

# Image-Processing Technologies for Service Robot “HOSPI-Rimo”

## サービスロボット “HOSPI-Rimo” の画像処理技術

Tatsuo Sakai\*  
酒井 龍雄

Anselm Lim Yi Xiong\*\*  
リム イー シオン アンセルム

To enhance the quality of service robots which perform work as substitutes for people, easy observation of the surrounding situation at the time of remote operation, and environment recognition - especially human detection - at the time of autonomous movement are required. This report explains the image-processing technologies for these purposes, and their application.

人々の作業を代替するサービスロボットの性能を向上させようとしたとき、遠隔操作時の周囲状況の容易な観測や、自律移動時の環境認識、特に人を検知する性能の向上が必要となる。本稿では、この目的のためにサービスロボット “HOSPI-Rimo” に搭載した画像処理技術、および、その応用について解説する。

### 1. Concept of “HOSPI-Rimo”

Recently, there has been an increase in demand for service robots to reduce hospital staff’s burdens. In this aging society, hospital staffs are increasingly burdened with additional works, such as watching over patients and patrolling to look out for roaming patients or suspicious individuals. Therefore, we have developed a service robot, HOSPI-Rimo, with the goal of reducing the staff’s burden. It also inherits autonomous mobility and affinity from Transfer robot HOSPI[1][2][3]. HOSPI-Rimo is able to act as a remote agent of hospital staffs, so that they do not need to be on the spot, as shown in Fig. 1.

To achieve this goal, HOSPI-Rimo needs to be easy to operate remotely and able to make a certain degree of autonomous decisions. Therefore, we introduce two functions, namely Omni-view and human detection with classification.

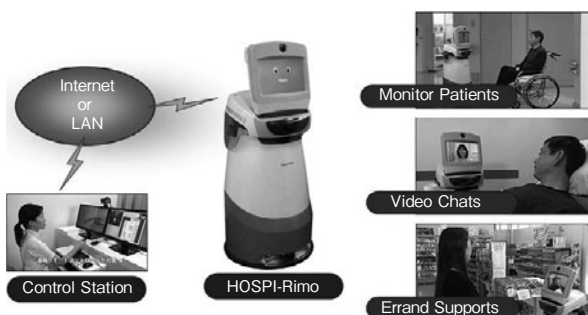


Fig. 1 HOSPI-Rimo

In the following sections, we will describe image processing technologies used in these functions.

### 2. Image processing technologies

#### 2.1 Omni-view

Omni-view is synthesized from images captured by four Panasonic car cameras (CY-RC70KD) mounted on the HOSPI-Rimo, as shown in Fig. 2. Images are transmitted wirelessly to user’s PC and mapped onto a 3D model constructed from four overlapping quarters of a hemisphere, via a mapping table pre-generated through calibration. Blending is then applied to all overlapping regions to create seamless and natural images. With the surroundings mapped onto the 3D model, the user can zoom in/out, pan or translate to his desired viewpoint in the 3D world. Omni-view images in Fig. 2 show a 3<sup>rd</sup>-person view, a top view and a close-up view from left to right.

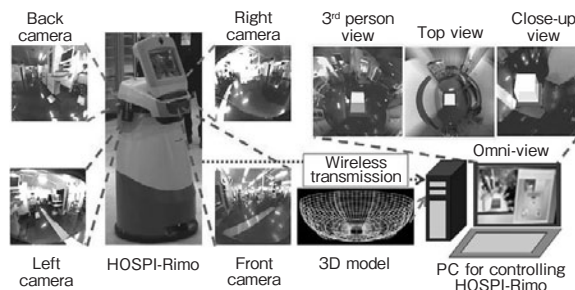


Fig. 2 Omni-view for HOSPI-Rimo

\* Hospital Robot Strategic Business Unit, Panasonic Healthcare Co., Ltd.  
パナソニック ヘルスケア株式会社  
院内ロボットストラテジックビジネスユニット

\*\* Panasonic R&D Center Singapore  
パナソニック R&Dセンター シンガポール

#### 2.2 Human Detection and Classification

Fig. 3 illustrates an overview of the Human Detection and Classification system. As shown in the figure, a scanning

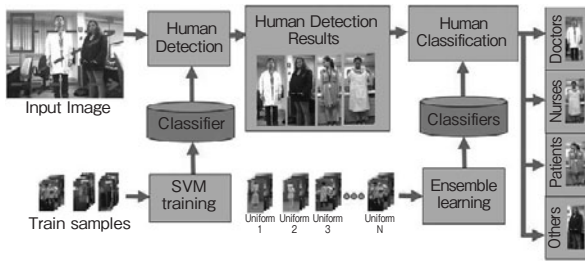


Fig. 3 Human detection and classification for HOSPI-Rimo

window process is applied on the input image. For each scan window, features are extracted and then correlated with basis functions trained earlier using Support Vector Machine (SVM), to generate a score measuring the confidence of containing a human.

Images captured by the car cameras in HOSPI-Rimo can also be utilized for automated human surveillance by Human Detection. The images have lower color contrast and more noise than images taken from conventional cameras like web cameras. We employ Histogram of Oriented Gradients combined with Multi-Resolution Local Binary Pattern (HOG-MRLBP) to increase the tolerance to this problem, exploiting the advantage given by using the additional spatial resolution in MRLBP[4]. Other methods like HOG-LBP-based human detection tend to fail in low contrast scenario (black suit with shadow in the background) as seen in Fig. 4, which compares results of detecting human with HOG-LBP, HOG-MRLBP, and HOG-MRLBP with dimension reduction.

Fig. 4 also shows a performance comparison of proposed method and various state-of-the-art benchmark methods like HOG[5], LBP [6] and HOG-LBP[7] on our testing data set of

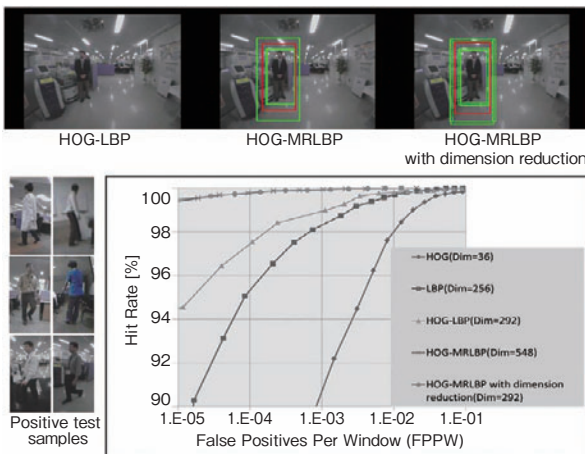


Fig. 4 Performance of various human detection methods

9,132 positive samples and 424,461 negative samples. HOG-MRLBP is shown to outperform the benchmark methods.

In HOG-LBP, the HOG portion contains 36 dimensions and the normal LBP portion contains 256 dimensions, having a total dimension of 292. In HOG-MRLBP, the LBP portion is replaced with MRLBP and it contains 512 dimensions, thus having a total dimension of 548. To speed up slow detection due to high dimension, we perform dimension reduction on the MRLBP portion from 512 to 256, resulting in 50 % improvement in speed and maintaining similar detection performance, as shown in Fig. 4. To do so, the absolute values of 512 MRLBP basis vector elements are sorted. We then create a 512-to-256 mapping table, by maintaining the 255 biggest-valued elements and combining the remaining 257 into the 256<sup>th</sup> element. Based on the mapping table, it is then re-trained as a new 292(36+256)-dimensional classifier.

To further speed up the system for real-time applications, optimization techniques like integral image, lookup tables, scan-window region of interest, and fast rejecters are deployed. The detection speed per VGA-size image improves from 3 sec to 400 ms. Due to characteristics of the cameras used, the Human Detection optimally senses humans located within 0.4 - 4.0 m.

Detected humans are normalized by determining a more precise head and shoulder location using a HOG-LBP trained head & shoulder detector. The normalized human image sample is then classified into Nurses, Doctors, Patients or others, utilizing Human Classification. We train the Classification module based on overlapping blocks of Hue, Saturation and Value histogram in region of interest using Ensemble Learning[8]. Using Doctor as an example, we trained an ensemble classifier by combining multiple smaller classifiers which was trained one against all via SVM. Sample results can be seen in Fig. 5. Based on the classification result, different warnings or signals can be sent to the nurse/security guard in duty.

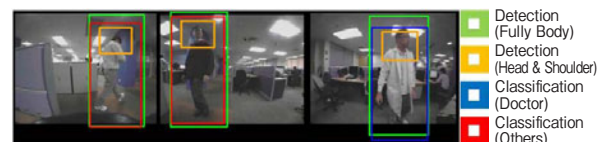


Fig. 5 Human detection and classification results

---

### 3. Future work

---

A feasibility test for remote control via Omni-view has been conducted in a real hospital. In the test, a nurse operated HOSPI-Rimo, and chatted with a patient via HD Visual Communications system (Panasonic KX-VC300) installed on HOSPI-Rimo. We achieved a desirable outcome from the test.

We are also planning for a feasibility test for the automatic human detection and classification function. In the test, a warning signal will be sent to a nurse, when roaming patients or suspicious individuals are detected. In this way, we are progressively advancing towards real deployment of the service robot at hospitals.

### References

- [1] T. Sakai et al., "Autonomous mobile robot system for delivery in hospital," Panasonic Electric Works Technical Report, vol.53, no.2, pp.62-67, 2005.
- [2] T. Sakai et al., "Transport system with autonomous mobile robots," Panasonic Electric Works Technical Report, vol.57, no.3, pp.16-22, 2009.
- [3] R. Takahashi et al., "Robotization of frequent small quantity conveyance - usefulness of autonomous transfer robot HOSPI," The Journal of Japan Society for Health Care Management, vol.12, p.261, 2011.
- [4] Timo Ojala et al., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," vol.24, no.7, pp.971-986, IEEE TPAMI, 2002.
- [5] Navaneet Dalal et al., "Histogram of oriented gradients for human detection," CVPR, 2005.
- [6] Y. Mu et al., "Discriminative local binary patterns for human detection in personal album," CVPR, 2008.
- [7] X. Wang et al., "An HOG-LBP human detector with partial occlusion handling," ICCV, 2009.
- [8] Thomas G. Dietterich, "Ensemble methods in machine learning," vol.1857, pp.1-15, Lecture Notes in Computer Science, ACM, 2000.