N. Instruments, "IMAQ Vision Concepts Manual," June 2003 ed: National Instruments, 2003.
- [13] P. J. Burt, J. R. Bergen, R. Hingorani, R. Kolczynski, W. A. Lee, A. Leung, J. Lubin, and H. Shvayster, "Object tracking with a moving camera," in *Visual Motion, 1989., Proceedings. Workshop on*, 1989, pp. 2-12.
- [14] C. R. del-Blanco, N. Garcia, L. Salgado, and F. Jaureguizar, "Object Tracking from Unstabilized Platforms by Particle Filtering with Embedded Camera Ego Motion," in *Advanced Video and Signal Based Surveillance, 2009. AVSS '09. Sixth IEEE International Conference on*, 2009, pp. 400-405.
- [15] Y. Rosenberg and M. Werman, "Real-time object tracking from a moving video camera: a software approach on a PC," in *Applications of Computer Vision, 1998. WACV '98. Proceedings., Fourth IEEE Workshop on*, 1998, pp. 238-239.
- [16] B. J. Nelson and P. K. Khosla, "An extendable framework for expectation-based visual servoing using environment models," in *Robotics and Automation, 1995. Proceedings., 1995 IEEE International Conference on*, 1995, pp. 184-189 vol.1.
- [17] A. H. A. Hafez and C. V. Jawahar, "Integration Framework for Improved Visual Servoing in Image and Cartesian Spaces," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, 2006, pp. 2320-2325.
- [18] C. Colombo, B. Allotta, and P. Dario, "Affine visual servoing: a framework for relative positioning with a robot," in *Robotics and Automation, 1995. Proceedings., 1995 IEEE International Conference on*, 1995, pp. 464-471 vol.1.

- [19] L. Overgaard, B. J. Nelson, and P. K. Khosla, "A multi-agent framework for grasping using visual servoing and collision avoidance," in *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, 1996, pp. 2456-2461 vol.3.
- [20] T. Martinez-Marin and T. Duckett, "Robot docking by reinforcement learning in a visual servoing framework," in *Robotics, Automation and Mechatronics, 2004 IEEE Conference on*, 2004, pp. 159-164 vol.1.
- [21] L. Ming, C. Pradalier, C. Qijun, and R. Siegwart, "A bearing-only 2D/3D-homing method under a visual servoing framework," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pp. 4062-4067.
- [22] T. Osa, C. Staub, and A. Knoll, "Framework of automatic robot surgery system using Visual servoing," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pp. 1837-1842.
- [23] A. M. Dearden and M. D. Harrison, "Impact and the design of the human-machine interface," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 12, pp. 19-25, 1997.
- [24] C. Gong, "Human-Machine Interface: Design Principles of Visual Information in Human-Machine Interface Design," in *Intelligent Human-Machine Systems and Cybernetics, 2009. IHMSC '09. International Conference on*, 2009, pp. 262-265.
- [25] W. Rong, Y. Liangjing, C. Chee-Kong, L. Kah-Bin, and S. Chang, "Intraoperative visual guidance and control interface for augmented reality robotic surgery," in *Control and Automation (ICCA), 2010 8th IEEE International Conference on*, pp. 947-952.
- [26] R. K. Begg and R. Wytch, "A television-based system for recording the kinematics of human gait," in *Engineering in Medicine and Biology Society, 1995., IEEE 17th Annual Conference*, 1995, pp. 1297-1298 vol.2.

- [27] P. Dario, E. Guglielmelli, B. Allotta, and M. C. Carrozza, "Robotics for medical applications," *Robotics & Automation Magazine, IEEE*, vol. 3, pp. 44-56, 1996.
- [28] L. Joskowicz, I. Tockus, Z. Yaniv, A. Simkin, and C. Milgrom, "Computer-Aided Image Guided Bone Fracture Surgery: Concept and Implementation," in *12th Computer Assisted Radiology and Surgery Symposium*, Tokyo, 1998.
- [29] Y. Nakajima, T. Tashiro, N. Sugano, K. Yonenobu, T. Koyama, Y. Maeda, Y. Tamura, M. Saito, S. Tamura, M. Mitsuishi, N. Sugita, I. Sakuma, T. Ochi, and Y. Matsumoto, "Fluoroscopic Bone Fragment Tracking for Surgical Navigation in Femur Fracture Reduction by Incorporating Optical Tracking of Hip Joint Rotation Center," *Biomedical Engineering, IEEE Transactions on*, vol. 54, pp. 1703-1706, 2007.
- [30] S. Delp, A. DiGoia, B. Jaramaz, T. S. Y. Tang, R. E. Ellis, and G. Fichtinger, "Fiducial Registration from a Single X-Ray Image: A New Technique for Fluoroscopic Guidance and Radiotherapy," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*. vol. 1935: Springer Berlin / Heidelberg, 2000, pp. 27-41.
- [31] L. Zollei, E. Grimson, A. Norbash, and W. Wells, "2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2001, pp. II-696-II-703 vol.2.
- [32] W. A. Hoff, R. D. Komistek, D. A. Dennis, S. Walker, E. Northcut, and K. Spargo, "Pose estimation of artificial knee implants in fluoroscopy images using a template matching technique," in *Applications of Computer Vision, 1996. WACV '96., Proceedings 3rd IEEE Workshop on*, 1996, pp. 181-186.

- [33] Z.-j. Du, M.-x. Kong, L.-x. Fu, and L.-n. Sun, "A Novel Fluoroscopy-Guided Robot-Assisted Orthopaedic Surgery System," in *Robotics and Biomimetics, 2006. ROBIO '06. IEEE International Conference on*, 2006, pp. 1622-1627.
- [34] J. Weese, G. P. Penney, P. Desmedt, T. M. Buzug, D. L. G. Hill, and D. J. Hawkes, "Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 1, pp. 284-293, 1997.
- [35] Y.-F. Wang, D. R. Uecker, and Y. Wang, "A new framework for vision-enabled and robotically assisted minimally invasive surgery," *Computerized Medical Imaging and Graphics*, vol. 22, pp. 429-437, 1998.
- [36] B. Bascle, N. Navab, M. Loser, B. Geiger, and R. Taylor, "Needle placement under X-ray fluoroscopy using perspective invariants," in *Mathematical Methods in Biomedical Image Analysis, 2000. Proceedings. IEEE Workshop on*, 2000, pp. 46-53.
- [37] J. R. J. Adler, M. J. Murphy, S. D. Chang, and S. L. Hancock, "Image-guided Robotic Radiosurgery," *Neurosurgery*, vol. 44, pp. 1299-1306, 1999.
- [38] A. Mallet, S. Lacroix, and L. Gallo, "Position estimation in outdoor environments using pixel tracking and stereovision," in *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on*, 2000, pp. 3519-3524 vol.4.
- [39] B. W. Minten, R. R. Murphy, J. Hyams, and M. Micire, "Low-order-complexity vision-based docking," *Robotics and Automation, IEEE Transactions on*, vol. 17, pp. 922-930, 2001.

- [40] T. Fong, C. Thorpe, and C. Baur, "Advanced Interfaces for Vehicle Teleoperation: Collaborative Control, Sensor Fusion Displays, and Remote Driving Tools," *Auton. Robots*, vol. 11, pp. 77-85, 2001.
- [41] P. Chung Hyuk and A. M. Howard, "Vision-based force guidance for improved human performance in a teleoperative manipulation system," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, 2007, pp. 2126-2131.
- [42] W. R. Hamel, D. Komistek Richard, and D. A. Douglas, "A Tracking Robot Concept for Human Musculoskeletal Diagnostics," in *The Third IARP - IEEE/RAS - EURON Joint Workshop on Technical Challenges for Dependable Robots in Human Environments*, 2004.
- [43] G. R. Preliasco, "Motion Control for a Tracking Fluoroscope System," in *Mechanical Aerospace Biomedical Engineering Department*. vol. Master of Science Thesis Knoxville, TN: The University of Tennessee, 2005.
- [44] R. B. Cunningham, "Mechanical Design and Integration of a Tracking Fluoroscope System," in *Meachnical Aerospace and Biomedical Engineering*. vol. Master of Science Knoxville, Tn: The University of Tennessee, 2005.
- [45] T. Oberg, A. Karsznia, and K. Oberg, "Basic Gait Parameters - Reference Data for Normal Subjects, 10-79 Years of Age," *Journal of Rehabilitation Research and Development*, vol. 30, pp. 210-223, 1993.
- [46] S. Verma, J. Kofman, and W. Xianghai, "Application of markerless image-based arm tracking to robot-manipulator teleoperation," in *Computer and Robot Vision, 2004. Proceedings. First Canadian Conference on*, 2004, pp. 201-208.
- [47] R. Cipolla and N. J. Hollinghurst, "Human-robot interface by pointing with uncalibrated stereo vision," *Image and Vision Computing*, vol. 14, pp. 171-178, Apr 1996.

- [48] Z. Fangwen, G. Zhenbang, and C. Ze, "Vision-based 3D Shape Reconstructure of a Non-Rigid Object," in *Mechatronics and Automation, Proceedings of the 2006 IEEE International Conference on*, 2006, pp. 1769-1773.
- [49] D. Kragic and H. I. Christensen, "Survey on visual servoing for manipulation," *Computational Vision and Active Perception Laboratory, Fiskartorpsv*, vol. 15, 2002.
- [50] D. Kragic and M. Vincze, "Vision for robotics," *Foundations and Trends in Robotics*, vol. 1, pp. 1-78, 2009.
- [51] Z. Kalal, J. Matas, and K. Mikolajczyk, "Online learning of robust object detectors during unstable tracking," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, 2009, pp. 1417-1424.
- [52] A. Nycz and R. W. Hamel, "Whole Arm manipulator as a universal master controller," in *ANS International Joint Topical Meeting: Sharing Solutions for Emergencies and Hazardous Environments*, Albuquerque, NM, 2008.
- [53] R. Zhou, W. R. Hamel, A. S. Hariharan, and M. W. Noakes, "Using the WAM as a Master Controller," in *1st Joint Emergency Preparedness and Response/Robotic and Remote Systems Topical Meeting* Salt Lake City, Utah, 2006.
- [54] M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot modeling and control*. Hoboken, NJ: John Wiley & Sons, 2006.
- [55] J. Nagle, "Congestion Control in IP/TCP Internetworks." vol. RFC 896, 1984.
- [56] J. Cohen, *Applied multiple regression/correlation analysis for the behavioral sciences*, 3rd ed. Mahwah, N.J.: L. Erlbaum Associates, 2003.

- [57] W. D. Pilkey, *Analysis and design of elastic beams : computational methods*. New York: Wiley, 2002.

- [58] M. W. Noakes, "Telerobotic Sensor-based Tool Control Derived From Behavior-based Robotics Concepts," in *Department of Mechanical, Aerospace, and Biomedical Engineering*. vol. PhD Knoxville: University of Tennessee 2011.

Appendices

Main Control Unit sample code

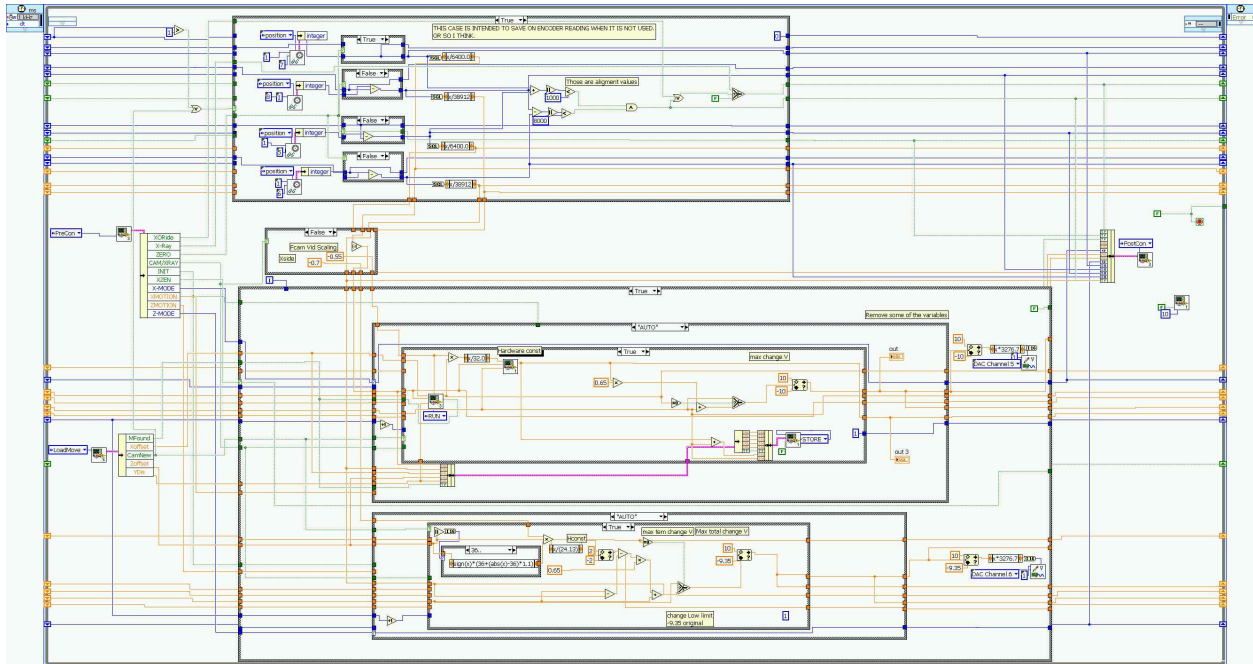


Fig. 71. The Main Motion Controller.

The main motion controller controls the horizontal and vertical drives. It adjusts the analog output of the computer according to the signal obtained from Image processing computer.

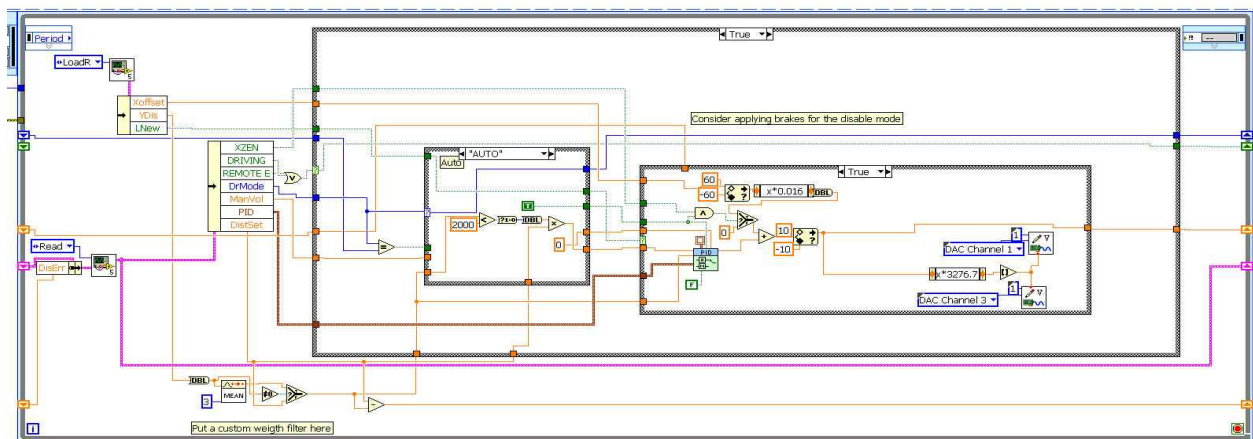


Fig. 72. Wheel motion controller.

The wheel motion controller based on the laser range signal keeps the subject at a constant distance with respect to the platform of the robot.

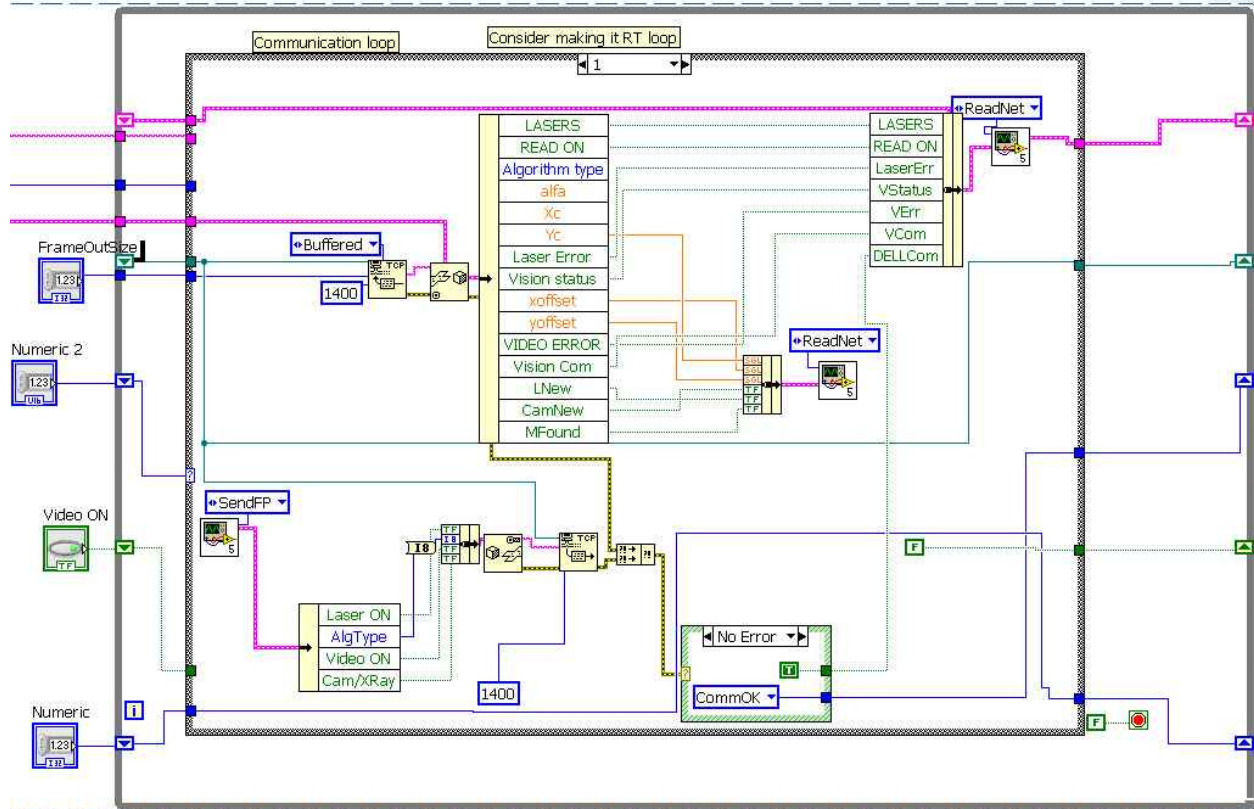


Fig. 73. Communication loop.

The communication loop transfers the commands and system variables to and from the Main Control Unit.

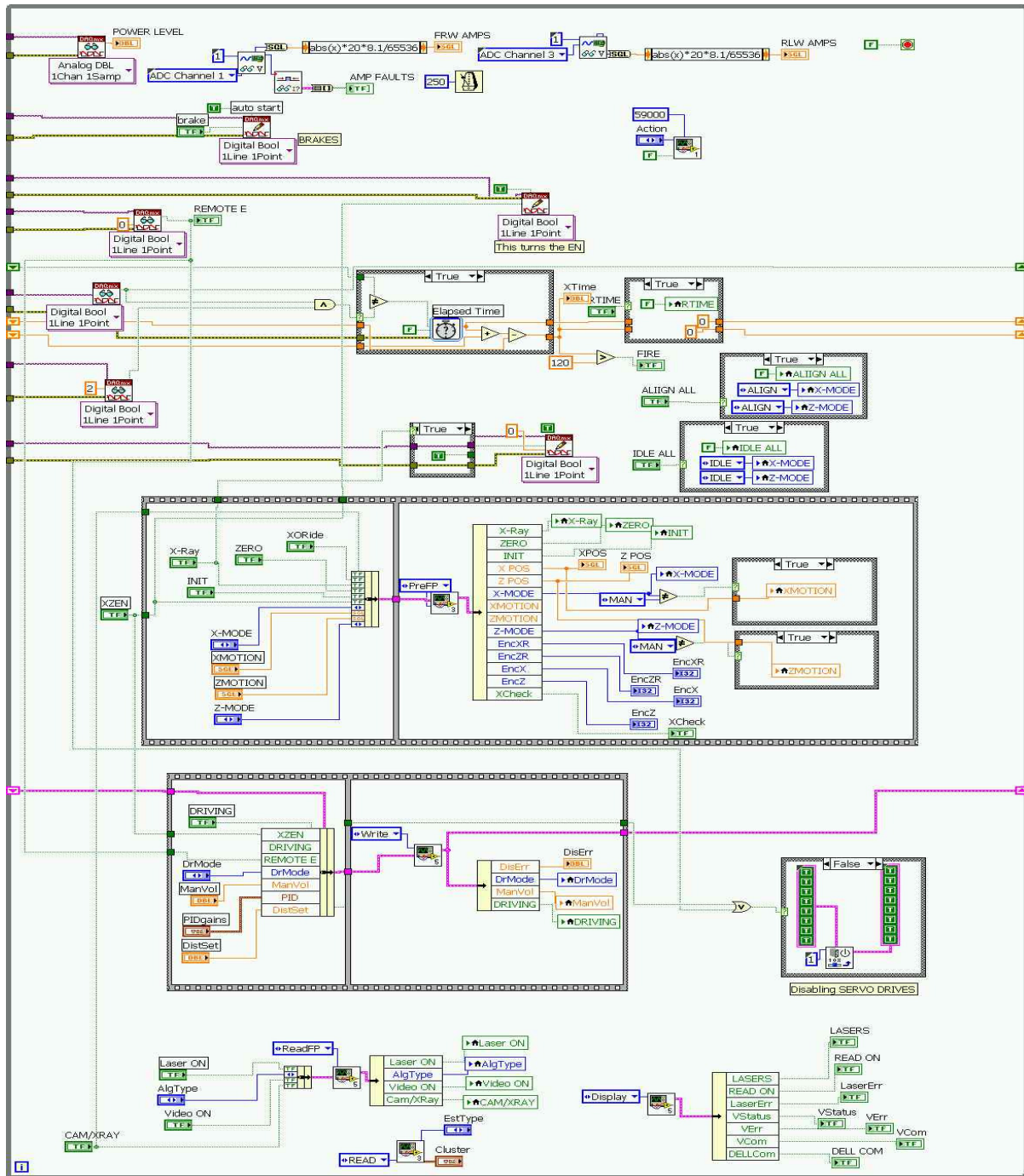


Fig. 74. Input/output and interface loop.

The input/output loop provides the connection and interface between the communication loop and the motion controller.

Image processing unit sample code

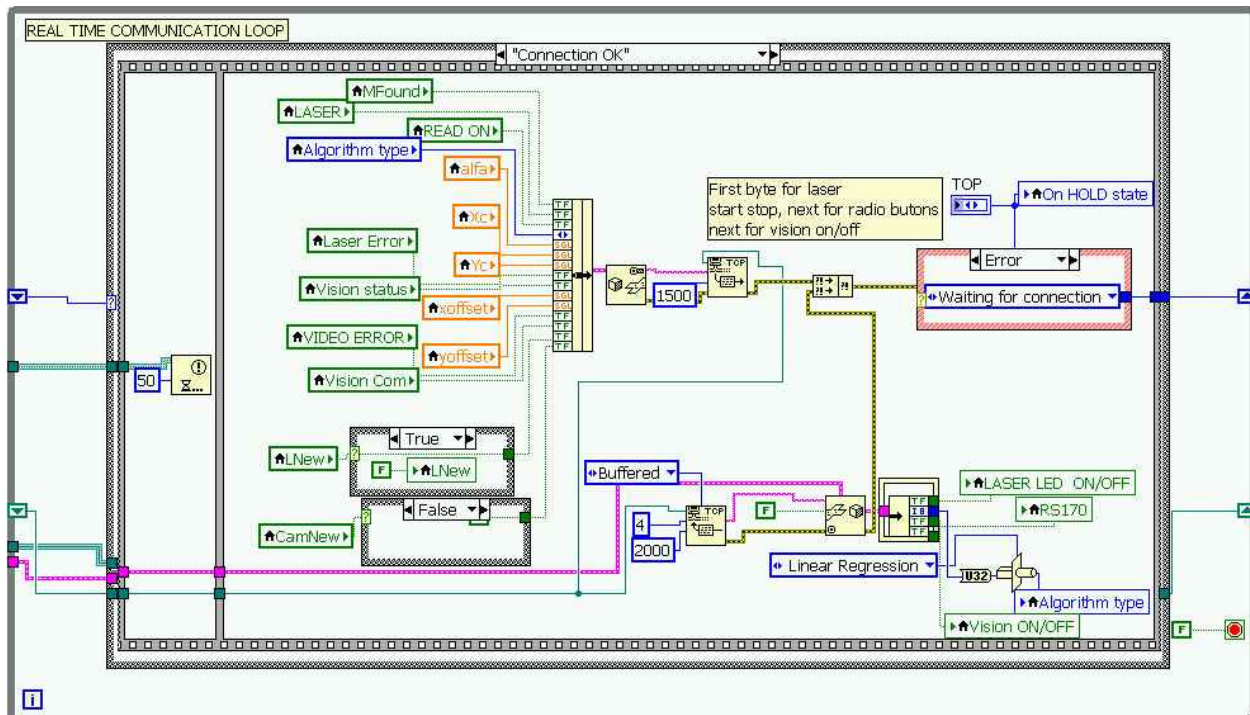


Fig. 75. Vision communication loop.

The vision communication loop exchanges the data acquired through the laser and image processing. It also allows for turning the lasers and image acquisition on and off.

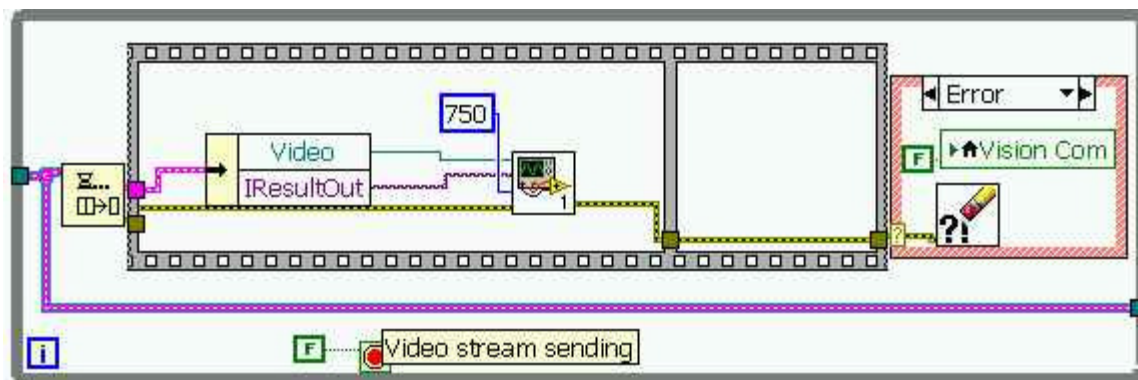


Fig. 76. Image transfer loop.

The image transfer loop is an independent process that allows for sending the images without interrupting the image processing or acquisition.

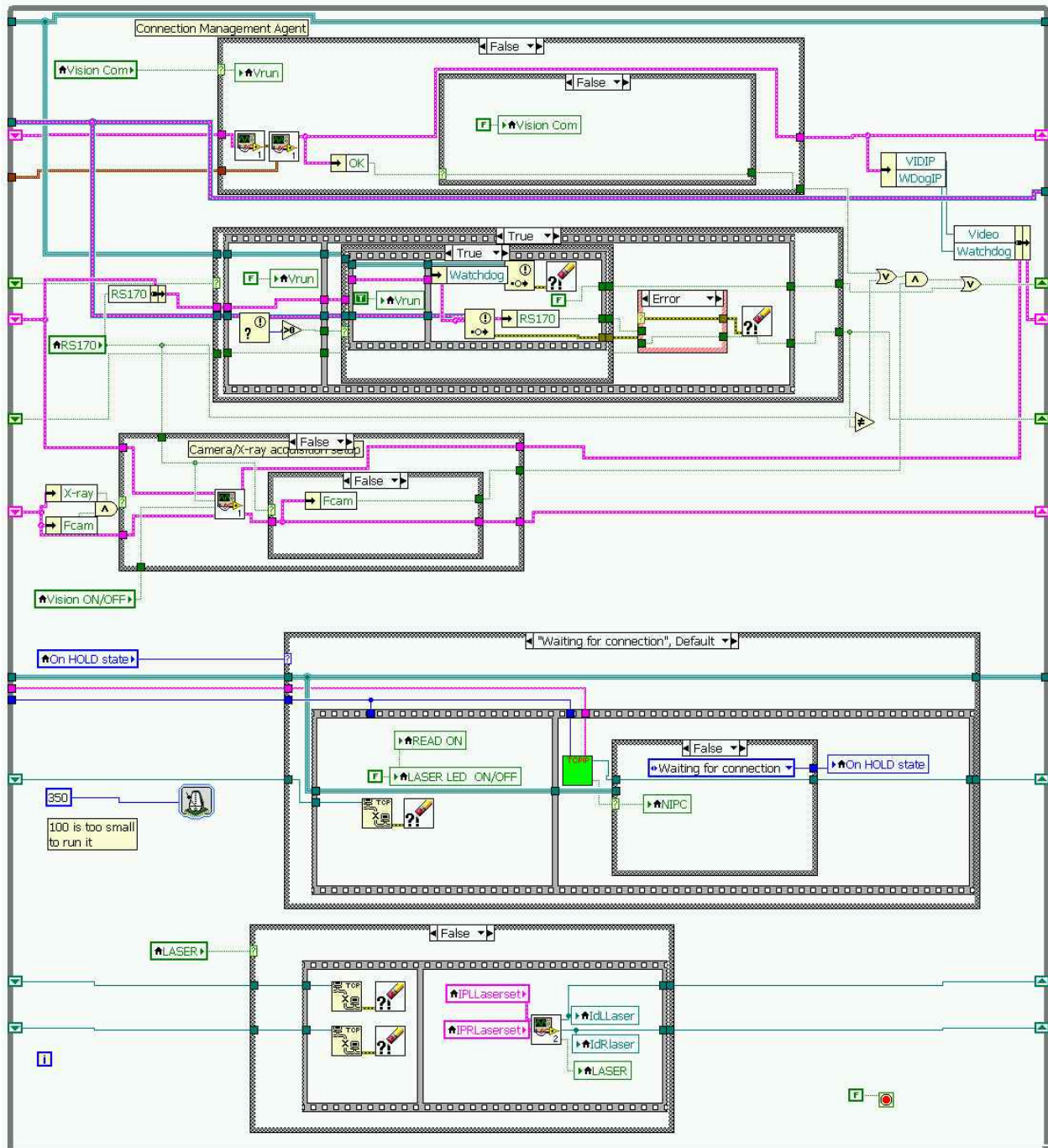


Fig. 77. Connection manager loop.

The connection manager loop is a low frequency process that automatically establishes TCP/IP connections and camera initializations. It also stores the camera specific settings.

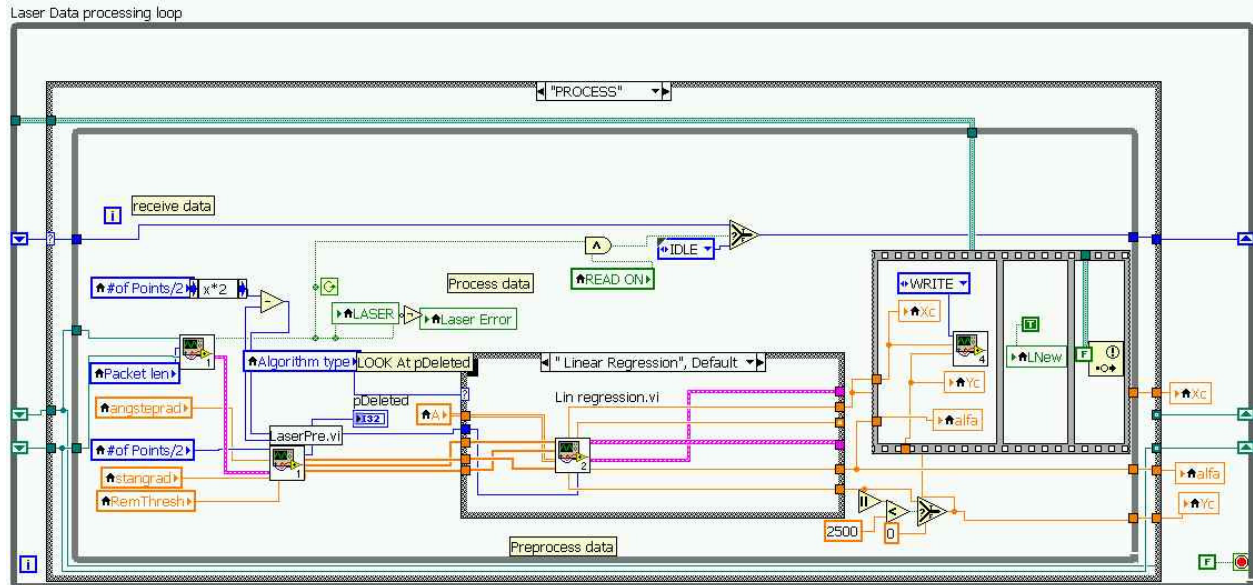


Fig. 78. Laser data processing loop.

The laser data processing loop, receives measurements information from two laser scanners, processes it and calculates the position of the subject.

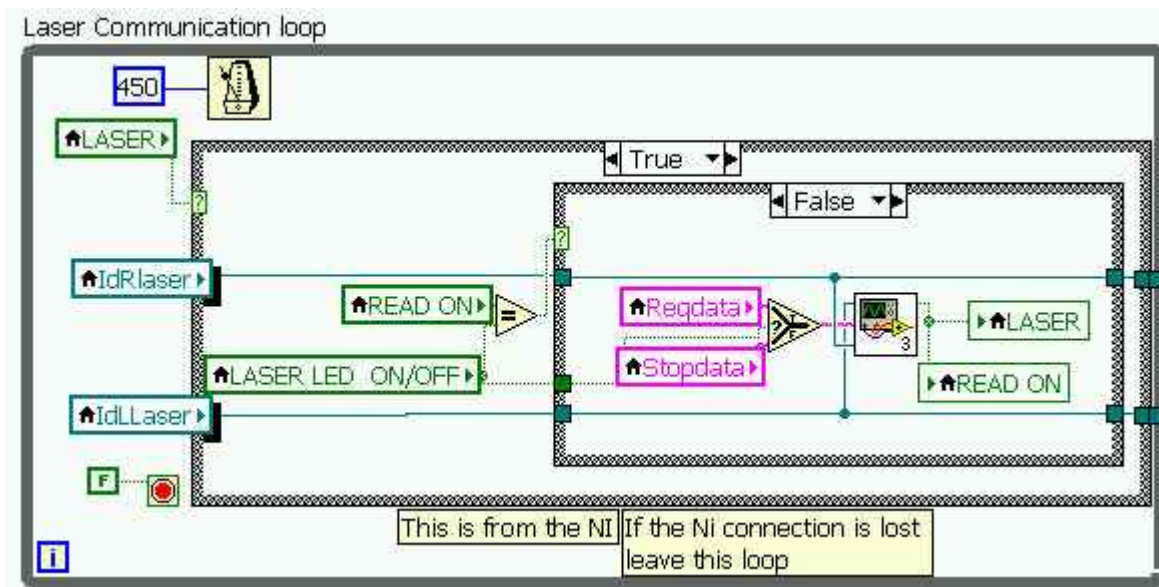


Fig. 79. Laser data communication loop.

The laser scanners communicate with the control computer through a set of commands. The communication loop allows for sending basic commands like measurement on/off.

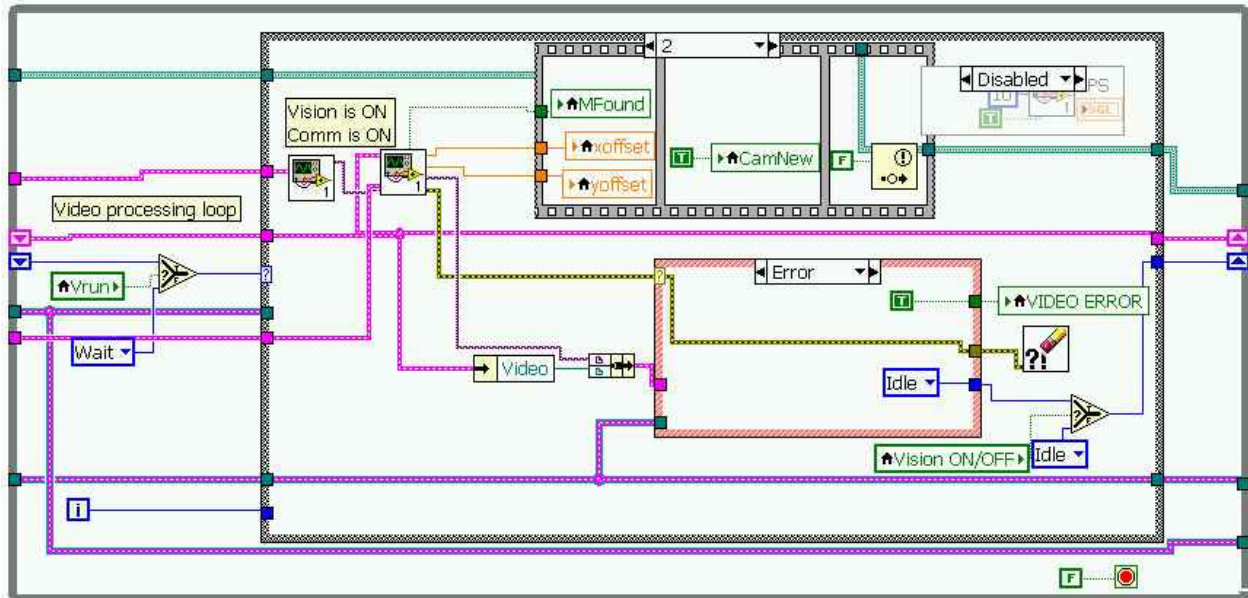


Fig. 80. Vision acquisition and processing loop.

The vision acquisition loop is a 60Hz rate loop receiving images from the camera, processing it and forwarding the results to the communication loop.

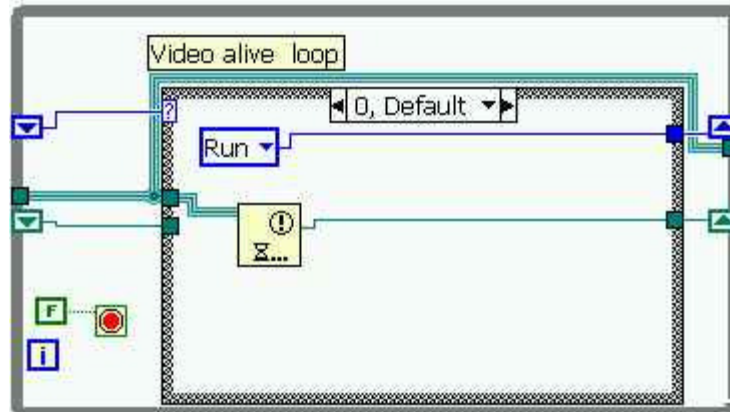


Fig. 81. Video watchdog loop.

The watchdog loop ensures the communication socket is alive. If it senses broken socket it informs the connection manager to reset and reestablish the connection.

Vita

Andrzej Nycz was born in 1979 in Jarocin in South-East Poland. He received his Master of Science degree in Mechanical Engineering from Warsaw University of Technology in June 2003. He started his employment as an engineer in 2002 as an Intern at The Industrial Institute for Automation and Measurement in Warsaw, Poland. In 2003 he joined the Machine Vision Division as a Research Engineer. His work focused on image processing for industrial applications of quality control. In 2005 he joined the Robotic Electro-Mechanical Systems Laboratory at The University of Tennessee as a Graduate Assistant. His worked focused on telerobotic systems control and visual servoing. His interests also include human-machine interfaces and human vision cognitive perception.