Realization of High-level Human Coaching System of RoboCup Middle Size League

Qingwu Fan, Jiaming Li, Sihan Li, Xudong Liu, Jing Huang

The Pilot College of Beijing University of Technology, Beijing 101101, China

hicx@bjut.edu.cn

Keywords: Robocup, Middle Size Group, QR code, speech recognition.

Abstract. Robocup middle size league has a relatively perfect autonomous decision-making system. To strengthen the human coach's role in grasping of the field situation, a high level human coaching system is presented by committee. This paper studies high level human coaching methods by QR code and human voice, and simulation results show that both the two schemes are feasible and reliable. It is possible to control the robots and make quick real time decisions during the match. Instructions' receiving and implementing have high accuracy and the application promotion value.

Introduction

MSL can be the most antagonistic game and has minimum difference to human soccer races. MSL Technical Committee alters the MSL rule and regulations every year to increase the difficulty which is intended to promote technology continually. MSL competition allows high level human coaching behaviors, teams can command the robots under the specify technology describing by the rules. In this paper, starting in the high level human coaching system, we studied and realized the high level human coaching by interpreted QR codes and voice coaching through dedicated microphones.

Middle Size Group Technology Overview

MSL rules and regulations limit the overall system structural diversity. The operational structure is divided into several subsystems. Each part is closely related to trainer system.

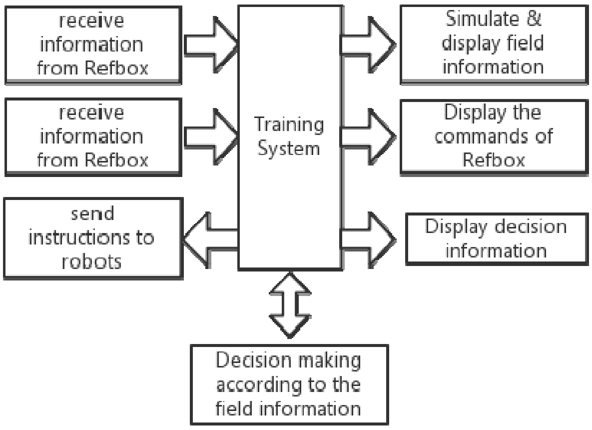
Visual subsystem is the most important environmental perception system. Most robots use omnidirectional cameras to get environmental information. Robots obtain 2D image from omnidirectional vision system, thus it's difficult to study in 3D coordinates. Water[1]、Nubot[2]、TU/e[3]have added a 3D visual device to get image from Kinect or some other similar equipment.

Decision making system is the core of the intelligence system, which has six[4], four[5], three[6], and two layers of decision-making model[7]. The conscious and independent policy makers are divided into groups. After group consciousness layer (run in trainers) integrate all information, it assessments situation and select formation for the team, then complete the role assignment, path planning, etc. After allocating the tasks, autonomous decision-making class (running on the robot) will correct the path choice and assign roles according to the vision subsystem of real-time data [错误!未定义书签。].

Control subsystem effect to robots' actual performance, including movement unit, control unit, stroke unit, and some special institutions. Nubot and TU/e improve the software and hardware of their ball control unit to coping with changing situations as well as strict rules.

Advanced human coach trainer design and implementation

Trainer. This article is based on two layers decision-making model, to realize efficient execution and dynamic transformation of group consciousness layer. Its main function is receiving instruction, making decision, and field information display. The model is shown in Fig.1.Trainer is responsible for collecting information uploading by field players, judging the current situation, and find a proper tactical to choose the best formation and role assignment. After receiving information, robots will modify the decision according to actual situation. Finally, players generate a direction controlling bottom wheel speed control circuit. The flow chart is shown in Fig.2. At the same time, the trainer is in charge of communicating with referee box. Trainer system need an interface wealth of information and control options to respond to the usual human testing.



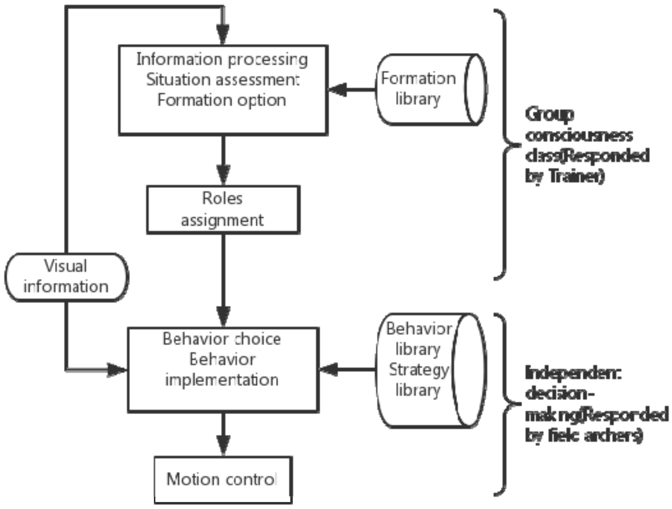
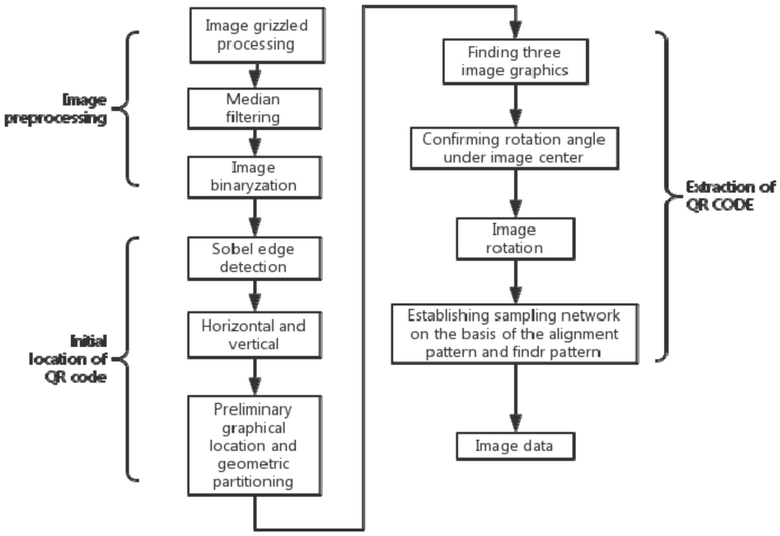


Fig.1 trainer model chart Fig.2two layer decision model flowsheet

Human-coached trainer. QR code, speech recognition, gesture recognition can all realize the human coaching. QR identification is most widely used, the best recognition rate and stability. Speech recognition instruction of short information can carry large temporary parameters. You can use the natural voice, is the most close to an implementation of human behavior. Gesture recognition has two kinds of recognition of 2D and 3D recognition, by identifying different gestures state and track its change to implement instructions recognition, but the operation is relatively complex. In this article, we use voice and QR code to implement information identification.

High level human behavior will intervene part of the trainer decision. After identification of human instruction will be conscious intervention groups, affect their choice after complete the information processing and situation assessment of the role of the formation and distribution.



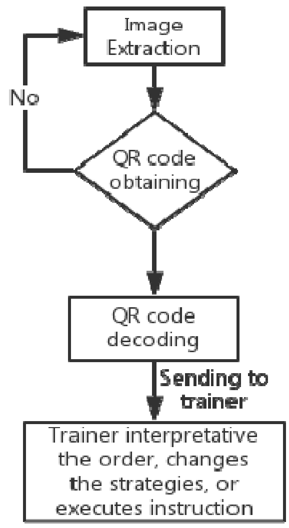


Fig.3.QR code image extraction flowsheet Fig.4.QR-based training system flowsheet

QR-based trainer. QR symbol is a square array composed of many square blocks. It's composed of code area and function area (detection patterns, separator, positioning and correction graphics), Functional area is not used for data encoding, the QR symbol has a white background.

Decoding process includes three steps, which includes image acquisition, image extraction and QR decoding. We can extract the QR code[8,9] as shown in Fig.3.Next process is: read the format, remove masks, code filtration, RS correction, data decoding, and instruction validation. After decoding, a preset instruction is obtained, which is transmitted to the trainer over TCP connection. The trainer, respond the instruction, implement the outside intervention, as shown in Fig.4.

Speech recognition based trainer. Speech recognition system is based on Windows speech recognition program, whose key technologies of mobile robot voice control. Microsoft Speech SDK (SAPI), containing SR (Speech Recognition) and SS (Speech synthesis) engine.

A list of directives is built to training the system to identify and respond to voice commands. The program captures the voice signal from monitoring computer microphone. After automatic speech recognition (ASR) converts the spoken word to text, correspondence command code is established and transferred, waiting for the voice confirmation. Flow is shown in Fig.5.

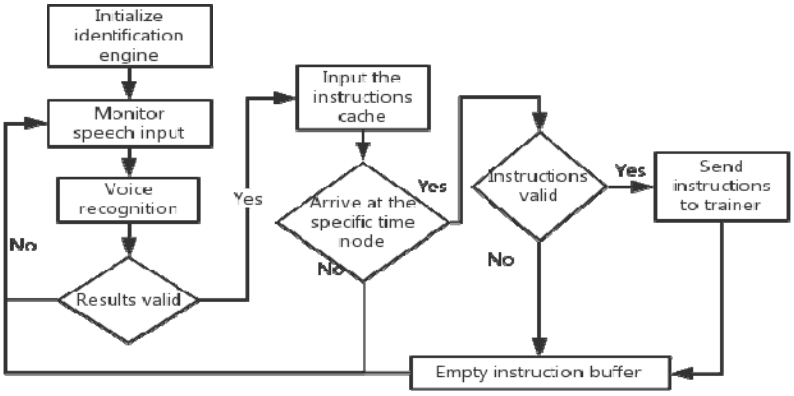


Fig.5.Speech recognition flowsheet

Command Design. Voice recognition system shares the similar command with QR recognition system. Take Chinese in voice recognition system for example, it abbreviates the command considering the accuracy of the QR recognition, thus greatly reduces the amount of information the QR code has to carry. The command system includes testing commands and commands for match. The way of QR code recognition is to show a QR code with predetermined commands to a camera, while one of the voice recognition is to say the command directly.

Predetermined Commands. Predetermined commands can be optimized since it has relatively fixed input information, thus making it of a higher accuracy than dynamic commands. Since the robot executes predetermined formation and tactics, these formation and tactics will be more stable to be carried out. Goalkeepers have their exclusive commands which control their reaction to the ball. It is designed to react in three ways-to guard the near point(a), distant point(b), and to face the ball directly all the time(c). Predetermined formation commands are: kick-off formation 1(d) and 2(e), free-kick formation 1(f) and 2(g), and throw-in formation 1(h) and 2(i). Predetermined tactics are: priority for defense (j) and priority for offense (k). All these predetermined formation and tactics command will be displayed on the trainer, which makes substituting formation according to situation on court possible. The QR for predetermined formation and tactics is shown in Figure 6.

Dynamic Commands. Since the predetermined formation and tactics cannot always be suitable under all circumstances, it can be supplemented by dynamic commands. These commands do not acquire generally fixed information, so its correspondent voice command and QR code graphics are not fixed, either. Under this mode, QR code generates graphics containing predetermined formation with QR code generator. And voice commands are directly issued by dictation. This implementation is influential in QR code. The longer commands are, the more information there are and the more complicated the graphics are. With the certainty of display area the recognition rate is low. Besides, due to the uncertainty of voice command, the accuracy of the recognition of key words is enhanced.

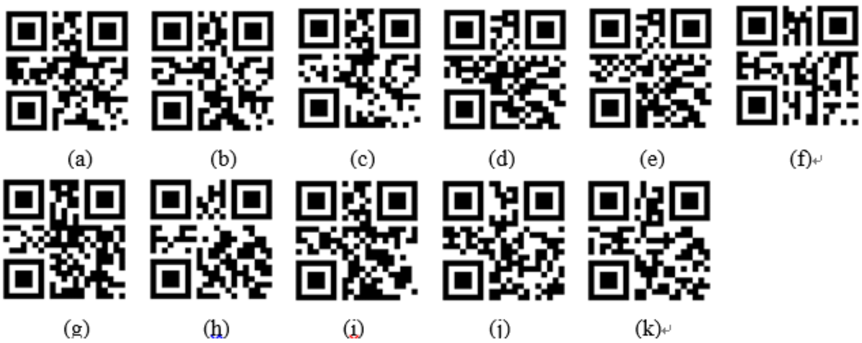


Fig.6.QR code for predetermined commands

To make sure the formation list changes along with the situation on court, the decision making system will only command the strategic device and its according position in that formation. When changing formation, only the predetermined strategic device should be involved; when changing individual behavior, the exact device of a certain ID should be involved.

The format for a dynamic command is: Type of command ID of a certain device (of ID of the strategic device) _Command 1_Command 2_Parameter 1_Parameter 2. The format is fixed so invalid parameters should be filled with an “X”. The design of the command is displayed in table 1.

Table1 Monomer and dynamic strategy instruction set

Command set	Instruction format					
	Type	Number	Instruction 1	Instruction 2	Parameter1	Parameter2
	UNIT	1-5	Attack	X	X	X
	UNIT	1-5	Defense	X	X	X
	UNIT	1-5	Move to	X	X	Y
	UNIT	1-5	Attack Priority	X	X	X
	UNIT	1-5	Defense priority	X	X	X
	Type	Number	Instruction 1	Instruction 2	Parameter1	Parameter2
	Tactic	1-5	M/O kick off	Move to	Coordinate X	Coordinate Y
	Tactic	1-5	M/O free kick			
	Tactic	1-5	M/O throw-in			

Testing

This section will be focusing on the testing of the decoding, voice recognition and network communication part of the QR decoding program. Whether the program designed is able to run properly as expected, and whether the stability can meet the need of a match will all be in the range of this test. The observation of whether the robots can perform correctly according to the command will be recorded. Result of the recognition and communication experiment is displayed in table 2.

As is shown in the table above, the accuracy of recognizing the QR code is 86% while the accuracy of recognizing a voice command is 68.5%. Therefore we can indicate that the recognition rate of a QR code is more stable than a voice command, since the recognition of a voice command can be easily influenced by environment and one’s pronunciation. As for the trainer device, due to it is connected with decoding system over TCP/IP protocol, the packet loss rate is basically zero. Also, the trainer device is connected with robots over UDP, which will send information about formation and individual behavior to robots repeatedly as soon as it receives the command, so judging from the statistics, the accuracy of the general process of communication reaches 100%.

Table2 Test Results

Test Results	Test Project				
	Project	QR Code Recognition	Code Instruction execution	Speech recognition	Speech Instruction execution
Test Times	executions	200	172	200	137
	successes	172	172	137	137

Conclusion

This paper aims to realize the program design on the trainer device based on the common behavior of human trainers. The two patterns introduced both have their advantages and disadvantages and space for improving, especially the recognition rate, which can be improved greatly with the optimization of the algorithms to stabilize the program. Furthermore, we can also optimize the algorithms of formations on the trainer device to deal with outside intervention, so that the command system will be more flexible, assuring robots a better performance on court.

References

[1] Water-Team. Team Description Paper[EB/OL]. http://blog.sina.com.cn/s/blog_96f5658d0102vds0. html.

[2] Junhao Xiao, Huimin Lu, Zhiwen Zeng, etc. NuBot Team Description Paper 2015[Z]. Beijing: National University of Defense Technology,2015.

[3] Cesar Lopez, Ferry Schoenmakers, Koen Meessen,etc. Tech United Eindhoven Team Description 2015[Z]. Netherlands: Eindhoven University of Technology,2015.

[4] Xinhe .Research on six-step reasoning model of robot soccer[A].Research and Discussion on Robot Soccer[C].1998:42-46

[5] Lijuan Wu, Chunhui Zhang, Xinhe Xu. Reasoning Model of Soccer Robot Strategy System[J]. Journal of Northeastern University, 2001, 22(6): 597-599

[6] Gerhard W.The Robocup challenge[J].Robotics and Autonomous System,1999,29:3-12

[7] HuLang. Research on the Key Technology of Decision-Making System for RoboCup Middle-Size Robot[D]. Chang an: Chang'an University, 2011. 15-16

[8] Mikio K, Takayuki Y, Tetsuo S, et al. Bar- code recognition system using image processing[C]/ /IEEE Symposium on Emerging Technologies & Factory Automation.[s. l.]: Institute of Electric and Electronic Engineer, 1997: 568- 572.

[9] Shellhammer S J, Goren David P, Pavlidis T. Novel signalprocessing techniques in barcode scanning[J]. IEEE Robotics and Automation Magazine, 1999, 6(1): 57- 65.