Proceedings of the 2011 IEEE
International Conference on Mechatronics
April 13-15, 2011, Istanbul, Turkey

Autonomous Humanoid Robot Navigation using Augmented Reality Technique

Omid Mohareri[#1], Ahmad B. Rad[#2]

[#]*School of Engineering Science, Simon Fraser University*
250 - 13450 – 102nd Avenue, Surrey, BC, Canada
[1]omoharer@sfu.ca
[2]arad@sfu.ca

Abstract— **This work presents a novel vision-based navigation strategy for autonomous humanoid robots using augmented reality (AR). In the first stage, a platform is developed for indoor and outdoor human location positioning and navigation using mobile augmented reality. The image sequence would be obtained by a smart phone's camera and the location information will be provided to the user in the form of 3D graphics and audio effects containing location information. To recognize a location, an image database and location model is pre-constructed to relate the detected AR-marker's position to the map of environment. The AR-markers basically act as active landmarks placed in undiscovered environments, sending out location information once detected by a camera. The second stage implements the same algorithm on an autonomous humanoid robot to be used as its navigation module. This is achieved by coupling the robot odometry and inertial sensing with the visual marker detection module. Using this system, the robot employs its vision system to enhance its localization robustness and allow quick recovery in lost situations by detecting the active landmarks or the so called AR-markers. The problem of motion blur resulting from the 6-DOF motion of humanoid's camera is solved using an adaptive thresholding technique developed to increase the robustness of the augmented reality marker detection under different illumination conditions and camera movements. For our experiments, we used the humanoid robot NAO and verified the performance of this navigation methodology in real-world scenarios.**

Keywords— **Humanoid robots, Autonomous navigation, Landmark based localization, Augmented Reality, Robotic vision**

I. Introduction

The research in the area of humanoid robotics is currently at the stage of producing robots which are expected to work in complex human environments and perform high-level human like tasks such as human assistance, home care and delivery. Recently, these platforms have become more attractive to the robotic research community as they present new perspectives like climbing stairs and accessing complex terrains, when compared to the wheeled mobile robots. One important task now is to equip these robots with world recognition, robust autonomous navigation and accurate localization capabilities which are prerequisites for them to operate in the real world.

In order to achieve long-distance autonomous navigation, global positioning within a given map is essential since the foot slippage is inevitable in bipedal locomotion and even minor positional and directional errors may cause critical failure in such navigation [1]. The main challenges that make the humanoid navigation problem different than that of a wheeled mobile robot are: (*i*) the excessive noise introduced to the system causing from the 6DOF motion of the on-board sensors while having the bipedal motion, (*ii*) limited sensor choice causing from the size and weight restrictions, and (*iii*) unreliable odometry data causing from the serious noise in the executed motion commands depending on ground friction and backlash in the joints [2-5]. Accordingly, reliable and accurate localization for such systems is an elaborate task and requires high amount of uncertainty management.

Compact and lightweight cameras are often used as the sensor for localization in the case of small humanoid robots which have only a limited payload [5]. Vision is reported to be an attractive and easy choice of sensor to achieve positioning in humanoids because of its compactness, accuracy and the richness of information it offers in addition to its intuitive appeal to the researchers because of being the main navigation tool for humans and most animals [1, 3, and 6]. Almost every humanoid robot developed so far has been equipped with cameras, a stereo camera pair in their robotic heads in imitation of humans. Moreover, many researchers prefer to have higher similarity in the ability of humanoids with humans, and are reluctant to install convenient sensors such as range sensors in humanoids [1]. Consequently, it is worth investigating new ways to realize vision-based autonomous navigation.

Augmented reality on the other hand, is an emerging area in computer science that fuses multiple technologies in the fields of computer vision, computer graphics and user interfaces into a single system that enhances a human or a machine's perception of an environment. The information or knowledge that is implicit in a scene yet not immediately obvious to the sensors is overlaid to provide a more complete picture; in a nutshell, it provides a "reading between the lines". AR provides extensive practical benefits to human/machine systems by superimposing additional information in the form of computer-generated 3D and 2D graphics and/or audio sounds on the user's view of the surrounding environment that effectively enriches the human/machine perception and facilitates the understanding of complex 3D scenarios [7].

Recently, AR technology has been used for indoor navigation applications instead of the commonly GPS-based, RFID-based or other sensor-based techniques. A vision based location positioning system using augmented reality for indoor navigation is proposed in [8]. This system automatically recognizes a location from the image sequence taken of the indoor environments and uses augmented reality to overlay the user's view with location information. To recognize a location, an image database and location model consisting of locations and paths between them in the environment was

978-1-61284-985-0/11/$26.00 ©2011 IEEE

constructed. Location was recognized by using prior knowledge about the layout of the indoor environment.

In this study, we merge two areas of augmented reality based navigation and autonomous humanoid robot navigation. The design basically incorporates a novel vision-based navigation system using mobile augmented reality (MAR) in a humanoid robot to navigate autonomously in an indoor environment. The project consists of two stages. In the first stage, a new technique is proposed for indoor and outdoor navigation using mobile augmented reality. The environment images are obtained by a smart phone's camera and the user's location is realized by performing the processes of marker detection; image sequence matching and location recognition. The location information will be expressed to the user in the form of 3D graphics superimposed on the user's view through the phone's camera and audio effects containing location information sent out of the phone's speakers. To recognize a location, an image database and location model is pre-constructed to relate the marker's location to the environment's map.

In the second phase, the developed software platform in the first phase is implemented on the humanoid robot to be used as the robot's visual-localization and navigation module. This method improves the navigation and localization performance by presenting location-based information to the robot through different AR markers placed in the robot environment. Each AR marker is associated to a specific set of navigation instructions and location information given to the robot once it is detected by the humanoid's vision module.

The remainder of this paper is structured as follows. Section II presents an overview of the related studies reported in literature in these research areas in the last few years. Section III describes the detailed explanations of the proposed AR-based navigation system followed by some experimental results and discussions. Section IV briefly describes the humanoid robot used as the experimental platform for the second stage of the project. Section V is dedicated to the details of the robot's navigation strategy and its implementation along with some experimental results illustrating the performance of this technique. Finally, Section VI expresses the paper conclusions

II. RELATED WORK

In the last few years, several techniques for autonomous humanoid robot navigation and localization have been reported in the literature. The methods vary in the type of sensor utilized for environment perception, the localization belief i.e. EKF, Particle Filters, Monte Carlo, etc. and the type of humanoid robot platform used for experimental validation. For example, Ido *et al.* [1] present a long-distance indoor navigation based on a view-based method using a camera mounted on humanoid robot HRP-2 [9]. They basically estimate the robot location by comparing the current view to the previously recorded images of the environment. Thompson and Kagami [2] used a laser range sensor on HRP-2 humanoid to obtain accurate 3D information and compared them to the 3D map of the environment to realize the robot's location. A 6DOF particle filter is developed to keep track of the robot pose. The same authors reported established 2D localization methods on humanoid robot JSK-H7 [10] using dense stereo vision depth maps and robot odometry [3].

As for simultaneous localization and mapping (SLAM) algorithms and methodologies, a visual 3D SLAM has been implemented on HRP-2 in [11] by utilizing the robot's camera along with the walking pattern generator and odometry data within an EKF structure. Another efficient visual-SLAM algorithm for humanoid robots in presented by Davison *et al.* in [6] and describes a real-time algorithm which can recover the 3D trajectory of a monocular camera, moving rapidly through a previously unknown scene. This work proved that it is possible to achieve real-time localization and mapping with a single freely moving camera as the only data source. Kwak *et al.* proposed Particle-based SLAM technique so that the SLAM posterior is estimated by multiple hypotheses [12]. The major difficulty of the particle-based SLAM with 3D grid maps is the high computational cost. To reduce the computational cost, the authors also proposed a scheduling method for the time when to match and for particles that engage in the matching process. Tellez *et al.* in [13] proposed a SLAM method based on the laser measurements and odometry information and applied it on Reem-B humanoid robot [14] for indoor environments. Two small lasers installed in the robot feet are used to capture distance data. The SLAM problem is solved by using a multi-laser SLAM solution together with a holonomic motion model.

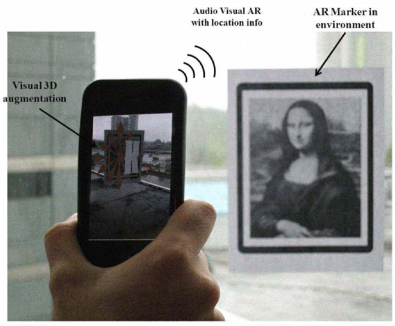
Lorch *et al.* [15] used footprints drawn on the ground surface to guide the robot's walking behaviour. Self-localization with relation to the foot-prints is achieved by tracking the position of blob features in the image with an EKF. In [16] the same authors of [15] extended their research work to track point objects and objects while walking. Ozawa *et al.* [17] presented 3D visual odometry and used dense feature maps to get the position of the camera. A well-known drawback of this incremental approach is the drift created by the accumulation of error. Simultaneously building a dense representation of the world and an accurate positioning of the camera seems to be quite difficult to achieve in real-time. Hornung *et al.* [4] proposed a Monte Carlo based localization method for indoor environemts by integrating the laser range measurements and odometry information and implemented their algorithm on NAO humanoid robot [18]. Based on the above literature review and to the author's knowledge, the presented work in this paper is the first landmark based visual localization and navigation technique developed for humanoid robots.

III. AR-BASED NAVIGATION SYSTEM

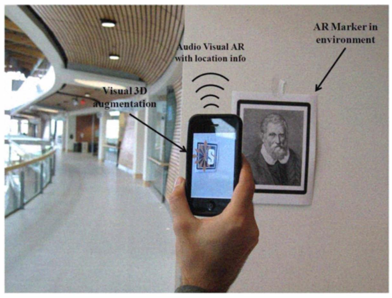
Recently, Augmented Reality technology has been used for indoor navigation applications instead of the popular sensor-based techniques. Kim *et al.* [8] developed a vision-based location positioning system using augmented reality for indoor navigation. The image sequence was obtained by a wearable mobile PC and a camera mounted on the user's cap, which transmitted the images to a remote PC for processing. The remote PC then performed marker detection, image sequence matching, and location recognition and transmits the recognized location information to the wearable mobile PC. A HMD (Head Mount Display) is used to annotate the direction sign on the user's view with location information.

The technique proposed in this project employs the minimum hardware comprising of a smart phone with camera to achieve the goal of localization and navigation in indoor and outdoor environments. The image sequence is obtained by a smart phone's camera and the whole processes of marker detection; image sequence matching and location recognition is done by the phone's processor, operating system and graphical modules. The location information will be provided to the user in the form of 3D graphics superimposed on the user's view through the phone's camera and audio effects containing location information sent out of the phone's speakers. In addition, a location model is pre-constructed to relate the marker's location to the environment's map. Using the smart phone's processing capability and implementing both audio and visual AR on the smart phone enabled us to get a fast and efficient location positioning performance in both indoor and outdoor environments with much cheaper hardware devices.

This system has several distinct advantages over all other vision-based and sensor-based location positioning systems. First, it offers an economical solution because the marker that identifies each location is simply a printed paper and the software can be installed in any class of smart phones. Second, it does not have the signal propagation, strength and multiple reflection restrictions of sensor-based localization techniques. The structure of the AR localization application and an example of its functionality is shown in Fig. 1.



Situation I: The user finds the AR-marker, tried to get some location info



Situation II: The user finds the AR-marker, tried to get some location info

Fig. 1. AR markers to be placed in the environment

Each marker unlocks a specific 3D model and sound effect that contains some location related information and clues to guide the user to the right path. A process diagram of the system is shown in Fig. 2. Given an image sequence taken from the smart phone's camera, this system annotates the recognized location information in the form of visual 3D graphics, audio effects or both. There are different ID numbers assigned to each marker that relate the marker position to the specific 3D model, audio file or both. Once a marker with a specific ID is detected, the ID is passed to the repository of 3D models or audio effects and the right effect will be chosen to be presented to the user.

When the marker detection process identifies a potential marker in the input image sequence, the image sequence matching process outputs location information by analysing and comparing both the features of the input image sequences and the images stored in the location dictionary. The results of the previous two processes and the location model are used to pinpoint the user's current location. Finally, the location annotation process asserts the direction signs and clues on the screen with current location information.

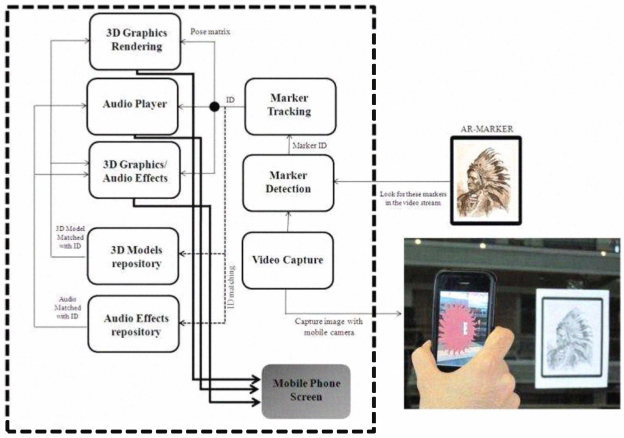


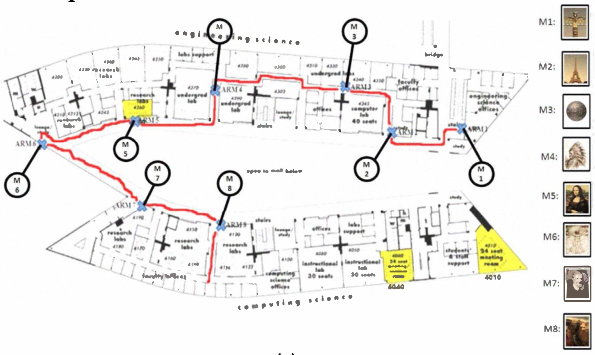
Fig. 2. Process diagram of the proposed system

In almost all of the augmented reality applications, the square shaped fiducial markers are the most commonly used markers [19]. The purpose is that a square shaped pattern provides at least four co-planar corresponding points so that the camera calibration can be achieved having a single marker in the scene. Each marker is coded with the pattern placed inside its inner square. The marker decoding is done using a simplified template matching procedure that compares three geometry invariants of the marker region with those of the pre-registered patterns in the code.

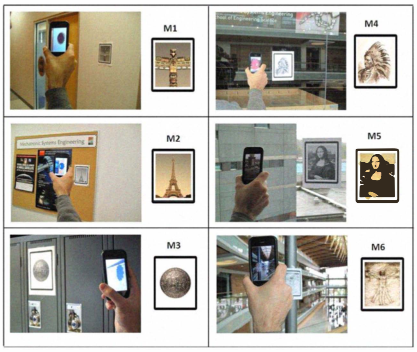
In order to reduce the tracking failure as a result of different illumination conditions, uniform shadows and reflections off a marker surface, and camera movements, an algorithm for selecting adaptive threshold values is implemented in this work. This algorithm which is originated from the work reported in [20] depends on the arithmetic mean of pixel intensities over a region of interest around candidate markers.

To verify the effectiveness of the proposed system, an experiment was performed using image sequences taken by the Apple iphone 3G [21] camera in an indoor environment. The images were taken at a rate of 30 frames per second and were digitized to a size of 640 × 480 (VGA resolution). The algorithm is developed on this platform using the iphone software development kit (iSDK) which is an objective-C based development environment and effective OpenCV vision library [22]. The marker detection process is applied to every frame of the input image sequence. To test the robustness of the implemented marker detection and adaptive thresholding techniques, different locations with different illumination conditions are tested. One of the experimental test results which show successful human navigation in an unknown environment using this technique is shown in Fig. 3. Eight different markers with distinct patterns which are placed in the indoor environment for this test are also shown in Fig. 3.

The proposed system in this section can be used both in human and robot location positioning applications. One novel application of this system is in autonomous robot navigation and localization. It's a potential application which is a cheap and efficient localization method for humanoid robots, and will be explained in the next section.



(a)



(b)

Fig. 3. (a) The human navigation and guidance in unknown environment using mobile AR. 8 different AR markers used in this experiment are shown on the right section of this figure (b) the human user finding different markers in the environment

IV. THE HUMANOID ROBOT NAO

NAO humanoid robot [18] is the test platform for the proposed autonomous navigation technique. NAO is a 57cm high humanoid robot with 25 degrees of freedom shown in Fig. 4.Its vision system consists of two VGA CMOS cameras with resolution of 640x480 and 30 frames per second located in the robot's forehead. Its camera's have a 58 deg field of view and focus range of 30 cm to infinity. The robot has a CPU in its head as the main processing unit: it's a GEODE 500 MHz board with 512Mo of flash memory and Wi-Fi connection for connection to the development remote computer.

The AR marker detection module is a vision module developed for NAO humanoid robot using C++ and OpenCV vision library to detect AR markers and return their size and position in the video stream. The same location positioning technique used for smart phones has been developed for the robot to be used as its localization module. Each AR marker is associated with some specific navigation instructions and location information. Once the robot detects a specific AR marker from a specific predefined distance, the robot can update its posterior estimate of the location according to that marker and decide on its next movement. The AR marker detection module has been tested under normal office lighting conditions (between 100 to 500 lux) and its detection range is between minimum of 0.035 rad = 2 deg to the maximum of 0.40 rad = 23 deg. The marker tilt angle could be +/- 60 degrees with 0 degree corresponds to the marker facing the camera. The special architecture of NAO operating system called NaoQi brought us the ability to run the developed applications on a remote development computer and run them on the robot once they are connected through a wireless connection.

The robot's odometry data are captured using the inertial unit placed in robot's chest as is shown in Fig. 4. The output data of this unit enables computing the chest speed and attitude (roll, pitch and yaw). This unit is composed of two axis gyrometers (5% precision with an angular speed of ~500°/s) and three axis accelerometer (1% precision with an acceleration of ~2G). The algorithm implemented to accomplish more accurate odometry data is to fuse both of these sensors data when the robot is in the dynamic mode to reduce the uncertainty. The accelerometer is the only absolute reference and generates acceptable torso angle in static mode. Whenever some motion is observed, the output angle is computed using the gyrometers which have satisfactory performance in dynamic mode. However, integration of gyrometers creates a bias in the angle computation, so in dynamic mode a fusion of the accelerometer and gyrometer is carried out to reduce the noise. A typical odometry data generated using this technique during the robot navigation is shown in Fig. 4.

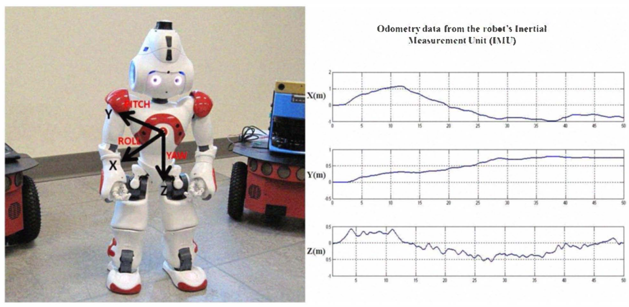


Fig. 4. The NAO humanoid Robot and its sample odometry data using the Inertial Measurement Unit

V. THE AUTONOMOUS HUMANOID ROBOT NAVIGATION STRATEGY

View-based or appearance-based navigation technique is one of the main autonomous navigation approaches in which an image is used to detect environment locations without going through a complete and accurate feature extraction. An image matching procedure is utilized in these techniques to perform localization and navigation without having explicit metric maps. Matsutmoto *et al.* [23] performed a view-based robot navigation on HRP-2 [9] using a set or sequence of view images. The robot recognizes its position by comparing its current view to the set or sequence of memorized views in its data-base. The major drawback of these view-based techniques is that the robot has to go through a learning or recording phase to update its knowledge base before it initiates the autonomous navigation. In this research, a new view-based navigation technique is developed using the AR location positioning approach developed for humans in section III. The advantage of the proposed method over similar view-based approaches is that the robot just needs to extract the AR-marker in the current view and match it to its database of AR-markers without going through an initial recording phase. The outline of this navigation strategy is shown in Fig. 5. The process of comparing an AR-marker in the current view with the AR-marker in the robot's database and calculating their correlation value is called the matching process. If the computed correlation value is larger than a specified threshold, the matching process is considered to be successful and the robot can get location information and update its posterior location belief.

One serious problem in the vision-based humanoid robot navigation techniques which is recently reported by Pretto *et al.* [24] is the 6-DOF robot motion that results to image blur and uncertainty in the view detection. The major cause of this phenomenon is the small oscillations of the robot camera in z-axis in a simple navigation scenario as it is shown in Fig. 4. The adaptive thresholding marker detection which uses the arithmetic mean of pixel intensities over a region of interest around candidate markers in the camera view is used in this work to increase the robustness to motion blur. In addition, the robot needs to only detect the AR-marker in its view and will not perform a thorough feature extraction once it detects the marker. This also improves the robot's robustness to motion blur while it's having bipedal locomotion. The robot autonomous navigation using this technique is tested in an indoor environment using five different AR-markers and the result is shown in Fig. 6. This proposed navigation strategy is added on top of a path planning module for obstacle avoidance and footstep planning developed in another stage of the project. The robot also keeps track of its local position using an EKF which incorporates the odometry measurements for the pose estimation with the camera measurements for the state correction stage.

A common model for processing landmarks in robot localization assumes that the sensor can measure the range and the bearing of the landmarks relative to the robot's local coordinate frame [25]. In addition, the feature extractor may generate a signature. The signature is a numerical value that characterizes the type of observed landmark, or a multi dimensional vector characterizing the landmark (e.g. height and colour).If we denote the range by r , the bearing by φ, and the signature by s, the feature vector which is the output of the feature extractor function f is given by:

$$f(z_t) = \{f_t^1, f_t^2, f_t^3, \dots\} = \left\{ \begin{pmatrix} r_t^1 \\ \varphi_t^1 \\ s_t^1 \end{pmatrix}, \begin{pmatrix} r_t^2 \\ \varphi_t^2 \\ s_t^2 \end{pmatrix}, \begin{pmatrix} r_t^3 \\ \varphi_t^3 \\ s_t^3 \end{pmatrix}, \dots \right\} \tag{1}$$

Here, z_t is the raw camera images which go to the image processing module for feature extraction. In the case of having AR-markers as the active landmarks in the robot view (x_m, y_m), bearing (φ_m) and the signature would be the marker ID (ID_m). Therefore the feature vector extracted from each camera view is:

$$f(z_t) = \{f_t^1, f_t^2, f_t^3, \dots\} = \left\{ \begin{pmatrix} x_{m,t}^1 \\ y_{m,t}^1 \\ \varphi_{m,t}^1 \\ ID_{m,t}^1 \end{pmatrix}, \begin{pmatrix} x_{m,t}^2 \\ y_{m,t}^2 \\ \varphi_{m,t}^2 \\ ID_{m,t}^2 \end{pmatrix}, \begin{pmatrix} x_{m,t}^3 \\ y_{m,t}^3 \\ \varphi_{m,t}^3 \\ ID_{m,t}^3 \end{pmatrix}, \dots \right\} \tag{2}$$

This feature vector is used to locate the robot in the environment based on the marker ID, size and orientation in the camera view. A typical feature vector which is the output of the AR-marker detection module corresponding to the situation shown in Fig. 7 is: $(Mark\ ID\colon 130,\ \varphi = 0.081,\ Width\colon 0.049,\ Height\colon 0.049)$. The proposed technique reduces the calculation costs as it eliminates the complex feature extractions and uses simple low resolution markers printed as black and white markers as active landmarks augmenting location information to the robot. The proposed AR marker detection technique tested on this platform is shown in Fig. 7. This method could be used as the correction or measurement update stage in incorporating any other kind of probabilistic filtering i.e. PF, Monte Carlo etc. for accurate pose estimation of humanoid robots.

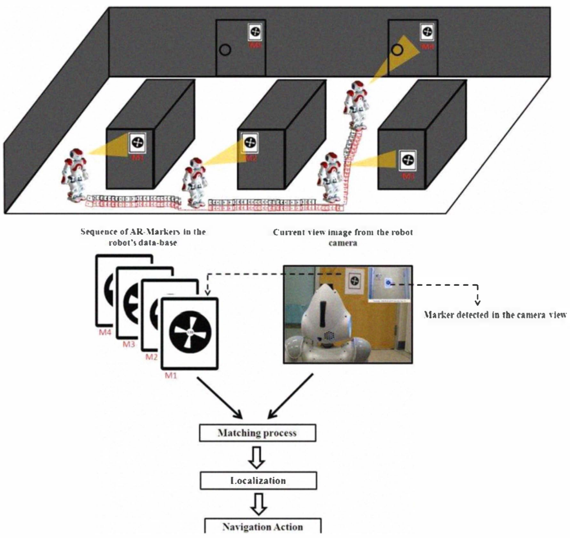


Fig. 5. The outline of the AR-based navigation strategy for NAO humanoid robot

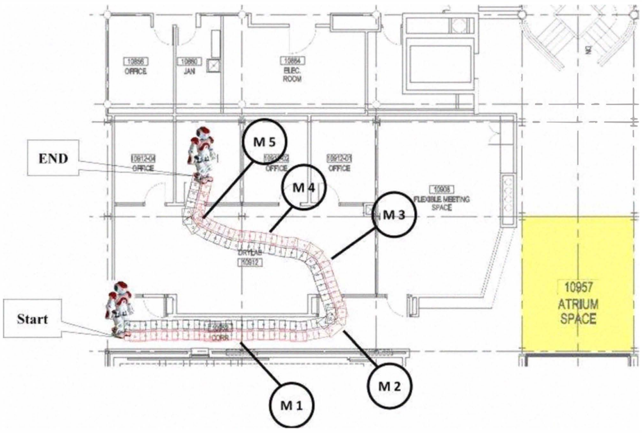


Fig. 6. The robot autonomous navigation test in indoor environment, incorporation of the robot odometry measurements with the AR-navigation technique, 5 different AR-markers were placed in this environment

VI. Conclusions

A novel location positioning technique using augmented reality is developed for human and robot navigation in unknown environments. This technique is implemented on a smart phone with camera and used by human users who want to navigate and find their way in new environments. Location positioning results show the efficiency of this system in comparison to similar works reported in this field because of its low cost, small weight and size in addition to its fast and high quality detection algorithm.

The same system is applied as a vision based navigation module to determine the position of a humanoid robot with respect to its environment. The AR markers placed in the environment can easily replace the radio beacons in the similar landmark localization methods. The proposed approach solves the problem of motion blur in the similar view-based navigation techniques by using the adaptive thresholding method for the marker detection task. The same AR navigation module will be used as a part of a visual simultaneous localization and mapping (Visual-SLAM) system which is under development for the same humanoid robot platform.

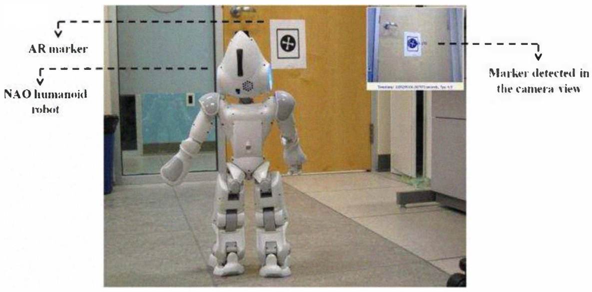


Fig. 7. A real-world scenario in which NAO humanoid robot detects an AR-marker in the lab environment and finds its way through

References

[1] [J. Ido, Y. Shimizu, Y. Matsumoto and T. Ogasawara, "Indoor Navigation for a Humanoid Robot Using a View Sequence," *International Journal of Robotic Research*, vol. 28, No. 2, pp. 315-325, Feb. 2009.

[2] S. Thompson, S. Kagami and K. Nishiwaki, "Localization for Autonomous Humanoid Navigation," *in Proc. of IEEE-RAS Int. Conf. on Humanoid Robots*, pp. 13-19, 2006.

[3] S. Thompson and S. Kagami, "Humanoid Robot Localization using Stereo Vision," *in Proc. of IEEE-RAS Int. Conf. on Humanoid Robots,* 2005.

[4] A. Hornung, K. M. Wurm and M. Bennewitz, "Humanoid Robot Localization in Complex Indoor Environments," *in Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2010.

[5] S. Obwald, A. Hornung and M. Bennewitz, "Learning Reliable and Efficient Navigation with a Humanoid," *in Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 2048-2053, pp. 2375-2380, 2010.

[6] A. J. Davison, I. D. Reid, N. D. Molton and O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, No. 6, pp. 1052-1067, June 2007.

[7] T. Chuan and C. Wang, "Registration based on scene recognition and natural feature tracking techniques for wide-area augmented reality systems," *IEEE Trans. on Multimedia*, vol. 11, No. 8, pp. 1393-1406, Dec. 2009.

[8] J. B. Kim and H. S. Jun, "Vision based location positioning using augmented reality for indoor navigation," *IEEE Trans. on Consumer Electronics*, vol. 5, No. 3, pp. 954-962, Aug. 2008.

[9] K. Kaneko, F. Kanehiro, S. Kajita, H. Hirukawa, T. Kawasaki, M. Hirata, K. Akachi and T. Isozumi, "Humanoid robot HRP-2," *in Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 1083-1090, 2004.

[10] K. Nishiwaki, S. Kagami, J. J. Kuffner, M. Inaba and H. Inoue, "Humanoid 'JSK-H7': Research Platform for Autonomous Behavior and Whole Body Motion ," *in Proc. of the Third IARP Int. Workshop on Humanoid and Human Friendly Robotics,* pp. 2-9, 2002.

[11] O. Stasse, A. J. Davison, R. Sellaouti and K. Yokoi, "Real-time 3D SLAM for Humanoid Robot considering Pattern Generator Information," *in Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2006.

[12] N. Kwak, O. Stasse, T. Foissotte and K. Yokoi, "3D Grid and Particle based SLAM for a Humanoid Robot," *in Proc. of IEEE-RAS Int. Conf. on Humanoid Robots,* 2009.

[13] R. Tellez, F. Ferro, D. Mora, D. Pinyol and D. Faconti, "Autonomous Humanoid Navigation Using Laser and Odometry Data," *in Proc. of IEEE-RAS Int. Conf. on Humanoid Robots,* 2008.

[14] R. Tellez, F. Ferro, S. Garcia, E. Gomez, E. Jorge, D. Mora, D. Pinyol, J. Oliver, O. Torres, J. Velazquez and D. Faconti, "Reem-B: An autonomous lightweight human-size humanoid robot ," *in Proc. of IEEE-RAS Int. Conf. on Humanoid Robots,* pp. 462-468, 2009.

[15] O. Lorch, J. F. Seara, K. H. Strobl, U. D. Hanebeck and G. Schmidt, "Perception Errors in Vision Guided Walking: Analysis, Modeling and Filtering," *in Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 2048-2053, 2002.

[16] J. F. Seara, K. H. Strobl, E. Martin and G. Schmidt, "Task-Oriented and Situation-Dependent Gaze Control for Vision Guided Humanoid Walking," *in Proc. of IEEE-RAS Int. Conf. on Humanoid Robots,* 2003.

[17] R. Ozawa, Y. Takaoka, Y. Kida, K. Nishiwaki, J. Chestnutt, J. Kuffner, J. Kagami, H. Mizoguch and H. Inoue, "Using visual odometry to create 3D maps for online footstep planning," *in Proc. of IEEE Int. Conf. on Systems, Man and Cybernetics*, pp. 2643-2648, 2005.

[18] D. Gouailier, V. Hugel, P. Blazevic, C. Kilner, J. Monceaux, P. Lafourcade, B. Marnier, J. Serre and B. Maisonnier, "Mechatronic design of NAO humanoid ," *in Proc. of IEEE Int. Conf. on Robotics and Automation,* pp. 769-774, 2009.

[19] R. Yu, T. Yang, J. Zheng and X. Zhang, "Real-Time Camera Pose Estimation Based on Multiple Planar Markers," *in Proc. Of 5th Int. Conf. on Image and Graphics,* pp. 640-645, 2009.

[20] T. Pintaric, "An adaptive thresholding algorithm for the augmented reality toolkit," *in Proc. Of Int. IEEE Augmented Reality Toolkit Work Shop,* 2003.

[21] *http://www.apple.com/ca/iphone/iphone-3gs/*

[22] *http://opencv.willowgarage.com/*

[23] Y. Matsutmoto, M. Inaba and H. Inoue, " Visual navigation using view-sequenced route representation," *in Proc. of IEEE International Conference on Robotics and Automation,* pp. 83-88, 1996.

[24] A. Pretto, E. Menegatti, M. Bennewitz, W. Burgard and E. Pagello, "A Visual Odometry Framework Robust to Motion Blur," *in Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 2250-2257, 2009.

[25] S. Thrun, W. Burgard and D. Fox, *Probabilistic Robotics*, the MIT press, 2005