VIEW-FINDER : Robotics Assistance to fire-Fighting services

Y. Baudoin, D. Doroftei, G. De Cubber, S.A. Berrabah, E. Colon, C. Pinzon
Royal Military Academy, Polytechnics
30 Av de la Renaissance, 1000 Brussels

A.Maslowski, J.Bedkowski
Research Institute of Automation and Control, Intelligent Mobile Systems Division
Aleje Jerozolimskie 202
Warsaw, POLAND

J.Penders
Sheffield Hallam University, UK

In the event of an emergency due to a fire or other crisis, a necessary but time consuming pre-requisite, that could delay the real rescue operation, is to establish whether the ground or area can be entered safely by human emergency workers. The objective of the VIEW-FINDER project is to develop robots which have the primary task of gathering data. The robots are equipped with sensors that detect the presence of chemicals and, in parallel, image data is collected and forwarded to an advanced Control station (COC). The robots will be equipped with a wide array of chemical sensors, on-board cameras, Laser and other sensors to enhance scene understanding and reconstruction. At the control station the data is processed and combined with geographical information originating from a web of sources; thus providing the personnel leading the operation with in-situ processed data that can improve decision making. The information may also be forwarded to other forces involved in the operation (e.g. fire fighters, rescue workers, police, etc.). The robots will be designed to navigate individually or cooperatively and to follow high-level instructions from the base station. The robots are off-the-shelf units, consisting of wheeled robots. The robots connect wirelessly to the control station. The control station collects in-situ data and combines it with information retrieved from the large-scale GMES-information bases. It will be equipped with a sophisticated human interface to display the processed information to the human operators and operation command.

1. The Crisis Management Information System (CMIS) [1,2]

The architecture of the CMIS is composed of two high level conceptual entities communicating with each other: the appliances (robots) performing their mission inside the crisis area and the combination of the base station (BS) and the Crisis Management Information System (CMIS), located close (if not at the same place) to the Control Operation Center (COC). The base station is in charge of managing the robots. The CMIS takes the information produced by the robots and the sensors they carry, and relayed through the BS, and provide an interface to this information for the crisis managers.

The following picture depicts the components of the system and the relationship between them: the Control Operational Center (COC) communicates with the robots, which process the data received from the sensors

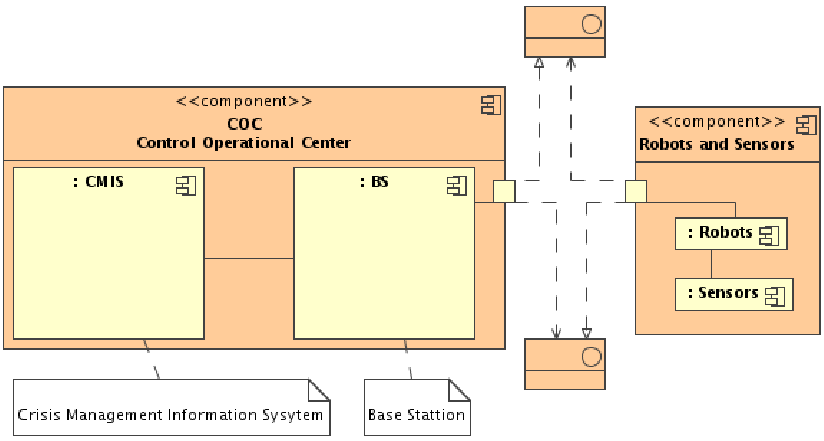

Figure 1: VF architecture

The View-Finder project, along with its appliances (robots), develops a crisis management system to support the actors involved in the crisis.

The crisis management information system (CMIS) developed in the context of View-Finder, must cover the whole Crisis Life Cycle (CLC) and therefore combines the following two components:

- the Disaster Management Action Plan (DMAP),
- the Crisis Management System (CMS)

A layered architecture of the CMIS is depicted in the following figure:

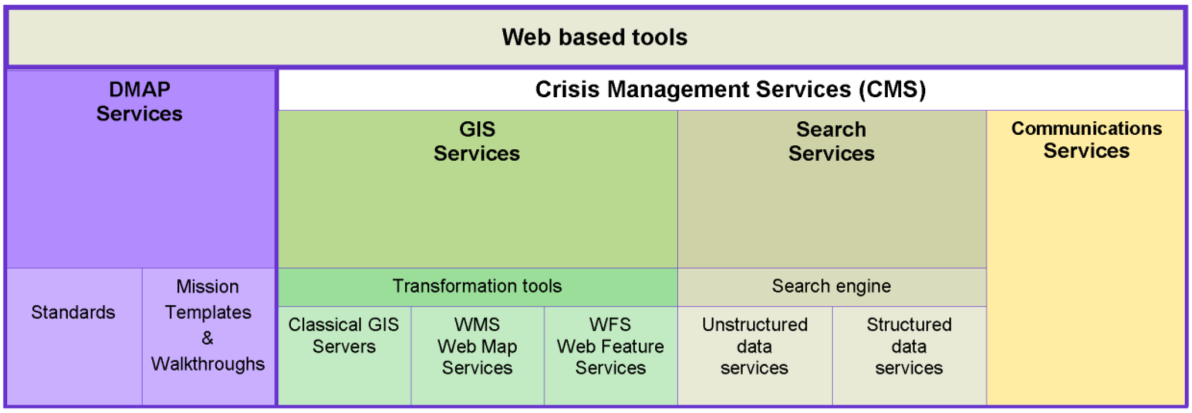

Figure 2: Layered architecture of the CMIS

1.1. *The DMAP*

The DMAP is a conceptual crisis management tool which shall be capable to provide specific answers at any stage of the crisis, for any kind of emergency situations, from large-scale national and cross-border crises, regional disasters down to local-scale incidents. The DMAP will provide to the end users three types of information:

1. Standards and terminology: to facilitate the communication in heterogeneous contexts;
2. Crisis events walkthroughs: provide the end users with access to the recommended methods to manage this particular crisis based on pre-recorded crisis events;
3. Mission templates: based on the material gathered from users, this tool gives a set of pre-compiled templates to facilitate the characterisation and follow-up of a crisis.

The DMAP features are available all along the utilisation of the CMS and can be invoked in particular situations

1.2. *The CMS*

The CMS integrates three services:

- Geographic information system (GIS): thematic maps will be provided in a specific Geographic Information System (GIS). The sources of these maps are heterogeneous (other GIS, data stored in common media support, databases); Modules for the processing of these data are foreseen (GIS Servers, WMS, WFS) as well as a module which will allow transforming the retrieved data in a seamless way.

- Search: A multi-criterion search tool will be available in order to give quick access to the data available in the system; these data could include information in many formats, for instance images retrieved from the GIS, text retrieved from legal documents or semi-structured information (xml), raw data in provenance of other databases and services.

- Communications: This module stores and makes available the gathered data to decision makers at different levels and provides to the system users, the authentication methods and tools which will allow them to keep informed about the crisis events and to keep in contact

2. The User’s requirements and scenarios [3]

On site of the incident, the Control Operational Centre will be deployed if the situation implies it (decision of the Fire-Fighting Operation and/or Incident Officer) and for so far the Robotics assistance fits the next basic requirements:

- Improvement of the **Security/Safety** of the Intervention Team
- **Rapid** Deployment of Intervention Team
- Operation in **Contaminated** Environments
- Implementation of Reliable Chemical **Detection** means
- Hazard Prediction Modelling (Mapping/Training/Testing through **Simulation** Tools)
- **Secure** visual **Data (**environment**, victims,** etc)

A typical scenario has consequently been proposed by/with the support of the End-Users:
The crash of an airplane on/near a military/civilian zone with the support of the 1 Wing Air Base (located in Belgium), as illustrated by the fig 3

Fig 3 Wing Intervention Area

3. The Robots [4]

Two robots will be used (from PIAP, Poland and RMA, Belgium): fig 4

		
Red ATRV Jr (PIAP)	Aircaft Crash Intervention	ROBUDEM (RMA)

Fig.4: the robots of the V-F project

The ARTVJr autonomous mobile robot has been improved for the ViewFinder Application from the software architecture and autonomous navigation point of view. The improvement [15] is based on the investigation of the State of the Art hardware drivers for ATRV Jr functionality and its implementation. The need of research is motivated by the incompatibility between ATRV Jr Mobility software and View Finder System components. The main goal of increased performance is achieved by combining State of the Art Player/Stage driver for the RFLEX–ATRV Jr onboard real time controller with ViewFinder system based on CORBA/CoRoBa/Mailman communication technologies. The usage of player’s driver determines independence of the ViewFinder System from indoor autonomous mobile platform. It means the robot is seen by the ViewFinder System as a set of functionalities, therefore another robot with respective functionalities can be used. Player server provides access to the robot basic functionality given by RFLEX real time OS, which consists of sensors: sonars, odometry, battery status, ptz camera, compass. The gateway from CORBA to Player is developed, therefore whole functionality of the mobile platform is available from CORBA components. For the purpose of robust wireless communication, the gateway between CORBA and Mailman is implemented. Therefore the robot is integrated with ViewFinder system via Mailman wireless communication.

The ROBUDEM is equipped with odometric sensors, a vision system consisting of a Bumblebee2 vision system consisting of 2 digital cameras, an RTK DGPS with three antennas, INS and ultrasonic sensors which give distance measurements to obstacles in front of the robot and before the wheels.

4. The mission planning and related command-control of the Robots

In the course of the View-Finder project, two robotics teams (RMS and PIAP) are working on the development of an intelligent autonomous mobile robot. The robot needs some degree of self-consciousness, meaning that it needs to be able to infer its current status in relation to the outside world from its sensor readings. This problem is also referred to as the Simultaneous Localization and Mapping (SLAM) problem. Classical SLAM solving techniques use input data from DGPS, Laser range scanners or ultrasound sensors, yet also visual data can be used. Furthermore, the robot must dispose of some sort of intelligence to execute the tasks or objectives it has been given. Last but not least, common frameworks had to be chosen in order to ease the communication between the three robots.

4.1. Framework: COROBA, MAILMAN [5]

As distributed hybrid robotics systems are becoming more complex and cooperating with each other, there is a need to promote the construction of new systems as composition of reusable building blocks. System modularity and interoperability are key factors that enable the development of reusable software. The V-F project includes the use of a designed generic control framework using CORBA as its communication middleware.

In distributed applications, programs need to invoke operations in other processes, often running in different computers. To achieve this and taking into account with the fact that our programs are written in object oriented languages, the following programming models have been chosen:

- The Remote Method Invocation (RMI) that allows objects in different processes to communicate
- The distributed Event-based programming model that allows objects to be notified when events they have registered interest in have been emitted.

Inspired by classical control applications, components have been divided in three categories: sensors, processors and actuators (Figure 5). Sensors have connections with the physical world and they output data to one Event Channel. Processors get their inputs from one Event Channel, they transform data and send the result to another Event Channel. Actuators have output connections with the physical world and received data from one Event Channel. Services can run on any machine in a network and are remotely managed by an administration application.

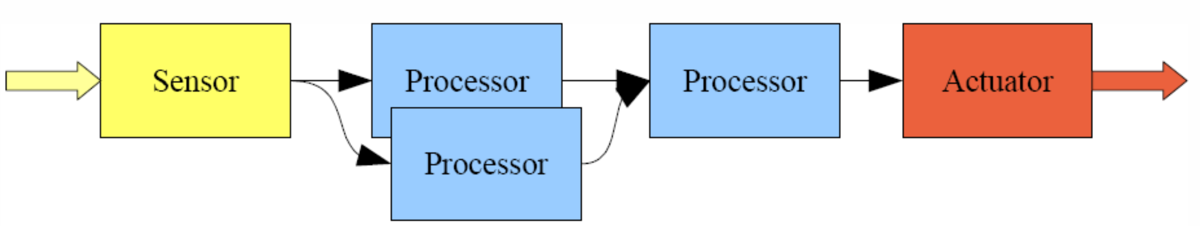

Fig. 5 Sequence of components

In the View-Finder, in addition to the classical Quality of service (QoS) criteria, the following are also important:

- Fair bandwidth allocation.
- Message prioritization.

The types of data that compose the traffic in the View-Finder network comprise multimedia (video) streams, sensor readings, control commands, status messages, navigation commands, and critical messages. Each of these data types has different characteristics and QoS requirements. Therefore the View-Finder Project uses the MAILMAN protocol communication. The reader may refer to [14].

4.2. Vision-based Simultaneous Localization and Mapping [6]

In the outdoor application the robot ROBUDEM uses a single monocular camera to extract natural features in the scene. These features are used as landmarks in the built map. The SLAM problem is tackled as a stochastic problem using an Extended Kalman Filtering to maintain a state vector, **X**, consisting of the robot state, $\mathbf{y_R}$, and map feature states, $\mathbf{x_{Li}}$. It also maintains a covariance matrix, **P**, which includes the uncertainties in the various states as well as correlations between the states.

A world coordinate frame W is defined such that its X and Z axes lie in the ground plane, and its Y axis point vertically upwards. The system state vector $\mathbf{y_R}$ in this case is defined with the 3D position vector (y_1,y_2,y_3) of the gravity center of the robot in the world frame coordinates and the robot's orientation roll, pitch and yaw about the *Z*, *X*, and *Y* axes, respectively $(\gamma, \theta, \varphi)$.

The dynamic model or motion model is the relationship between the robot's paste state, $\mathbf{y}_R^{t-1}$, and its current state, $\mathbf{y}_R^t$, given a control input u^t

$$\mathbf{y}_R^t = \begin{bmatrix} y_1^t \\ y_2^t \\ y_3^t \\ \gamma^t \\ \theta^t \\ \varphi^t \end{bmatrix} = \mathbf{f}(\mathbf{y}_R^{t-1}, u^t, \mathbf{v}^t) = \begin{bmatrix} y_1^{t-1} + (\mathbf{v}^{t-1} + \mathbf{V})\cos(\gamma^{t-1})\Delta t \\ y_2^{t-1} + (\mathbf{v}^{t-1} + \mathbf{V})\sin(\gamma^{t-1})\Delta t \\ y_3^{t-1} \\ \gamma^{t-1} + (\omega^{t-1} + \mathbf{\Omega})\Delta t \\ \theta^{t-1} \\ \varphi^{t-1} \end{bmatrix}$$

Where $\mathbf{f}$ is a function representing the mobility, kinematics and dynamics of the robot (transition function). $\mathbf{v}$ and w are a random vector describing the unmodelled aspects of the vehicle (process noise such as wheel sleep or odometry error).

Feature are selected using SIFT algorithm and are represented by their 3D position vectors. When a feature is first detected, measurement from a single camera position provides good information on its direction relative to the camera, but its depth is initially unknown. In our application, to estimate the 3D position of the detected features, we use an approach based on *epipolar geometry*. This geometry represents the geometric relationship between multiple viewpoints of a rigid body and it depends on the internal parameters and relative positions of the camera. Features are not deleted from the map when they leave the field of view, but remain in the map and can be re-observed when the camera moves back and they become visible again. In some cases it is necessary to delete features which are not being reliably matched on a regular basis: some features detected will be frequently occluded or may contain parts of objects at very different depths. These features will lead to failed correlation attempts and can be removed from the map automatically.

To avoid using outlier features, the moving object mask detected by the motion segmentation procedure introduced in [8] is used. Subsequently, during map building, the detected features on the moving parts are excluded.

For feature matching we used a measurement test based on the discrepancy between a predicted measurement that each feature would generate and an actual sensor measurement and the mahalanobis distance between features descriptors. The epipolar constraint is also taken into account in our application for feature matching.

The main open problem of the current state of the art SLAM approaches and particularly vision based approaches is mapping large-scale areas. Relevant shortcomings of this problem are, on the one hand, the computational burden, which limits the applicability of the EKF-based SLAM in large-scale real time applications and, on the other hand, the use of linearized solutions which compromises the consistency of the estimation process. To overcome these limitations, we proposed an approach to build a global representation of the environment based on several size limited local maps built using the previously described approach. Two methods for local map joining are proposed, the first method consists in transforming each local map into a global frame before to start building a new local map. While in the second method, the global map consists only in a set of robot positions where new local maps started (i.e. the base references of the local maps). In both methods, the base frame for the global map is the robot position at instant t_0. The obtained map can be superimposed to a satellite image of the navigated area by matching the GPS data corresponding to frame coordinates of the local maps.

4.3. Behaviour based navigation [9]

The control architecture describes the strategy to combine the three main capabilities of an intelligent mobile agent: sensing, reasoning and actuation. These three capabilities have to be integrated in a coherent framework in order for the mobile agent to perform a certain task adequately. To combine the advantages of purely reactive and planner-based approaches, our research work aims at implementing a hybrid control strategy which fuses a behaviour-based controller for autonomous navigation with automation aspects for gas-detection.

The performance of the behaviour-based controller depends on the implementation of the individual behaviours as well as on the method chosen to solve the behaviour fusion or action selection problem. The action selection problem can be formulated as a multiple objective decision making (MODM) problem.

Mathematically, a multi-objective decision problem can be represented in the following way:

$$\arg\max_{\mathbf{x}} \left[o_1(\mathbf{x}), \ldots, o_n(\mathbf{x}) \right]$$

Where $o_1(\mathbf{x}), \ldots, o_n(\mathbf{x})$ are a set of system objectives, tasks or criteria and where $\mathbf{x} = \left(x_1, \ldots, x_n \right) \in R^n$ is a n-dimensional decision variable vector. The degree of attainment of a particular alternative ***x***, with respect to the k^{th} objective is given by $o_k(\mathbf{x})$. $X \subseteq R^n$ defines the set of feasible alternatives. This problem is in the literature often known as the Vector Optimization Problem (VOP).

The approaches towards solving the vector optimization problem for action selection all have their advantages and disadvantages. Solving the VOP using reliability analysis has the big advantage of incorporating direct information from the system under control into the control process. On the other hand, this architecture does not offer a human decision maker the ability to interact with the decision process. As autonomous agents more and more have to interact with humans on the field, exchanging knowledge and learning from each other, this is a serious shortcoming. Common techniques for solving the VOP while taking into account a decision maker's preferences take into account these issues by offering a human operator the ability to input some objectives or ideal points. These approaches, however, suffer from the disadvantage that no reliability data from the sensing and other processes is taken into account while performing action selection. One could thus argue that while reliability analysis-based approaches are too robot centric, these second set of approaches is too human-centric.

In our project, we then decided to integrate the advantages of both theorems. This can be achieved by minimizing the goal programming and reliability analysis constraints in an integrated way, following [10]:

$$\arg\min_{\substack{\mathbf{x} \in X \\ w_i \in \mathbf{w}}} \left[\lambda \left(\sum_{i=1}^{n} w_i \left| o_i(\mathbf{x}) - o_i^* \right|^p \right) + \left(1 - \lambda \right) \sum_{i=1}^{n} \left| w_i - \sum_{j=1}^{\textit{number of outputs}} \left(1 - \sigma_{\varpi_j}^{b_i} \right) \right|^p \right],$$

with λ a parameter describing the relative influence of both constraints. w_i and σ weights resulting from the reliability measures of each experienced (the principle behind the calculation of those parameters lies on the stability analysis of each behaviour over the time. The parameter λ indirectly influences the control behaviour of the robot. Large values of λ will lead to a human-centered control strategy, whereas lower values will lead to a robot-centered control strategy. The value of λ would therefore depend on the expertise or the availability of human experts interacting with the robot.

In summary, the degree of relevance or activity is calculated by observing the history of the output of each behaviour. This history-analysis is performed by comparing the current output to a running average of previous outputs, which leads to a standard deviation, which is then normalized). It is obvious that this method increases the numerical complexity of finding a solution to the VOP, but this does not necessarily leads to increased processing time, as the search interval can be further reduced by incorporating constraints from both data sources.

As the Robudem is equipped with multiple sensors with very different spectral properties, various types obstacles can be perceived: the architecture of fig.6 has thus been developed, where, among the behaviours, the traversability assessment and the victim detection have been particularly implemented on the ROBUDEM.

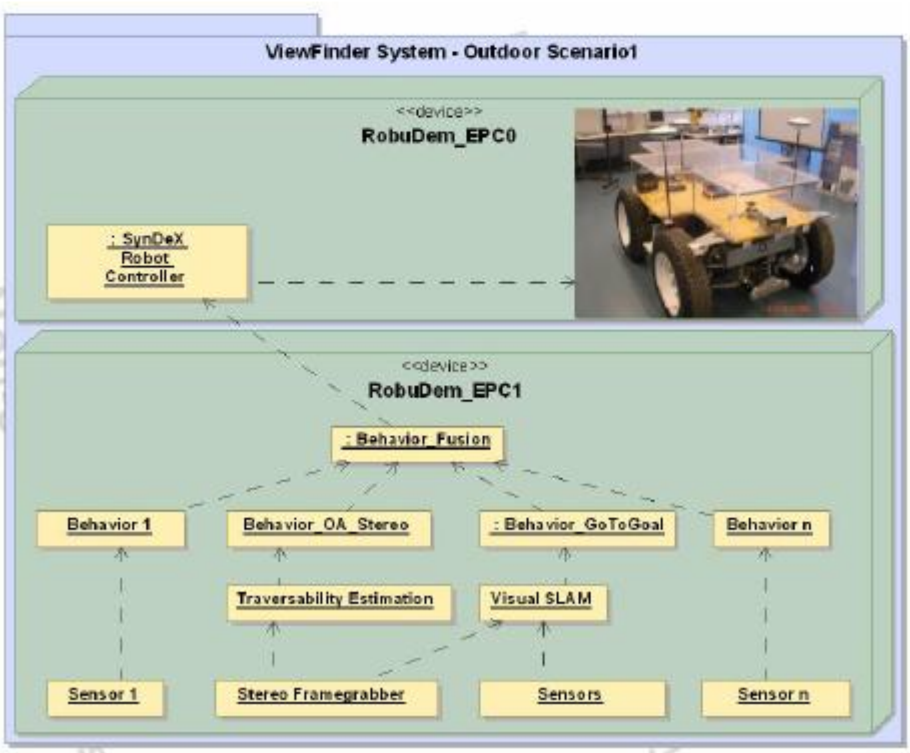

Fig. 6 ROBUDEM Architecture

4.4. Victim Detection

In a first attempt at victim detection, we used the standard Viola-Jones [13] detector for face and upper body detection. The first tests were executed on indoor and good quality images. These tests were very successful, 90% of the faces and 80% of the upper bodies were detected. All together the hit rate reached the 95% while the false alarm rate stayed under 25%. However, the target hardware, the RobuDem, is going to operate in outdoor environment where the background is various and the illumination is unpredictable. So, outdoor experiments were strongly suggested. Although, the results were better than expected, the false alarm was increased dramatically while the hit rate was decreased to 70% for the upper body and to 30% for the face detection. The conclusion from these tests is that in outdoor environment the face detection based person detection is not viable. Usually it only consumes the computation time without giving any results or any correct results. If the detection was more detailed, the system became too slow with minor success. If the detection was tuned to be faster, the hit rate decreased under 10%. The upper body detection is more robust, it adopts itself to different illuminations much better. However, it gives much more false alarms.

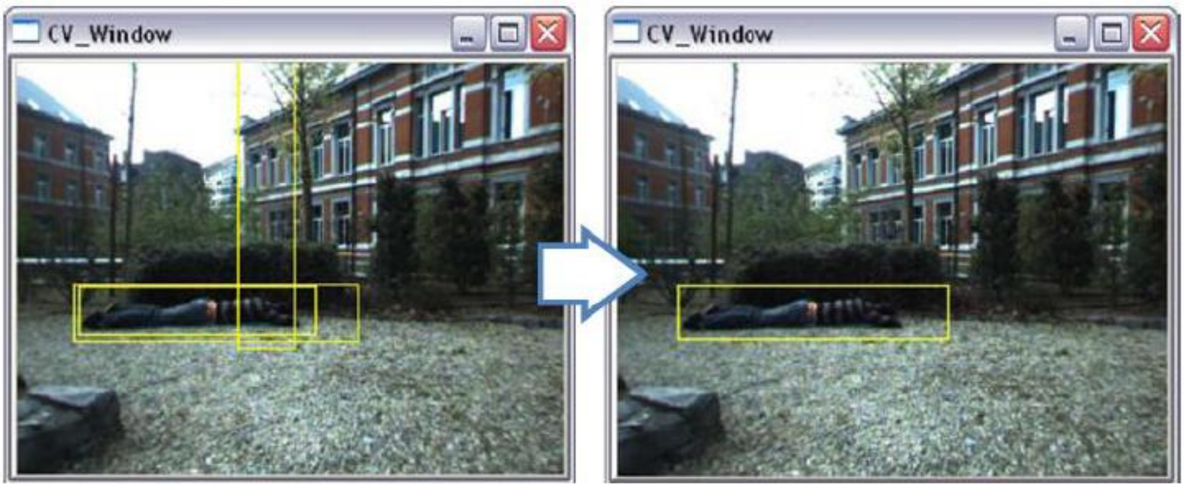

Fig. 7 Before and after merging the neighbour detection areas

4.5. Traversability Analysis

Detecting obstacles from stereo vision images may seem simple, as the stereo vision system can provide rich depth information. However, from the depth image, it is not evident to distinguish the traversable from the non-traversable terrain, especially in outdoor conditions, where the terrain roughness and the robot mobility parameters must be taken into account. Our approach is based on the construction and subsequent processing of the *v-disparity* image [12], which provides a robust representation of the geometric content of road scenes. The v-disparity image is constructed by calculating a horizontal histogram of the disparity stereo image. Consider 2 stereo frames and the computed disparity image, as shown in Figure 8. Then, the v-disparity image can be constructed by accumulating the points with the same disparity that occur on a horizontal line in the image.

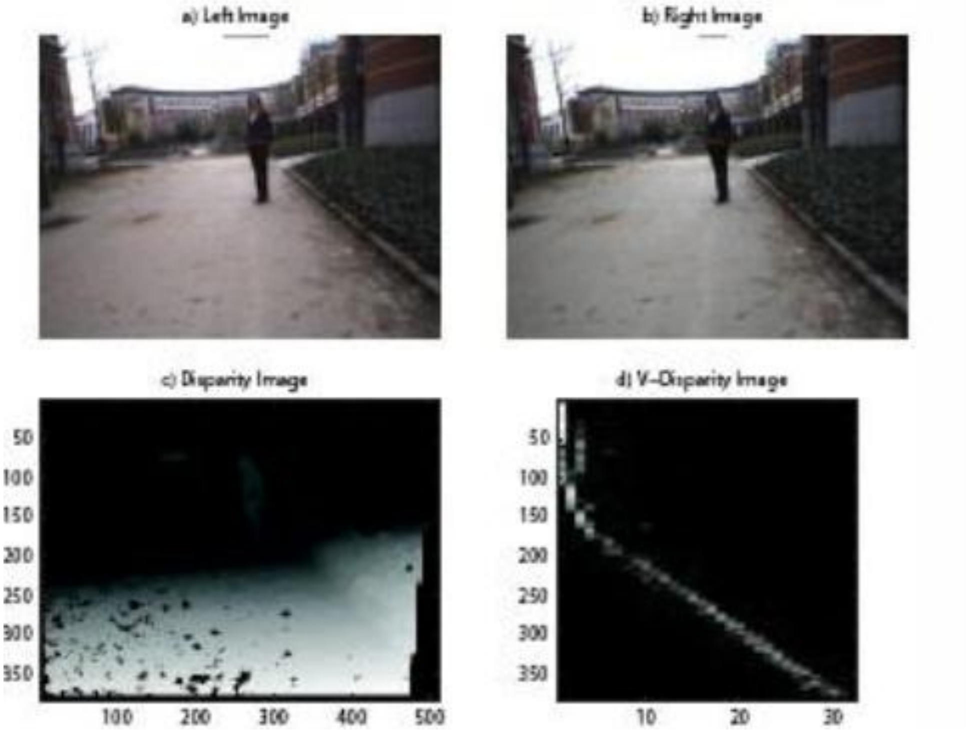

Fig 8. Disparity, V-disparity Image

The classification of the terrain in traversable and non-traversable areas goes out from the assumption that the majority of the image pixels are related to traversable terrain of the ground plane. The projection of this ground plane in the v-disparity image is a straight line, from the top left to the bottom right of the v-disparity image. Any deviations from this projection of the ground plane are likely obstacles or other non-traversable terrain items. As such, the processing of the v-disparity image comes down to estimating the equation of the line segment in the v-disparity image, corresponding to the ground plane. This is done by performing a Hough transform on the v-disparity image and searching for the longest line segment. Then, one must choose a single parameter which accounts for the maximum terrain roughness. As this parameter depends only on the robot characteristics, it only needs to be set once. This parameter sets the maximum offset in v-disparity space to be considered part of the ground plane. Any outliers are regarded as obstacles, which enables to compile an obstacle image.

5. Conclusions.

The very essence of the VIEW-FINDER Intelligent Information System is to integrate disparate elements involved in a crisis situation into an info-structure that will allow information to be exchanged readily between all of those elements: crisis centres, relevant forces dealing with the crisis (fire fighters, de-bombing squads, police, etc.), robotics platforms and sensors. This paper introduced some results of the major outdoor tasks entrusted to the partners of the View-Finder project.

6. References

[1] W. Mees: Information management for localized disaster relief operations. ISMCR 2007 21-23 June 2007 - Warsaw, Poland.

[2] Carlos Pinzón, Yvan Baudoin, Andrzej Maslowski : Disaster Management Action Plan and Integration of Robotics in existing Protection service structures, IARP-RISE'2008, 7-9 Feb 2008 – Benicassim, Spain

[3] Y.Baudoin, J.Penders: Robotics Assistance to Protection services: Users requirements, IARP WS RISE'2008, Benicassim, 7-9 Jan 2008

[4] Y.Baudoin, et Al: Systems requirement Document, Deliverable 1.2, View-Finder Project,Contract no.: 045541 Dec 2006.

[5] E.Colon, H.Sahli, Y.Baudoin: Distributed Control of Robots with CORBA, ISMCR'2005, Brussels, 8-10 November 2005.

[6] S.A.Berrabah, J.Bedkowski: Robot Localization based on Geo-referenced Images and Graphic Methods, IARP WS RISE'2008, Benicassim, Spain

[7] A. J. Davison, I. D. Reid, N. D. Molton, O. Stasse, MonoSLAM: Real-Time Single Camera SLAM, IEEE Trans. PAMI 2007

[8] S.A. Berrabah, G. De Cubber, V. Enescu, H. Sahli, MRF-Based Foreground Detection in Image Sequences from a Moving Camera , IEEE International Conference on Image Processing (ICIP 2006), Atlanta, GA, USA, Oct. 2006, pp.1125-1128

[9] Daniela Doroftei, Eric Colon,Yvan Baudoin, Hichem Sahli Development of a behaviour-based control and software architecture for a visually guided mine detection robot, European Journal of Automated Systems (JESA), 2008

[10] Daniela Doroftei, Eric Colon, Geert De Cubber A behaviour-based control and software architecture for the visually guided Robudem outdoor mobile robot, ISMCR2007, Warsaw, Poland, 21-23 June 2007

[11] G.De Cubber, D.Doroftei IARP WS Environmental Surveillance, Baden-Baden, Germany, 23-25 Jul 2008

[12] R. Labayrade, D. Aubert, J. P. Tarel, "Real Time Obstacle Detection on Non Flat Road Geometry through V-Disparity Representation", IEEE Intelligent Vehicles Symposium, Versailles, June 2002

[13] P. Viola and M. Jones. Robust *Real-time Object Detection*. In Intl. Workshop on Statistical and Computational Theories of Vision, July 2001.

[14] Alessandro Muzzetta, IES Solutions Wireless Communications, CoRoBa, CORBA
and Mailman in ViewFinder, IES technical Report Oct, 2007

[15] J. Będkowski, P. Kowalski, G. Kowalski, A. Masłowski, E. Colon: *Improvement of ATRVJr Software Architecture for VeiwFinder Application* International Workshop On Robotics for risky interventions and Environmental Surveillance RISE'2009 12-14 Jan 2009 Brussels CD-ROM

Acknowledgment: This research is funded by the View-Finder FP6 IST 045541 project.