Robotica (2012) volume 30, pp. 813–826. © Cambridge University Press 2011
doi:10.1017/S0263574711001020

Performance analysis of a human–robot collaborative target recognition system

Y. Oren†, A. Bechar‡* and Y. Edan†

†*Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer-Sheva, Israel*
‡*Institute of Agricultural Engineering, The Volcani Center, A.R.O. Bet-Dagan, Israel*

(Accepted August 25, 2011. First published online: October 3, 2011)

SUMMARY

This paper presents a comprehensive analysis of time and action operational costs on an objective function developed by Bechar *et al.* (A. Bechar, J. Meyer and Y. Edan, "An objective function to evaluate performance of human–robot collaboration in target recognition tasks," *IEEE Trans. Syst. Man Cybern. Part C* **39**(6), 611–620 (2009)) for collaborative target recognition systems. Different task types, system reaction types, and environments were evaluated. Results reveal two types of task and system reactions – one focused on minimizing false alarms, and the second on detecting a target when one is presented. In addition, the analysis reveals a new property of the objective function based on a specific ratio between the weight differences that generalizes the model's objective function and facilitates its analysis. Results indicate that human decision time strongly influences system performance.

KEYWORDS: Operational level; Human–robot collaboration; Objective function; Performance analysis; Operational cost; Target recognition.

1. Introduction

Current robotic systems have limited recognition performance in real-world environments that are unknown,[1] unstructured, and continuously changing.[2] Sensory limitations[3] further reduce performance. Complexity in target recognition tasks is further increased when dealing with natural objects, such as in medical and agricultural environments, due to the object's high degree of variability in shape, texture, color, size, orientation, and position.

Humans' acute perception capabilities enable them to deal with a broad scope of vague and unstructured definitions.[4] Moreover, humans have superior recognition capabilities and can easily adapt to changing environmental and object conditions.[5] However, a human operator is not consistent, tends to fatigue, and is subject to distraction.[6] By taking advantage of human perception skills and the accuracy and consistency of the autonomous system, an integrated system can be simplified, resulting in improved performance.[7]

Several cooperative systems have been developed. Sheridan[8] divides automation into 10 levels, from fully autonomous, with no human intervention to fully manual. Scholtz[9] describes five roles that a human can take when interacting with a robot: supervisor, operator, mechanic, peer, and bystander. Howard[10] focused on role allocation in human–robot collaboration for space missions. Bechar *et al.*[11] defined four human–robot collaboration levels for target recognition tasks in unstructured environments using an objective function developed to determine the expected value of task performance, given the parameters of the system, the task, and the environment.[12]

In target recognition, the general task is to detect the objects and distinguish between the required objects (targets) and non-target objects. The ability to discriminate between target and non-target is limited by the distance between the means of the two distributions, defined as variable d′, which is also defined as sensitivity.[13] When $d' = 0$, the two distributions completely overlap and are impossible to be distinguished. As the value of d′ increases, they become easier to be distinguished. The location of the threshold is often defined in terms of the cut-off point or the likelihood ratio between the target and non-target probability distributions as measured at the threshold position and denoted as β. Discrimination analysis is commonly performed using the Receiver Operating Characteristic (ROC) curve[14] (e.g., sensor fusion for land mine detection;[3] discriminating disease cases from normal cases in medical applications;[15] and comparing diagnostic tests.[16] In an ROC curve, each detector or diagnostic represents a single curve on the hit-FA space where the sensitivity influences the convexity of the curve.

Common performance measures for evaluating target recognition include[6,17] probability of target recognition (hit), probability of non-target recognition (false alarm), and recognition time.[18] However, the performance measure values are task-dependent.[19] For example, systems for land mine detection tasks are designed to prioritize detecting targets, while systems for medical-oriented tasks prioritize minimization of false alarms. This can be implemented by assigning different weights to the performance measures according to the task characteristics using an objective function.[19] Alternatively, the Pareto optimal sets[20] can be used when the solution consists of different objective values that cannot be compared.

This paper is based on Bechar's model[12] for human–robot collaboration in a target recognition system. Previous analyses[2,12,19] were limited, since they excluded operational costs and critical parts regarding misses and correct

* Corresponding author. E-mail: avital@volcani.agri.gov.il

rejections. This paper presents in-depth analyses expanded to include operational costs, which are critical for implementation in real-world environments. In addition, the behavior of different types of systems in different environments and different tasks was extensively analyzed.

2. Model Description

2.1. Operational levels

The previous model[19] includes four operational levels based on Sheridan's[8] scale of "action selection and automation of decision" from fully manual to autonomous operation. (1) H: The HO detects and marks the desired targets single-handedly. (2) HR: The HO marks targets, aided by recommendations from the robot, i.e., some objects are automatically marked by a robot recognition algorithm; the HO must approve the correctly recognized targets and mark targets that the robot missed. (3) HOR: The HO supervises the robot. Objects are identified automatically by the robot's recognition algorithm and marked. The human assignment is to cancel false detections and to mark targets that were missed by the robot. (4) R: Fully autonomous, all targets are marked by the robot's recognition algorithm.

The model was developed in order to evaluate the performance of human–robot system in task operation, which requires target recognition operation.

2.2. System objective function

The objective function[19] considers four major parameter groups – human, robot, environment, and task. The objective function (V_{Is}, Eq. (1)) includes five parts, one for each of the four possible outcomes: hit (H), false alarm (FA), miss (M), correct rejection (CR), and the fifth for the operational costs:

$$V_{\mathrm{Is}} = V_{\mathrm{H}} - V_{\mathrm{M}} - V_{\mathrm{FA}} + V_{\mathrm{CR}} - V_{\mathrm{T}}. \tag{1}$$

Equations (2) through (5) illustrate the explicit part of each possible outcome:

$$V_{\mathrm{H}} = N \times P_S \times p_{\mathrm{H}} \times W_{\mathrm{H}}, \tag{2}$$

$$V_{\mathrm{M}} = N \times P_S \times p_{\mathrm{M}} \times W_{\mathrm{M}}, \tag{3}$$

$$V_{\mathrm{FA}} = N \times (1 - P_S) \times p_{\mathrm{FA}} \times W_{\mathrm{FA}}, \tag{4}$$

$$V_{\mathrm{CR}} = N \times (1 - P_S) \times p_{\mathrm{CR}} \times W_{\mathrm{CR}}. \tag{5}$$

All these equations are composed similarly, where N is the number of objects in the analyzed image and P_S is the probability for an object to be a target, and therefore $(1 - P_S)$ is the non-target probability. These parameters characterize the environmental conditions. The third parameter, p_x, symbolizes the probability of the system for one of the possible outcomes, where index x can be H, M, FA, or CR. This parameter considers the human and robot characteristics. The last parameter, W_x, symbolizes the weight's value for each possible outcome. The value of each weight depends on the task.

Operational costs include the time (V_{t}) and action (V_{c}) costs as illustrated in Eq. (6).

$$V_{\mathrm{T}} = \underbrace{t_{\mathrm{S}} \times W_{\mathrm{t}}}_{V_{\mathrm{t}}} + \underbrace{\overbrace{(N \times P_S \times p_{\mathrm{H}} + N \times (1 - P_S) \times p_{\mathrm{FA}})}^{W_{\mathrm{a}}} W_{\mathrm{C}}}_{V_{\mathrm{c}}}. \tag{6}$$

The time cost (V_{t}) is composed of t_{S}, which is the system time required to analyze an image, and W_{t}, which is the cost of one time unit. The action cost (V_{c}) is influenced only by the number of hit or false alarm outcomes (noted as W_{a}), since there is an actual action of the robotic system for these outcomes. W_{c} is the cost of one object recognition operation.

The system time, t_{S}, consists of the human operator (HO) time to confirm the robot hits (t_{Hrh}), the HO time to hit additional targets (t_{Hh}), the HO time to correct the robot false alarms (t_{FArh}), the HO time to mark false alarms (t_{FAh}), the HO time lost when a robot hit is missed (t_{Mrh}), the HO time invested when the robot missing a target (t_{Mh}), the HO time to correctly reject a robot false alarm (t_{CRrh}), the HO correct rejection time (t_{CRh}), and the robot time to process the image and perform hits or false alarms (t_{r}). Equation (7) explicitly illustrates t_{s}.

$$\begin{aligned} t_{\mathrm{S}} = {} & N \times P_S \times p_{\mathrm{Hr}} \times p_{\mathrm{Hrh}} \times t_{\mathrm{Hrh}} + N \times P_S \times (1 - p_{\mathrm{Hr}}) \\ & \times p_{\mathrm{Hh}} \times t_{\mathrm{Hh}} + +N \times (1 - P_S) \times p_{\mathrm{FAr}} \times p_{\mathrm{FArh}} \times t_{\mathrm{FArh}} \\ & + N \times (1 - P_S) \times (1 - p_{\mathrm{FAr}}) \times p_{\mathrm{FAh}} \times t_{\mathrm{FAh}} + N \times P_S \\ & \times p_{\mathrm{Hr}} \times (1 - p_{\mathrm{Hrh}}) \times t_{\mathrm{Mrh}} + N \times P_S \times (1 - p_{\mathrm{Hr}}) \\ & \times (1 - \mathrm{p}_{\mathrm{Hh}}) \times t_{\mathrm{Mh}} + N \times (1 - P_S) \times p_{\mathrm{FAr}} \\ & \times (1 - p_{\mathrm{FArh}}) \times t_{\mathrm{CRrh}} + N \times (1 - P_S) \times (1 - p_{\mathrm{FAr}}) \\ & \times (1 - p_{\mathrm{FAh}}) \times t_{\mathrm{CRh}} + t_{\mathrm{r}} \end{aligned} \tag{7}$$

Each of the human time variables represents a superposition of a decision time, t_{d}, and a motoric time, t_{m}, in accordance with the collaboration level.

The system objective function[12] evaluates the gains, rewards, costs, and penalties of the operational costs and the recognition performance measures, assigned with monetary value units (i.e., \$, £, or €). To adapt the objective function to a variety of tasks, systems, and environments, the objective function was defined as a weighted superposition of the performance measures. The *task* is defined by the goal to be accomplished (i.e., maximize hits; minimize false alarms; minimize misses; minimize execution time or combinations: e.g., maximize hits and minimize false alarms; minimize execution time and false alarms); the *system* is the physical means to operate the task and has specific characteristics such as operating time and costs; the *environment* is characterized by its targets and other objects (i.e., non-targets). Different weights of the first four performance measure parts (V_{H}, V_{M}, V_{FA}, and V_{CR}) represent different tasks; however, different weights of the cost part (V_{T}) represent different systems and target probabilities, and the number of objects represents different environments. For the same task and environment, systems with different costs will have different performances.

For example, in medical tasks, the hit and false alarm probabilities are more important than the recognition time; therefore, there is a high reward for a hit, high penalty for a FA, and a low cost of time unit is assigned. In the detecting

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

Table I. Summary of the variable parameters analyzed in the research.

Parameter	Description	Range	Step
W_{FA2H}	Ratio between the value of false alarm and hit	0.1,1,10	–
W_{CR2M}	Ratio between the value of correct rejection and miss	0.1,1,10	–
W_{M2H}	Ratio between the value of miss and hit	0.1,1,10	–
d'_r	The robot sensitivity	0.5–3	0.5
d'_h	The human sensitivity	0.5–3	0.5
P_S	Target probability	0.1–0.9	0.1

land mines task, the goal is a high hit rate, while the false alarm rate and recognition times are usually less important; hence, the reward for a hit will be high and the cost of each time unit and penalty for a false alarm will be relatively low. In a melon-harvesting task, the goal is a high hit rate; however, false alarms are not favored, since this will lower performance; hence, both rewards for hits and penalties for false alarms are high.

In real-time systems, recognition time is critical and hence the cost of each time unit will be high. In off-line activation of expensive systems, the cost of each time unit will be low; however, the cost of each action will be high.

An example for different environments is given in the cases of a melon-harvesting robot performing at the beginning of the season when there is a low number of objects with low target probabilities as opposed to the peak season, which will be characterized by high number of objects and high target probability.

2.3. Signal detection theory

A previously developed and modified Signal Detection Theory (SDT)[11,12,19] for two detectors (human and computer) was used to simplify the analysis. The performance of the first detector (robot) is determined by its sensitivity, d'_r, and its criterion, d'_r. The second detector (human) uses its sensitivity, d'_h, and two criteria – one for objects already marked by the robot, d'_{rh}, and another for objects unmarked by the robot, d'_h.

3. Methodology

3.1. Overview

This work focused on in-depth evaluation of different tasks, systems, and environments. A comprehensive numerical analysis was conducted by examining the influence of human, robot, environment, and task parameters on collaborative systems.

A numerical analysis of the objective function was conducted on a PC with Matlab 7.1™. The focus was determination of the best operational level for various systems, tasks, and environments, and investigating the operational costs. In addition, a sensitivity analysis of different parameters was conducted.

3.2. Best operational-level analysis

The influence of different human and robot sensitivity combinations (d'_h and d'_r) and different target probabilities (P_S) on the best operational level of various systems and tasks was analyzed.

Tasks that prioritize detection will set a high value for the hit weight (W_H). In contrast, tasks oriented to minimize false alarms will have a high value for the false alarm weight (W_{FA}). In addition, there are costs associated with system operational time and action. By setting different values for the objective function weights (W_H, W_{FA}, W_{CR}, and W_M) different tasks can be analyzed. To simulate different tasks, three ratio parameters between the objective function weights were set: the ratio between false alarm and hit weights, W_{FA2H}; the ratio between correct rejection and miss weights, W_{CR2M}; and the ratio between miss and hit, W_{M2H}. To create a drastic difference between the objective function weights, the values for these ratios were set as 0.1, 1, and 10. The value of hit weight, W_H, was set to 50, and all the other weights were determined according to the ratios. The values of different parameters of the simulation were extracted from a preliminary experiment performed by Bechar and Edan.[2] The probability for target (P_S) ranged from 0.1 to 0.9. The human sensitivity, d'_h, and the robot sensitivity, d'_r, ranged from 0.5 to 3. The operational cost weights were constant where the cost for one system action was set to $W_c = 2$ and the cost for one time unit was set to $W_t = 2000\,\text{hr}^{-1}$. The number of objects in each environment was set to $N = 1000$. The decision time for all human time parameters was set to $t_d = 5$ s/object, and the human motoric time was set to $t_m = 2$ s/(recognized object). The robot time was set to $t_r = 0.01$ s/object. In addition, all analyses were performed for optimal likelihood ratios. The optimal likelihood ratios, d'_r, d'_h, and d'_{rh}, were determined in the range between the logarithm of -4 and the logarithm of 4, in order to cover the available hit and false alarm probabilities. The simulation variables are summarized in Table I.

3.3. Operational costs analysis

This analysis aims to investigate the influence of operational costs on each operational level and on the system's overall performance. A comprehensive analysis of both operational cost parts – time and action – was conducted. The model's operational cost (V_T) consists of costs associated with time (V_t) and operation (V_c). The operational cost as a function of time, V_t, is affected by a number of decisions that the system (robot and human) has to make, how fast these decisions are made, and the cost per time unit of system operation. The operational cost as a function of operation, V_c, is affected by the cost per one operation of the robotic arm (action) and the number of times an action is required. Since the robotic arm moves only when there are hits and false alarms, this cost is only present when these outcomes occur.

4. Best Operational-Level Analysis

The numerical analysis focused on the influence of different human and robot sensitivity combinations (d'_h and d'_r) and

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

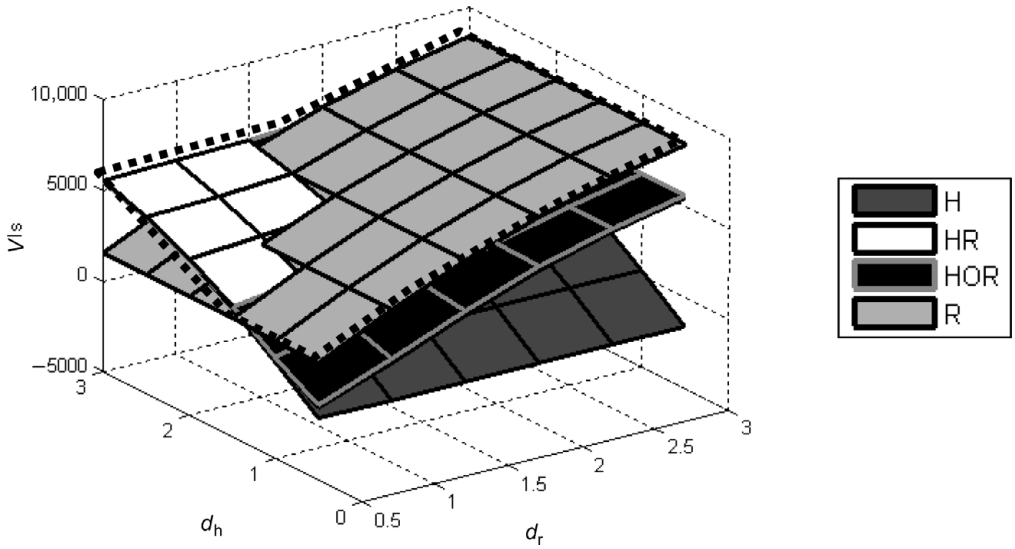

Fig. 1. The influence of human and robot sensitivities on the system's performance for each operational level – an example.

different target probabilities (P_S) on the best operational level of different systems and tasks. System performance for different human and robot sensitivity combinations is shown in Fig. 1. Each surface represents one of the system's possible collaboration levels. This figure illustrates the influence of the sensitivities on the objective function score (z-axis) and the highest score for each sensitivity combination, which is composed of the largest surface created from the surface intersections of different operational levels (its perimeter is marked with a dashed line). Furthermore, each intersection in this area represents a shift between the operational levels to maintain optimal performance.

4.1. Best operational level – definition

The best operational level is defined as the operational level that, under specific task parameters, achieves the highest objective function score (V_{Is}). Analysis was conducted using a 2-D graph that illustrates the best operational level map (Fig. 2), where each operational level is represented by a different pattern. In the case presented, the HR, HOR, and R levels are the best levels in different areas of the sensitivity space. This example indicates that for the same task the best operational level can change. Different task, human, robot, or environment parameters will result in different operation level maps.

4.2. Best operational level – results

Due to the multitude of results, the preliminary objective was to determine a common system "behavior" (the reaction to different parameters) between the different tasks analyzed. After investigating all the operational level maps that were produced by the analysis,[21] it was found that the maps could be classified into two main types of tasks based on the influence of the target probability on the best operational level

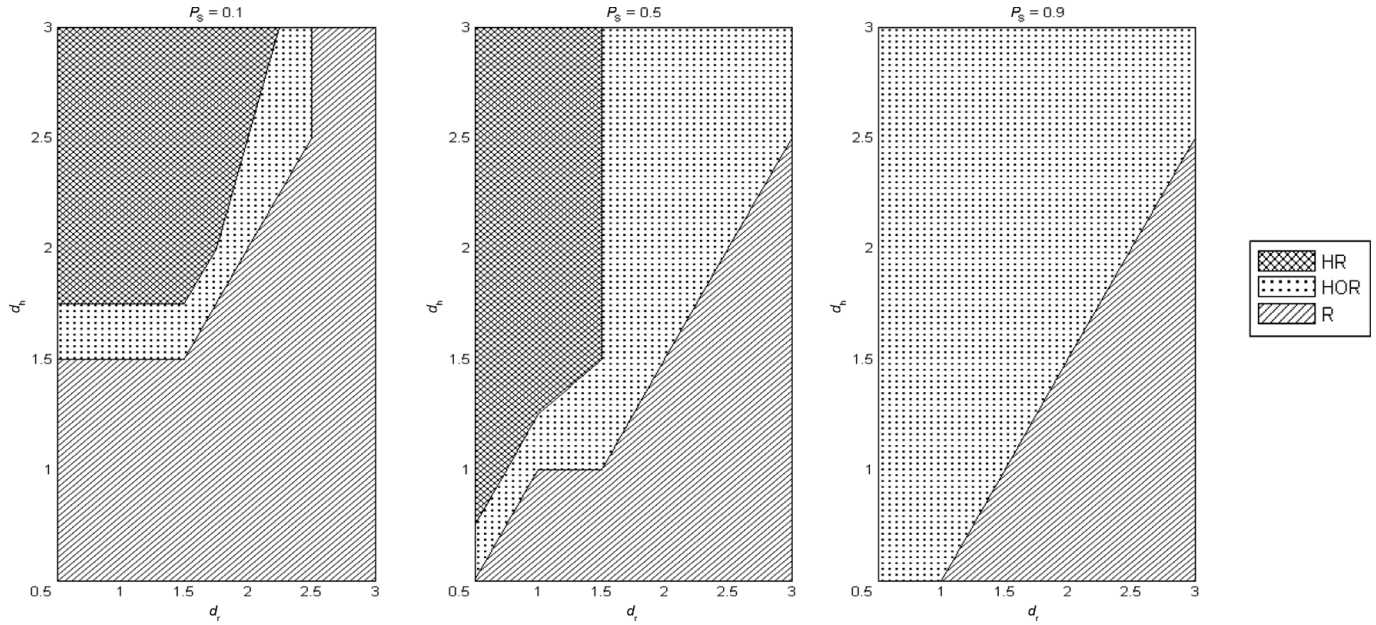

Fig. 2. Example of the system's best operational level maps for different target probabilities in a *Type 1* task.

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

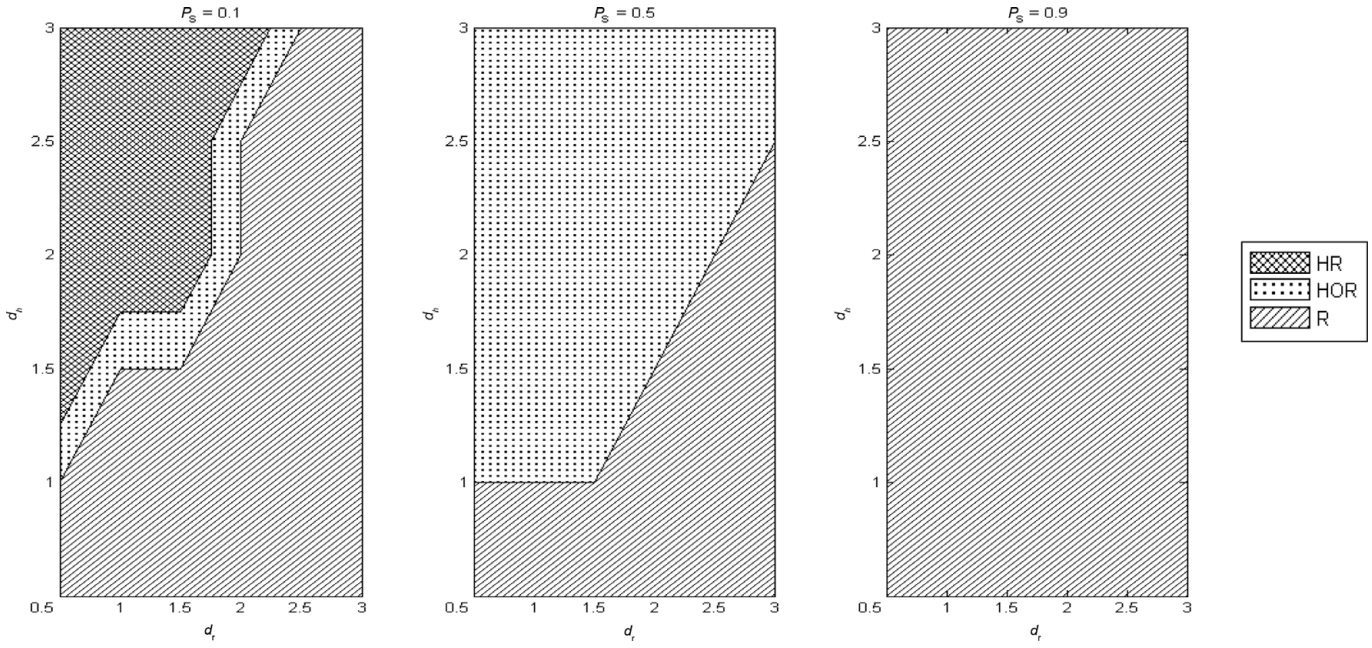

Fig. 3. Example of the system's operational level maps for different target probabilities in a *Type 2* task. System Properties Analysis.

maps; the objective function mathematical model appears to have an inherent symmetry with respect to the weight pairs and ratios. Although the threshold setting and the weight pairs and ratios are continuous variables, at a specific value (of the weight pairs and ratios), the system's reaction to the task changes from one type to another. It was found that the two types had opposite tendencies as a function of the target probability, i.e., the tendency of changes in the best operational level maps as a function of the target probability was opposite for the two types. The first type, denoted herein as '*Type 1*,' consists of reaction to tasks focused on minimizing false alarms. This goal can be reached by setting proportionately greater rewards for correct rejections and/or greater penalties for false alarms. The second type denoted herein as '*Type 2*,' consists of reaction to tasks focused on detecting targets when one is presented. This goal can be achieved by setting proportionately higher rewards for hits and/or higher penalties for misses.

4.2.1. Type 1: High priority to minimizing false alarms. Results indicated that as the target probability, P_S, increases, the area in which best performance is achieved by the R level, in the best operational level map, decreases. Furthermore, results indicated that the HR is the best level only when the human sensitivity, d'_h, is greater than the robot sensitivity, d'_r. It should be considered that when the target probability, P_S, is high, the HOR level is preferable in most of the sensitivity space. However, when robot sensitivity is greater than human sensitivity, the best level is R. The H level (human performs solely) was never the best operational level (Fig. 2).

4.2.2. Type 2: High priority to target recognition. As opposed to *Type 1*, in *Type 2* an increase in the target probability, P_S, increases the area of the R level in the best operational level map. Moreover, for high and intermediate target probabilities, the R level was found to be the best level when the sensitivity of the robot was greater than that of the human. The HR was found to be the best operational level only when the target probability was low, and the human sensitivity was greater than the robot sensitivity. For very low target probability ($P_S = 0.1$), the HR level is the best operational level in more cases than the HOR level, although as target probability increases, the HOR level performs better in more cases than the HR level (Fig. 3). Similar to the findings in *Type 1*, in the system's reaction to *Type 2* tasks, the manual mode (H) was never the best level.

Analysis of the objective function model for the best operational level (see Table I) reveals symmetry between hits and false alarms ratio, W_{FA2H}, and between correct rejection and miss ratio, W_{CR2M} (i.e., the same operational level map was generated when $W_{\mathrm{FA2H}} = \mathrm{X}$, $W_{\mathrm{CR2M}} = \mathrm{Y}$ and $W_{\mathrm{FA2H}} = \mathrm{Y}$, $W_{\mathrm{CR2M}} = \mathrm{X}$, where X and Y are the ratio values independent of the type). This implies that systems in different tasks with identical weight ratios can have identical best operational level maps. In order to examine these findings a further analysis of the objective function score was conducted. The objective function score was analyzed by analyzing contour graphs in the human and robot sensitivity space for different target probabilities (Fig. 4).

While there is uniformity in the best operational level maps of systems with different ratios, their objective function scores are different. The difference is derived from the way these systems achieve their task goal – minimizing false alarms for *Type 1* and recognizing targets for *Type 2*. In *Type 1* tasks, while some of the systems achieve the task goal of minimizing false alarms by giving a large penalty for false alarms, other systems achieve it by giving a large reward for correct rejections. In *Type 2* tasks, while some of the systems achieve the task goal of recognizing targets by giving a large reward for a hit, other systems achieve it by giving a large penalty for a miss.

Examples of these results are presented in Figs. 4 and 5: The objective function score maps of two systems in *Type 1* tasks show identical operational level maps as presented in Fig. 2.

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

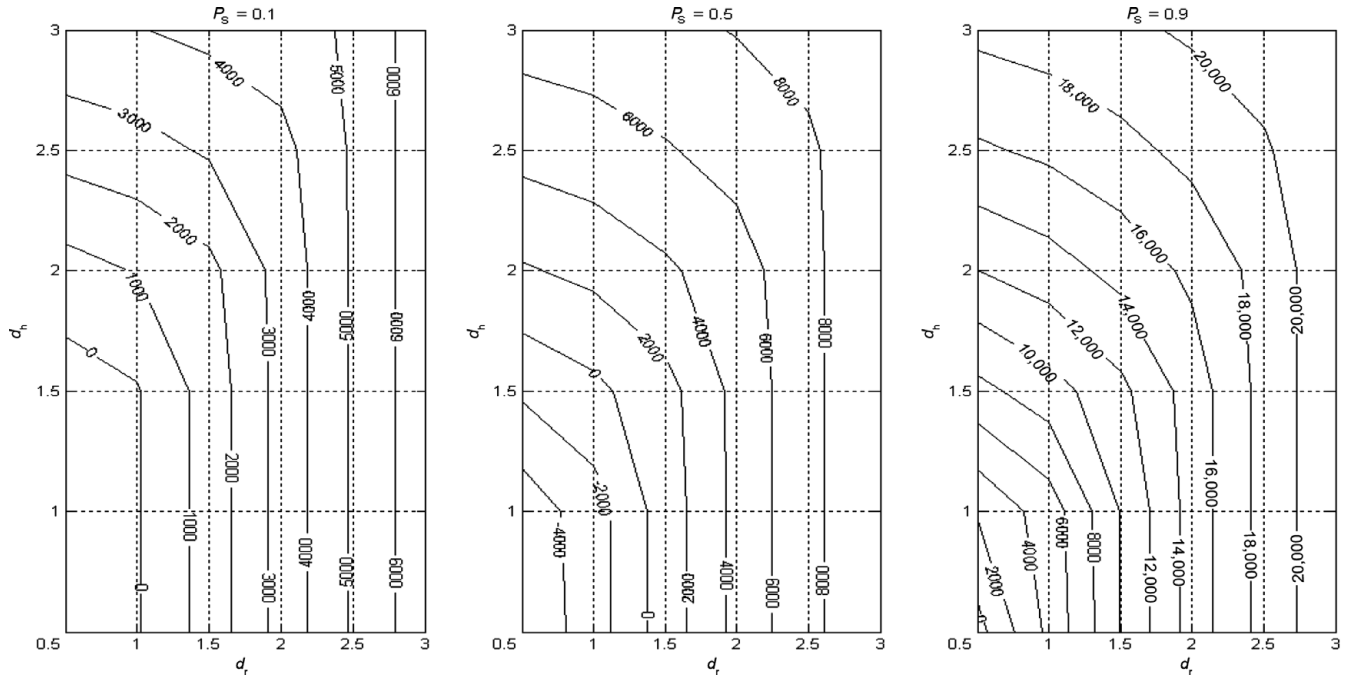

Fig. 4. Example of a *Type 1* objective function graph where the system's goal is achieved by giving high penalty for FA proportionally to CR.

Actually, these findings indicate that there is symmetry between hit weight (W_H) and miss weight (W_M), and symmetry between correct rejection weight (W_{CR}) and false alarm weight (W_{FA}). This yields that if we set $W_H = X_1$, $W_M = X_2$, $W_{CR} = Y_1$, $W_{FA} = Y_2$, ($X_1 \neq X_2$, $Y_1 \neq Y_2$), then the same best operational level maps will appear for $W_H = X_2$, $W_M = X_1$, $W_{CR} = Y_1$, $W_{FA} = Y_2$, $W_H = X_2$, $W_M = X_1$, $W_{CR} = Y_2$, $W_{FA} = Y_1$, $W_H = X_1$, $W_M = X_2$, $W_{CR} = Y_2$, $W_{FA} = Y_1$. Note that each set represents a different system. An example of the last finding is presented in Table II.

These results led to further analysis, aimed at testing the shift in the best operational level maps, which occurs in the transition between symmetrical weights. This revealed that for all the analyzed systems the same best operational level map was received, i.e., the symmetry attribute depends on the difference between the weights. In fact, this further analysis extended the symmetry attribute, and it was found

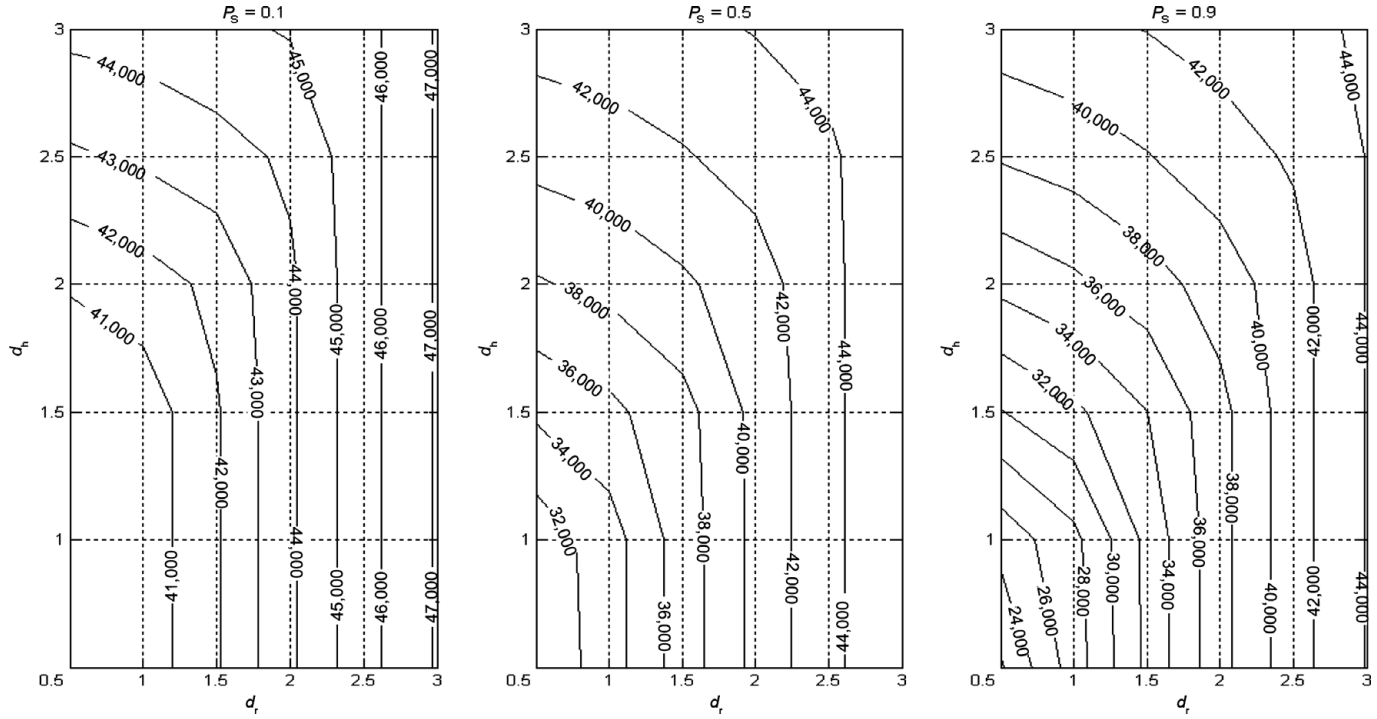

Fig. 5. Example of a *Type 1* objective function graph where the system's goal is achieved by giving high reward for CR proportionally to FA.

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

Table II. Examples of different systems with the same best operational level maps.

W_H	W_M	W_{CR}	W_{FA}
20	5	10	7
5	20	10	7
20	5	7	10
5	20	7	10

that the type (*1* or *2*) depends upon the ratio between $\Delta_1 = W_{CR} - W_{FA}$ and $\Delta_2 = W_H - W_{M.}$. If $\Delta_1/\Delta_2 > 1$, the system reaction is classified as *Type 1*. If $\Delta_1/\Delta_2 < 1$, the system reaction is classified as *Type 2*. In cases where $\Delta_1/\Delta_2 = 1$, the system reaction can be considered as *Type 1* or *Type 2*. This can be assumed as a new property of the objective function. These findings are presented in Fig. 6.

All analyses were conducted with optimal β_s presented in Eq. 8.

$$\beta^* = \overbrace{\frac{(1-P_S)}{P_S}}^{A} \times \overbrace{\left(\frac{W_{CR}-W_{FA}}{W_H-W_M}\right)}^{B} \tag{8}$$

The optimal β includes two parts (Eq. (8)[22]): the ratio between the *a priori* probabilities (noted as A) and the ratio between the weights (Δ_1/Δ_2, noted as B). Analysis of the objective function and Eq. (8) revealed that the value of optimal β is determined by the system type (i.e., Δ_1/Δ_2) and the target probability (P_S). Parameter P_S depends on the environment and cannot be controlled by the system's designers. Thus, this parameter has not been considered as a parameter that defines the system reaction type, but it does affect the system performance. On the other hand, the ratio between the weights (Δ_1/Δ_2) depends on the system reaction and the tasks and can be controlled by the system designers.

5. Operational Cost Analysis

The operational cost analyses were conducted for systems reacting to *Type 1* and *Type 2* tasks where the weight ratios were set to $\Delta_1/\Delta_2 = 10$ and $\Delta_1/\Delta_2 = 0.1$, respectively.

5.1. System time (t_s) analysis

The system time is affected by the target probability (P_S), and the human and robot sensitivities (d'_r and d'_h), thresholds (d'_s), and reaction times (see Section 2.2, Eq. (7)). The robot's reaction time was considered negligible (since it depends on computer hardware and algorithm complexity). The robot time is considered to be deterministic. On the other hand, the human decision time, t_d, depends on many variables, such as the operator's skills and fatigue and image complexity, and therefore is considered variable.

The analysis focused on the effects of each of the parameters influencing t_s, which were described above: P_S, d'_r, d'_h, and t_d. *Type 1* analysis is presented in Fig. 7. These

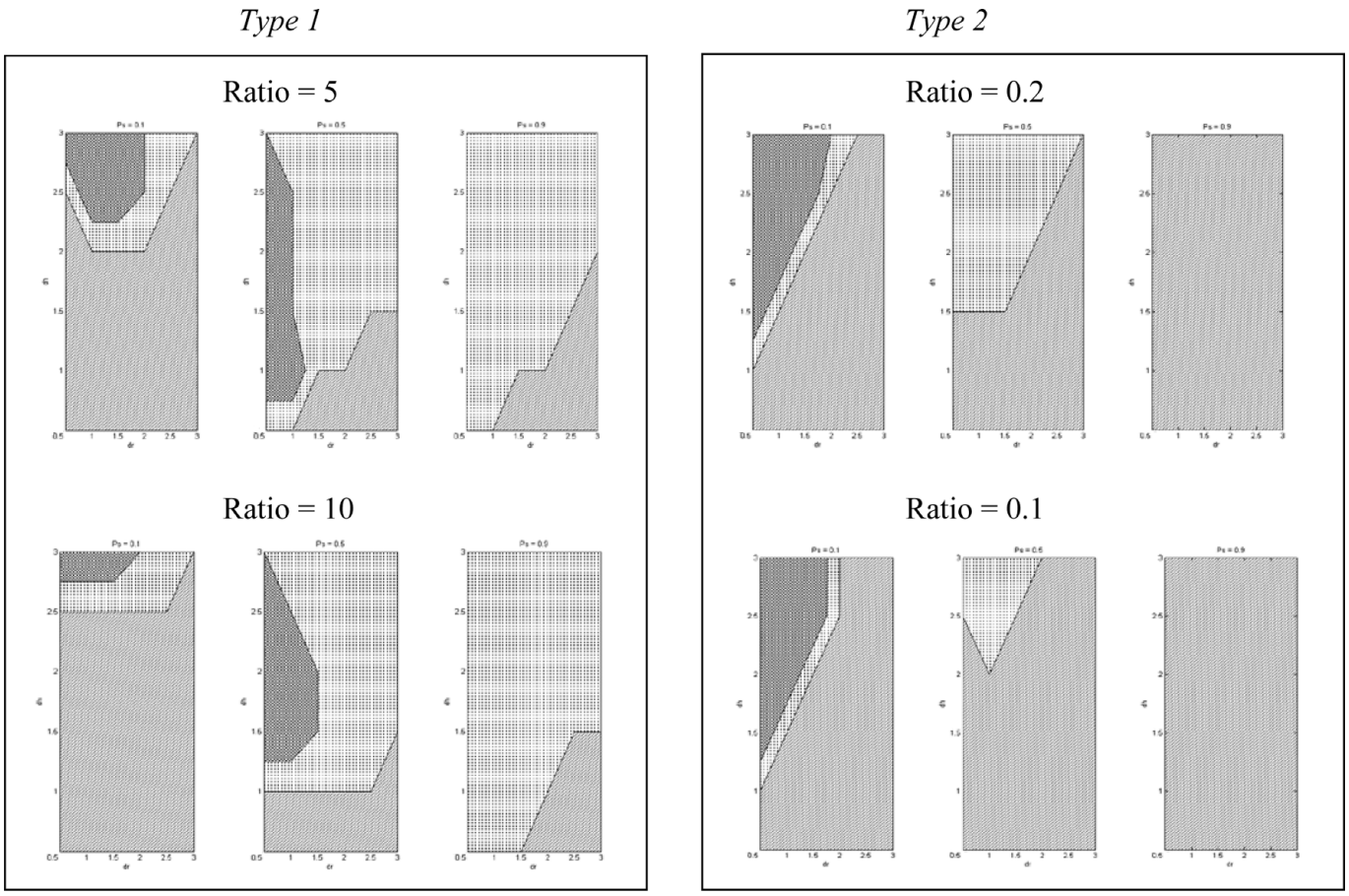

Fig. 6. Example of the symmetry property. On the left graphs, $\Delta_1/\Delta_2 > 1$ (i.e., *Type 1*). On the right graphs, $\Delta_1/\Delta_2 < 1$ (i.e., *Type 2*). Different ratios will produce different best operational level maps.

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

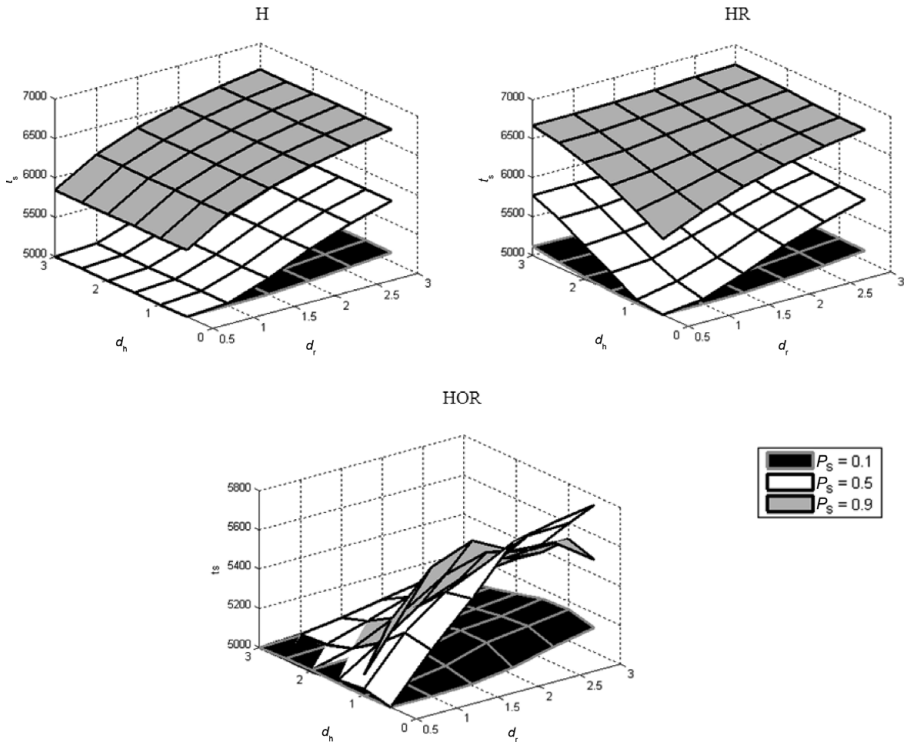

Fig. 7. Time cost analysis of *Type 1*. The effect of different sensitivities and target probabilities on system time.

graphs demonstrate the effect of target probability and human and robot sensitivities on each of the operational levels. As explained previously, robot time is treated as a constant and therefore level R is not presented. Each surface represents a different target probability, *x*- and *y*-axes represent the robot and human sensitivities, respectively, and the *z*-axis is the system time (t_s).

In the H operational level, the "hit" and "false alarm" outcomes are associated with longer reaction times than miss and correct rejection, since these outcomes also involve a motoric motion. Results show that for level H, as target probability and human sensitivity increase, t_s becomes higher. These results are supposedly counter-logical, as one would expect that as human sensitivity rises, his/her discerning ability improves, causing decision-making time to shorten, and therefore system time to diminish. Actually, as human sensitivity improves, the probability of recognizing a target (probability for hit) rises, which leads to a larger number of longer reactions, and therefore the total time the system operates is longer.

The difference between the HR and the HOR operational levels in the model's objective function manifests itself by the t_s value. Thus, the graphs of these operational levels are different. In the HR level, the robot first marks objects it considers being targets and then the human decides whether these objects are targets. Thus, similar to the H level, hit and false alarm outcomes are associated with longer reaction times than miss and correct rejection. Therefore, the HR level behaves similarly to the H operational level. In the HOR level, the robot first marks all recognized targets; then the human marks objects the robot has missed, and rejects wrong entries of the robot. Thus, in the HOR operational level, contrary to the H and HR levels, miss and correct rejection outcomes are associated with longer reaction times than hit and false alarm. Consequently, the HOR level has opposite trends than the H and HR levels.

The results for *Type 2* are presented in Fig. 8. Similar to the results of *Type 1*, as target probability increases, the value of t_s increases in the H and HR levels. However, contrary to the results of *Type 1*, as human and robot sensitivities increase, t_s decreases in these operational levels. This result is derived from the difference between *Type 1* and *Type 2*. In *Type 1*, an increase in the system's (human and robot) sensitivity mostly influences the number of hits, which increase as well and therefore t_s increases. However, in *Type 2*, an increase in the system's sensitivity mostly influences the number of false alarms, which decrease, thus t_s decreases.

The HOR graph behaves similarly to the graph of *Type 1* except that the surfaces of $P_S = 0.1$ (black surface) and $P_S = 0.9$ (gray surface) are opposite. This result reflects the opposite tendencies of the two types as a function of the target probability.

5.2. *Human decision time (t_d) analysis*

In order to investigate the influence of human decision time, t_d, on system performance, its value was varied from 2 to 14 s. *Type 1* analysis for the best operational level maps as a function of P_S, t_d, d'_h and d'_r is presented in Fig. 9.

For low target probabilities ($P_S = 0.1$), a change in decision time is critical and for values equal and above $t_d = 8$ s the best operational level is always R level. In contrast, for medium (0.5) or high (0.9) target probabilities, a change in decision time has only low influence on the best operational level maps. In order to compare the influences of the t_d and P_S, further analysis was performed, which examined these parameters together (Fig. 10).

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

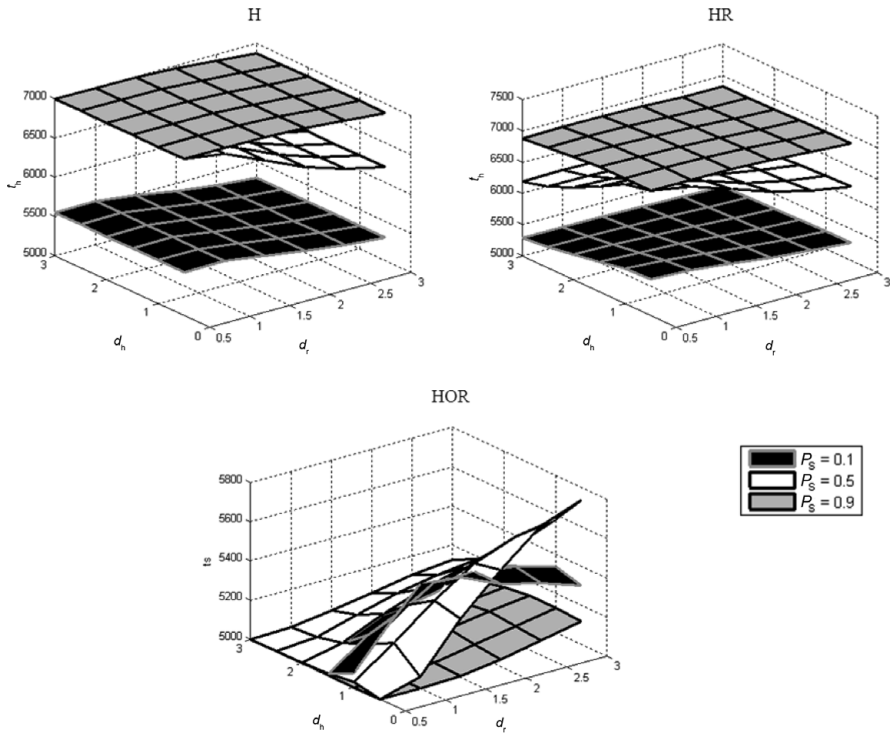

Fig. 8. Time cost analysis of *Type 2*. The effect of different sensitivities and target probabilities on system time.

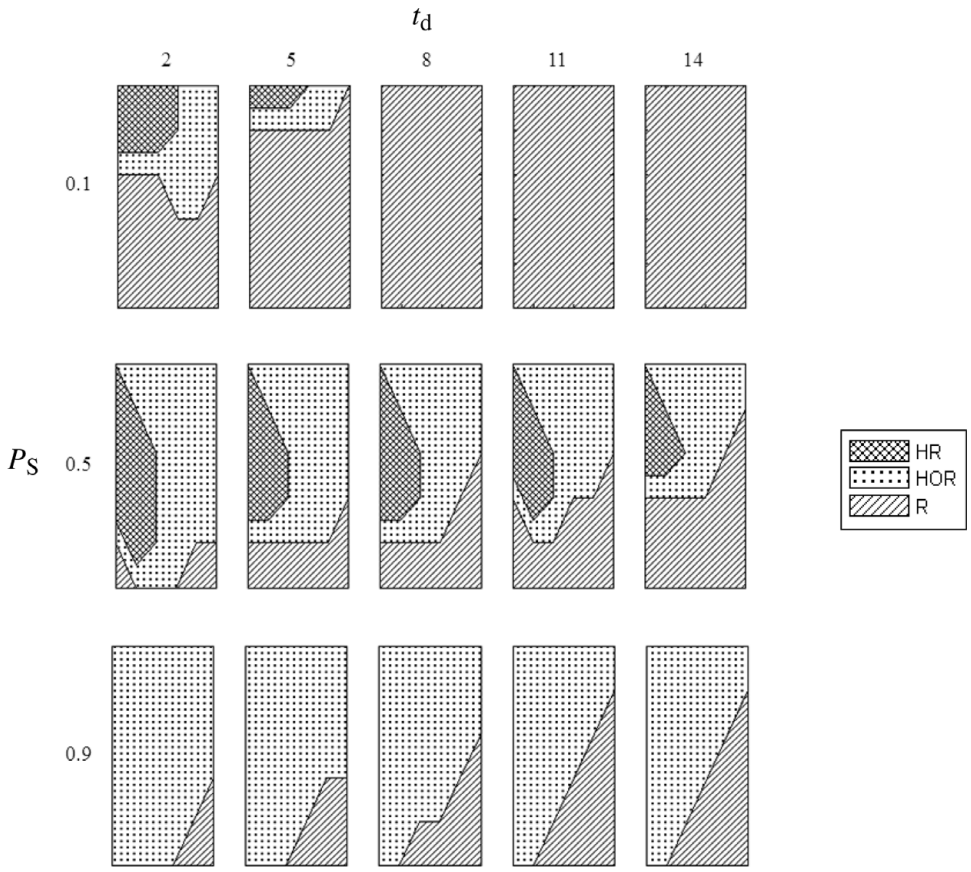

Fig. 9. *Type 1* results for analysis of best operational level maps for different t_d and P_S. Columns are for different t_d and rows for different P_S. Each map is represented in the sensitivity space where the x-axis is d_r' and the y-axis is d_h'. For instance, the upper left graph presents the best operational level of *Type 1* system reaction as a function of robot (x-axis) and human (y-axis) sensitivities where $t_d = 2$ and $P_S = 0.1$.

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

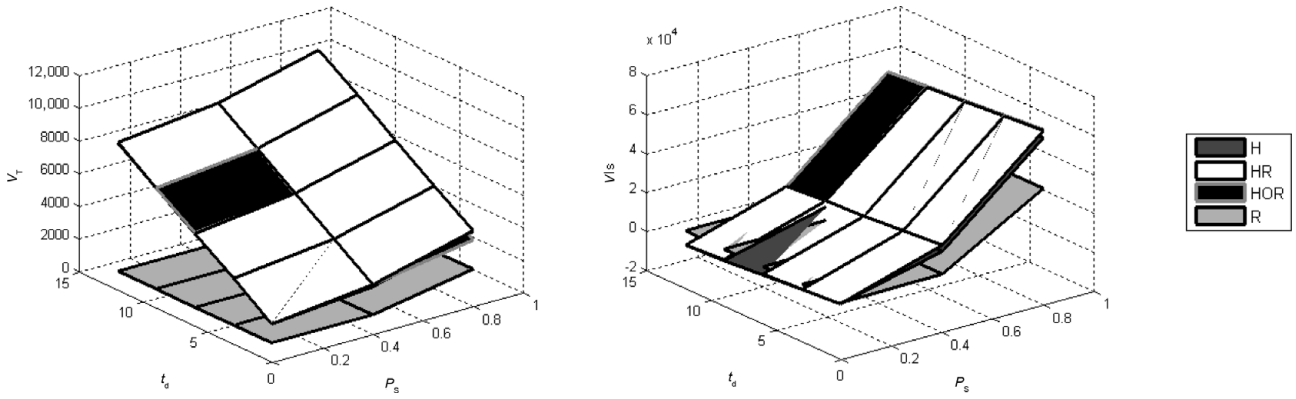

Fig. 10. Decision time analysis of *Type 1*. Effect of t_d and P_S on V_T (left graph) and V_{Is} (right graph).

The effect of P_S and t_d on the objective function score (V_{Is}, right graph) and on the operational costs (V_T, left graph, values are absolute) are shown in the figure. Results show that the effect of decision time, t_d, on V_T is larger than the effect on P_S, i.e., an increase in this parameter leads to a sharper increase in operational costs. In contrast, P_S is the parameter that has a larger effect on the total score, V_{Is}. Thus, t_d's level of influence on the system's overall performance depends upon the ratio V_T/V_{Is}. As this ratio increases, t_d is relatively more effective than P_S. For both graphs t_d has an identical effect on levels H, HR, and HOR. *Type 2* results are presented in Figs. 11 and 12.

For low (0.1) and medium (0.5) target probabilities, a change in decision time has a greater influence on the best operational level maps. As t_d increases, the R level is more preferable.

Similar to the results of *Type 1*, while t_d has a larger effect than P_S on V_T, the target probability has a larger influence than t_d on the total score, V_{Is}. Thus, t_d's effect is independent of task type.

5.3. *Object recognition weight (W_c) Analysis*

The cost associated with object recognition, W_c, depends, among other things, on the type of assignments the system has to perform and the robotic system's complexity. In order to examine the effect of W_c, this parameter received variable values, ranging from 2 to 18. *Type 1* analysis of the best operational level maps for different values of W_c and different target probabilities is presented in Fig. 13.

Changes in W_c have little influence on the best operational level maps for all target probabilities. Similar to the time cost analysis, for a further examination of W_c, its influence was

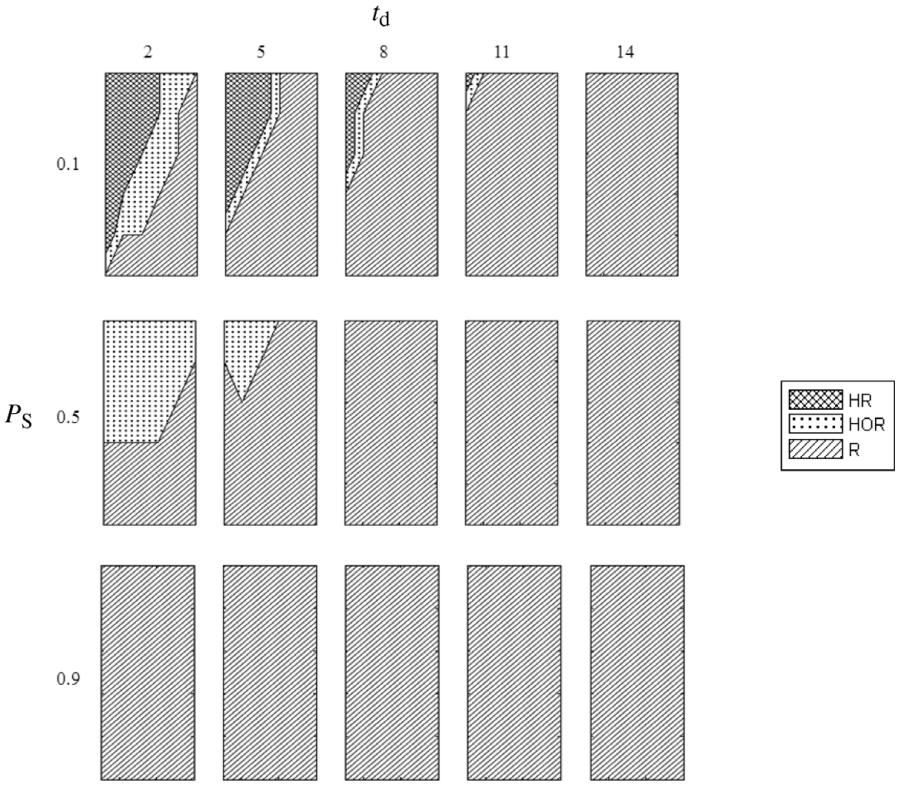

Fig. 11. *Type 2* results for analysis of best operational level maps for different t_d and P_S. Columns are for different t_d and rows for different P_S. Each map is represented in the sensitivity space where the x-axis is d'_r and the y-axis is d'_h.

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

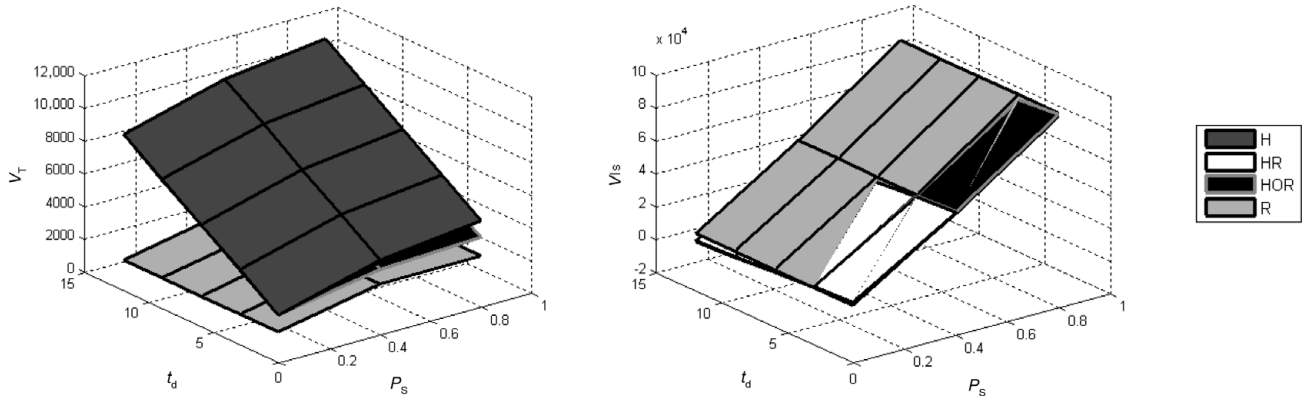

Fig. 12. Decision time analysis of *Type 2*. Effect of t_d and P_s on V_T (left graph) and V_{Is} (right graph).

compared with the influence of P_S on the operational costs and on the total objective function score (Fig. 14).

Increase of W_c has little influence on the operational costs (Fig. 14, left) and total system performance (Fig. 14, right) for all models' operational levels. However, a combination of high P_S and high W_c increases the operational costs dramatically. Greater target probability increases W_a (number of hits and false alarms), which is the factor that is multiplied by W_c. Hence, the combination of high P_S and high W_c has a great influence on the operational costs (Fig. 14, left). W_c's analysis for "*Type 2*" is presented in Figs. 15 and 16.

These results correspond to those of *Type 1*. Changes in W_c have only a slight effect on the best operational level maps for all the target probabilities. These findings match the results of *Type 1*. Thus, the influence of W_c is independent of the type.

6. Conclusions

An investigation of the behavior of the objective function consisting of all modules (hit, false alarm, miss, correct rejection, and operational cost) was conducted. Different types of tasks and systems were evaluated, resulting in two behavior types based on the natural qualitative bipartition of model behaviors: *Type 1* is focused on minimizing false alarms and *Type 2* is geared toward recognizing targets. These system reactions and tasks can be quantified in terms of a simple ratio of differences of model weights. The objective function has an inherent symmetry with respect to the weight pairs and ratios; this symmetry was shown analytically. The symmetry is expressed by the fact that systems reacting to different tasks have identical best operational level maps. This finding led to a further investigation, which indicated that only the ratio Δ_1/Δ_2, where $\Delta_1 = W_{CR} - W_{FA}$ and

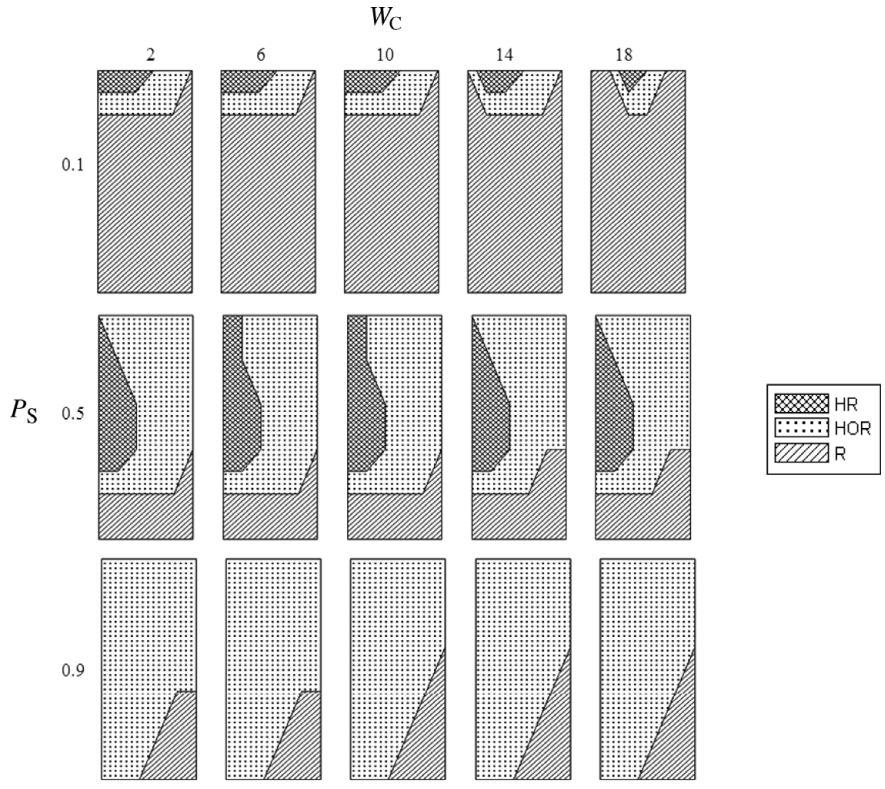

Fig. 13. *Type 1* analysis of the best operational level maps for different W_c and P_S. Columns are for different W_c and rows for different P_S. Each map is represented in the sensitivity space where the x-axis is d'_r and the y-axis is d'_h.

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

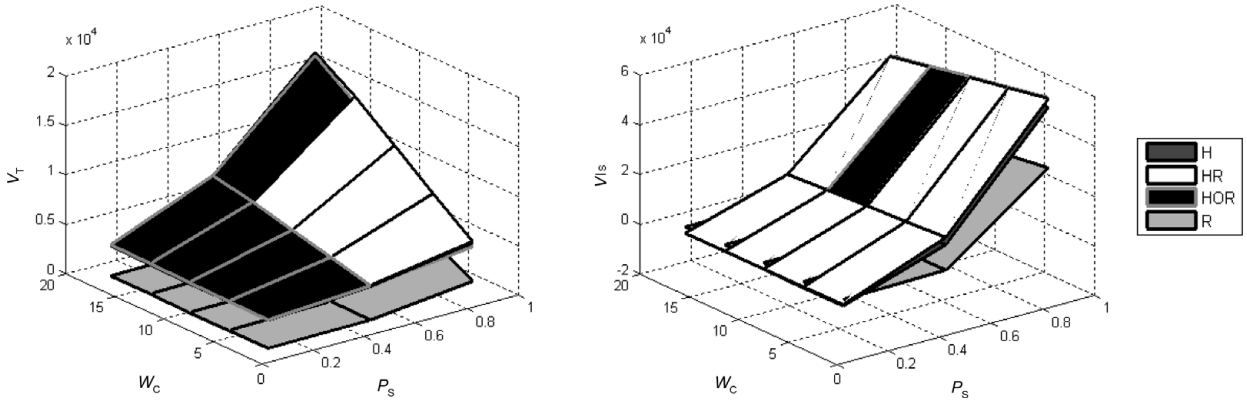

Fig. 14. Effect of W_c and P_S on V_T (left graph) and V_{Is} (right graph) of *Type 1* system reaction.

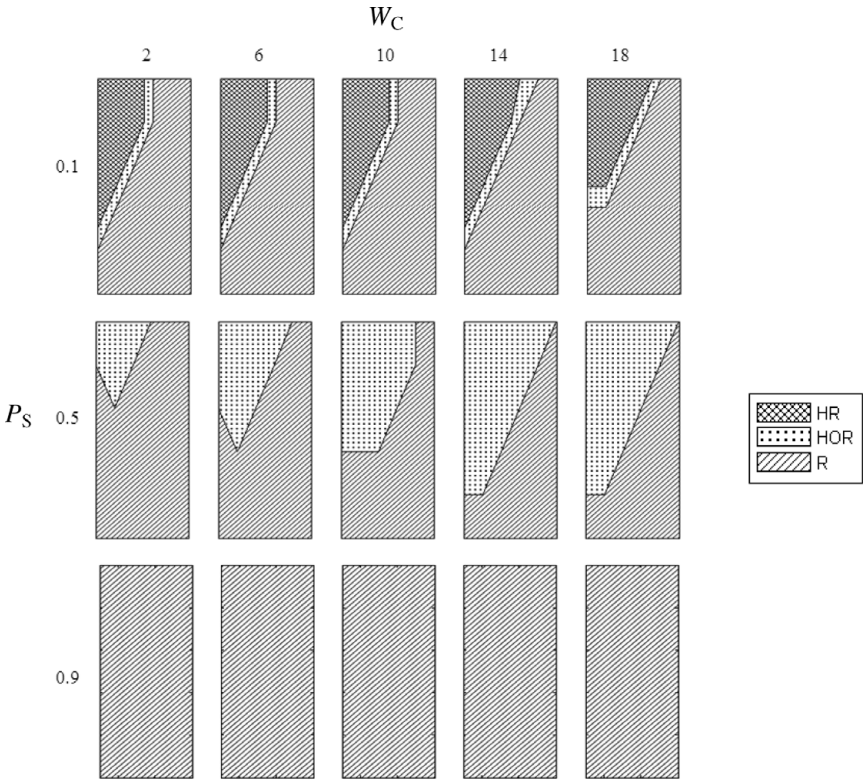

Fig. 15. *Type 2* analysis of the best operational level maps for different W_c and P_S. Columns are for different W_c and rows for different P_S. Each map is represented in the sensitivity space where the x-axis is d_r' and the y-axis is d_h'.

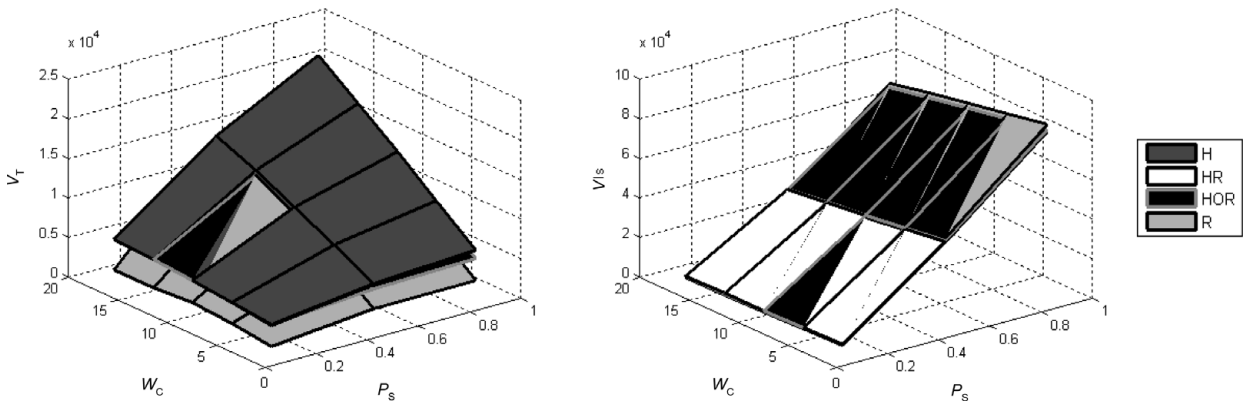

Fig. 16. Effect of W_c and P_S on V_T (left graph) and V_{Is} (right graph) of *Type 2* system reaction.

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press

$\Delta_2 = W_H - W_M$, determines the type of the task and the system reaction. In addition, since the *a priori* probability (P_S) is a characteristic of the environment and cannot be controlled by the system's designers, this parameter has *not* been considered as a parameter that defines the type.

The last finding can enable generalization of the results that were found for specific tasks and systems in this work to any task and system with the same ratio. Moreover, this finding can be assumed to be a new property of the objective function score and assist in the understanding of further analysis.

The human operator decision time was found to be the most influential parameter on the best operational level map. Improvement in the human decision time (i.e., shortening this time) leads to preference of the HR and HOR collaboration levels over level R. Hence, research should be directed to methods to reduce human decision time.[23,24]

The manual level (H) was never the best level for the optimal cases; this may be the result of high operational costs and a relatively low recognition rate in the systems analyzed. This implies that optimal collaboration between human and robot in target recognition tasks, with similar conditions, can improve system performance. It appears that improvement in recognition rate and hence rise in profits gained by this collaboration, outweighs the rise in operational cost attributable to adding the robot to the system.

Best operational level results showed opposite tendencies between the two types found. In system reaction to *Type 1* tasks, as target probability increased, the R level was preferable in most cases. In system reaction to *Type 2* tasks, the trend was reversed: as target probability increased, the HR and HOR collaboration levels were preferable in most cases. *Type 1* system reaction places great value on not committing errors, that is to say, they place great importance on results in situations where no target is present, or target probability is low. In turn, *Type 2* system reaction places great value on results in which a target is present. Even though very different tendencies were discovered by the functional analysis, several important similar tendencies found between them should be pointed out: in both systems as the probability of the prominent object (non-target in *Type 1* and target in *Type 2*) increases, the R level will be preferable in more cases. It can be assumed that this trend stems from the reciprocation between operational costs and recognition profits.

This research provides tools to develop an integrated human–robot target recognition system. The system can be designed to fit a specific task and environment. The work presented here can be applied off-line and even in the absence of an actual system. For instance, in real-time systems, recognition time is crucial, the cost of each time unit will be high, and the system will be allowed to operate only in the R and HOR collaboration levels. On the other hand, in detecting land mines, the goal is a high hit rate, while the false alarm rate and recognition times are usually less important; the reward for hit will be high and the cost of each time unit and penalty for false alarm will be relatively low, and the system designers will prevent the system from operating in collaboration levels that will impair the performance such as in the R collaboration level. Furthermore, this methodology can be used to analyze on-line system performance and to recommend the best collaboration and the human performance on-line.[25]

Acknowledgments

This research was partially supported by the Paul Ivanier Center for Robotics Research & Production Management and by the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering, Ben-Gurion University of the Negev.

References

1. S. Everett and R. Dubey, "Human–machine cooperative telerobotics using uncertain sensor and model data," **In**: *Proceedings of the IEEE International Conference on Robotics and Automation*, Leuven, Belgium (May 16–20, 1998) pp. 1615–1622.
2. A. Bechar and Y. Edan, "Human–robot collaboration for improved target recognition of agricultural robots," *Ind. Robot* **30**(5), 432–436 (2003).
3. F. Cremer, K. Schutte, J. G. M. Schavemaker and E. den Breejen, "A comparison of decision-level sensor-fusion methods for anti-personnel landmine detection," *Inf. Fusion* **2**(3), 187–208 (2001).
4. Y. C. Chang, G. S. Song and S. K. Hsu, "Automatic extraction of ridge and valley axes using the profile recognition and polygon-breaking algorithm," *Comput. Geosci.* **24**(1), 83–93 (1998).
5. L. F. Penin, R. Aracil, M. Ferre, E. Pinto, M. Hernando and A. Barrientos, "Telerobotic System for Live Power Lines Maintenance: ROBTET," **In**: *Proceedings of the IEEE International Conference of Robotics and Automation*, Leuven, Belgium (May 16–20, 1998) pp. 2210–2115.
6. Z, Sun, G. Bebis and R. Miller, "Object recognition using feature subset selection," *Pattern Recognit.* **37**(11), 2165–2176 (2004).
7. H. Medicherla and A. Sekmen, "Human–robot interaction via voice-controllable intelligent user interface," *Robotica* **25**, 521–527 (2007).
8. T. B. Sheridan, *Telerobotics, Automation, and Supervisory Control* (MIT Press, Cambridge, MA, 1992).
9. J. Scholtz, "Theory and Evaluation of Human Robot Interaction," *Proceedings of the IEEE Hawaii International Conference on System Sciences*, Hawaii (Jan. 6–9, 2003).
10. A. M. Howard, "Role Allocation in Human–Robot Interaction Schemes for Mission Scenario Execution," **In**: *Proceedings of the IEEE International Conference on Robotics and Automation*, Orlando, Florida (May 15–19, 2006) pp. 3588–3594.
11. A. Bechar, Y. Edan and J. Meyer, "Optimal Collaboration in Human–Robot Target Recognition Systems," **In**: *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Taipei, Taiwan (Oct. 8–11, 2006) pp. 4243–4248.
12. A. Bechar, J. Meyer and Y. Edan, "An objective function to evaluate performance of human–robot collaboration in target recognition tasks," *IEEE Trans. Syst. Man Cybern. Part C* **39**(6), 611–620 (2009).
13. Y. M. Ye and J. K. Tsotsos, "Sensor planning for 3D object search," *Comput. Vis. Image Underst.* **73**(2), 145–168 (1999).
14. E. Bicho, P. Mallet and G. Schoner, "Target representation on an autonomous vehicle with low-level sensors," Int. J. Robot. Res. **19**(5), 424–447 (2004).
15. M. H. Zweig and G. Campbell, "Receiver-operating characteristic (ROC) plots – a fundamental evaluation tool in clinical medicine," *Clin. Chem.* **39**(4), 561–577 (1993).
16. P. F. Griner, R. J. Mayewski, A. I. Mushlin and P. Greenland, "Selection and interpretation of diagnostic tests and procedures. Principles and applications," *Ann. Intern. Med.* **94**(4), 557–592 (1981).
17. A. Steinfeld, "Interface lessons for fully and semi-autonomous mobile robots," *Proc. IEEE Int. Conf. Robot. Autom.* **3**, 2752–2757 (2004).

18. D. Shin, R. A. Wysk and L. Rothrock, "A formal control-theoretic model of a human-automation interactive manufacturing system control," *Int. J. Prod. Res.* **44**(20), 4273–4295 (2006).
19. A. Bechar, "Human–Robot Collaboration Methods for Target Recognition in Unstructured Environments," *Ph.D. Dissertation* (Ben-Gurion University of the Negev, Beersheba, Israel, 2006).
20. K. Deb, L. Thiele, M. Laumanns and E. Zitzler, "Scalable multi-objective optimization test problems," **In**: *Proceedings of the 2002 Congress on Evolutionary Computation 1*, Honolulu, HI, USA (May 2002) pp. 825–830.
21. J. P. Ko and J. M. Lee, "Tactile tele-operation of a mobile robot with a collision vector," *Robotica* **24**, 11–21 (2006).
22. J. A. Swets, R. M. Dawes and J. Monahan, "Better decisions through science," *Sci. Am.* **283**(4), 82–87 (2000).
23. M. Maltz and J. Meyer, "Use of Warnings in an attentionally demanding detection task," *Hum. Factor* **43**(2), 217–226 (2001).
24. M. Maltz and D. Shinar, "New alternative methods of analyzing human behavior in cued target acquisition," *Hum. Factors* **45**(2), 281–295 (2003).
25. I. Tkach, A. Bechar and Y. Edan, "Switching between collaboration levels in a human–robot target recognition system," *IEEE Trans. Syst. Man Cybern. Part C* (accepted).

https://doi.org/10.1017/S0263574711001020 Published online by Cambridge University Press