Atti del Dipartimento di Sistemi Elettrici e Automazione

Distributed Dynamic Programming for Adaptive On-line Planning of AUVs Team Mission with Communication Constraints

A. Caiti, G. Casalino, A. Turetta, R. Viviani

Distributed Dynamic Programming for Adaptive On-line Planning of AUVs Team Missions with Communication Constraints

A. Caiti, G. Casalino, A. Turetta, R. Viviani

Abstract—*An algorithm for adaptive on-line planning of environmental exploration missions with a team of Autonomous Underwater Vehicles (AUVs) is proposed. The algorithm has the primary goal of determining an estimate of the sampled environmental quantity with an estimation error below a prescribed threshold. The additional degree of freedom of the algorithm is exploited to spread the team over the exploration area, in order to minimize mission time while, at the same time, the communication connectivity of the team is preserved. A distributed dynamic programming approach is employed in order to satisfy these two conflicting requirements.*

I. INTRODUCTION

Networked systems have recently experienced phenomenal growth in numerous engineering fields. Thanks to the increased capability in creating small and low-cost mobile platforms and the progresses in communication systems, the deployment of networked mobile agent in a number of applications becomes possible. In particular, multi-robot systems offer potential advantages in performance, robustness, and versatility for sensor-driven tasks such as survey, exploration, and environmental monitoring. There are many examples in the literature of applications of robot cooperation, from robotic soccer teams [1], to the exploration of unknown environments [2], [3]; other important examples come from the military field, where teams of unmanned vehicles (land, aerial, underwater) are asked to perform operations such as border patrol [4] or mine counter measurements [5].

New interesting aspects about the topic of multi-robot cooperation have been introduced in [6] in particular with regard to formation control [7] and leader following [8]. Research on Autonomous Underwater Vehicle (AUV) cooperation in the ocean environment had a somewhat slower start with respect to other fields, mainly as a consequence of the communication and localization constraints posed by the marine environment; in particular, the absence of a GPS-like localization system, and the severe range and bandwidth limitation of underwater communication make it very hard to simply transpose general robotic techniques to the underwater scenario.

An attempt to translate formation control in the marine context has been reported in [9]. In [10] a new behaviour-based approach has been introduced and applied to the control of a fleet of surface vehicles, while in [11], [12] formation control has been extended to the coordinate path-following of marine vehicles, considering communication delays and failures. An interesting approach within this line of research has been proposed and investigated in [13], as referred to ocean gliders with the task of identifying the gradient of environmental oceanographic quantities.

Specific applications require different types of communication technologies: wireless local area networks can be easily established among surface vehicles through radio links, while acoustic modems are generally used for communication between underwater vehicles. Underwater acoustic communication suffers from transmission delay, multi-path fading and limited range and bandwidth.

This paper tackles the problem of a behaviour-based adaptive mission planning for a team of AUVs cooperating in an environmental monitoring mission (e.g., measurement of the temperature field) subject to communication constraints, with regard to range limitations. Mission adaptation aims to maintain a desired accuracy on the reconstruction of the environmental field, through estimation of the local smoothness properties of the field itself as new measurements become available. Following our previous research [16], [17], the accuracy of the environmental map is ensured exploiting the approximation properties of Radial Basis Functions (RBFs). The innovative contribution of this work is in the exploitation of an additional degree of freedom of the adaptive algorithm: this is employed to spread the AUV team over the investigation area (hence attempting to minimize the mission time) while at the same time preserving the connectivity of the team in the form of a chained communication structure where the distance between two adjacent nodes is maintained below a prescribed threshold. This is accomplished with the application of a distributed dynamic programming approach. While in this paper we focus on a serial chained communication structure, the approach can be extended to more complex chained structure or to reconfigurable structure. Moreover, the proposed method is well suited to be performed by a team of Folaga class AUVs [15], equipped with an acoustic modem or a wireless LAN communication device.

The paper is organized as follows: in the next section the problem is formally stated and the cooperative adaptive sampling algorithm is introduced in its general form, considering the preservation of the network links as additional requirements for the vehicles. In section 3 the adaptive constrained exploration algorithm is described. In section 4 simulation results are presented and, in the last section, conclusions are given.
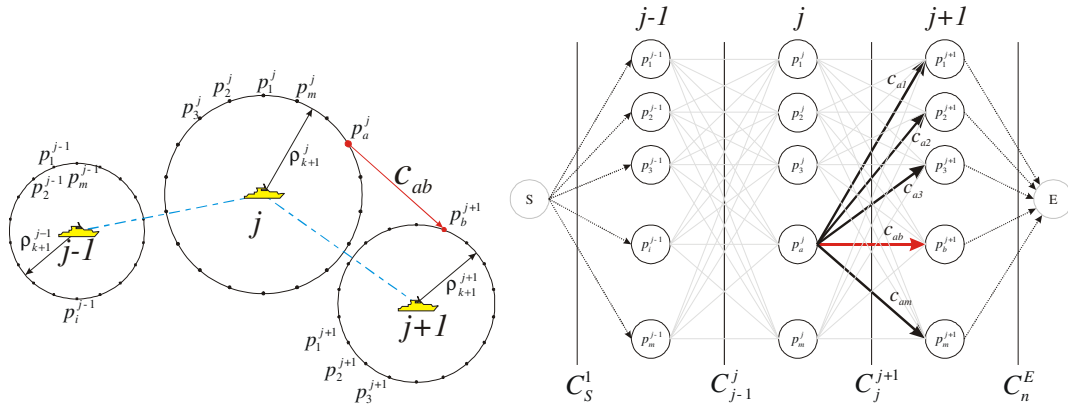

Fig. 1. On the left part, the network topology and the discretized set of candidate measuring points. On the right part, the selection process of the next sampling points seen as *minimum-cost path* problem.

II. PROBLEM STATEMENT

The general framework of the adaptive sampling problem has been presented in our previous works [16], [17]; here the basic concepts are briefly recalled, for completeness, and the general framework is extended to the case in which additional range constraints among the vehicles are enforced. Let us suppose we have the availability of n AUVs, each one equipped with a network device to communicate with the others up to a maximum range D_{max}, and a sensor able to point-wise sample an environmental quantity θ at the geographical coordinates $\mathbf{x} = (x, y)$. The time-scale variation of the oceanographic field is supposed to be larger than

the sampling mission time-scale . Let A be the geographical domain of interest, i.e., $\mathbf{x} \in A$. Let $\mathbf{x}_{j,k_j}$ be the k_j-th measurement point of the vehicle j; let $M^{(j)} = \bigcup_{h=1}^{n} \bigcup_{i=1}^{k_h} \{x_{h,i}\}$ be the set of sampling points known by the j-th vehicle after its last measurement, so $I^{(j)} = \left\{ M^{(j)}; \ \theta = \theta(\mathbf{x}) \middle| \mathbf{x} \in M^{(j)} \right\}$ is the information set available to the vehicle j. Let S be an estimation algorithm that computes an estimate $\hat{\theta}_j$ of the quantity θ over the whole region A on the basis of the current available information $I^{(j)}$, i.e. $\hat{\theta}_j(\mathbf{x}) = S\left(I^{(j)} \right)$. On the basis of the estimation algorithm S and the available information set, one can define the estimation error:

$$\varepsilon_j(\mathbf{x}) = \left| \theta(\mathbf{x}) - \hat{\theta}_j(\mathbf{x}) \right| \tag{1}$$

The main objective of the mission is to survey the region so that the estimation error is everywhere below a given threshold. Moreover, it is desirable to spread the team over the region A, in order to obtain the maximum possible area coverage in the minimum time; however, since the vehicles share (some) information, communication connectivity must be preserved among the team. While in [16], [17] it was assumed that a broadcasting node was always available to all vehicles (as for instance in the case of re-surfacing and radio communication), in this work we assume that the communication takes place underwater through acoustic modems. To maintain communication links, it is assumed that the team has a chained serial structure: the team members are labeled from 1 to n, and each vehicle is the node of a mobile serial network, with static topology. This means that the j-th vehicle is only connected to the vehicle $(j-1)$-th and $(j+1)$-th, see the left part of figure 1. The index j indicates the position of the vehicle into the serial network.

The general cooperative strategy to attain the two mission goals proposed in this paper is based on the approximation properties of RBFs and on the application of dynamic programming in a distributed fashion. Let us suppose the vehicles have an initial configuration such that the communication constraints are satisfied. As soon as the vehicle j completes the k_j-th sampling measurement and communicates it to all the others in the team, the information set $I_k^{(j)}$ becomes available. On the basis of the information set $I_k^{(j)}$ each vehicle computes its exploring radius $\rho_{k+1}^{(j)}$, i.e. a circumference centred at the last sampling point $(x_k, y_k)^{(j)}$ of radius $\rho_{k+1}^{(j)}$. The exploring radius $\rho_{k+1}^{(j)}$ is chosen so that the error of the estimation map is below the required threshold at every point inside the circle, assuming the local smoothness of the environmental field at the point $(x_{k+1}, y_{k+1})^{(j)}$ to be the same of that estimated at the point $(x_k, y_k)^{(j)}$. The selection of the specific point $(x_{k+1}, y_{k+1})^{(j)}$ on the circumference of centre $(x_k, y_k)^{(j)}$ and radius $\rho_{k+1}^{(j)}$ does not influence the expected accuracy of the estimated oceanic field: hence this additional degree of freedom can be used to satisfy the other mission requirements. Let us suppose that each vehicle shares the information about its exploring radius $\rho_{k+1}^{(j)}$ with the others mobile agents, and that it can choose from a finite set, $P_m^{(j)} = \left\{ p_1^{(j)}, p_2^{(j)}, \ldots, p_m^{(j)} \right\}$, of m points belonging to its exploring circumference. It is possible

to define a cost function related to the choice of the point $p_a^{(j)}$ by the vehicle number j with respect to other vehicles' choices. The vehicles in the team select their next sampling point as the solution of a distributed dynamic programming problem that minimizes the cost function.

This cooperative approach is well suited to the operation of an AUV, since the sampling points are chosen in a convenient neighborhood of the current vehicle position. The choice of both the estimation algorithm and the exploring radius $\rho_{k+1}^{(j)}$ have been discussed in [16], [17]; now the selection of the next sampling point $(x_{k+1}, y_{k+1})^{(j)}$ that allows to satisfy the additional constraint is investigated, and in the following section a solution method based on dynamic programming is presented.

III. ADAPTIVE CONSTRAINED COOPERATIVE EXPLORATION ALGORITHM

Let us consider the the vehicles as the nodes of a mobile serial network, with static topology, as defined in the previous section. The selection of the sampling points is synchronized in accordance with the following adaptive planning algorithm:

A. Error estimation and Computation of the exploring radius

On the basis of the available information set I, the oceanographic quantity θ is estimated exploiting the approximation properties of RBFs. In the RBFs framework the estimation error can be bounded by the following equation:

$$\varepsilon(\mathbf{x}) = \left| \theta(\mathbf{x}) - S(\mathbf{x}) \right| \leq \left\| \theta(\mathbf{x}) \right\|_{\Phi} F_{\Phi}\left(h_{\rho}(\mathbf{x}) \right) \qquad (2)$$

where $\theta(\mathbf{x})$ is the environmental quantity measured at $\mathbf{x}$, $S(\mathbf{x})$ the estimation of $\theta(\mathbf{x})$ on the basis of the available information set I, $h_{\rho}(\mathbf{x})$ is the *local fill distance* and it depends on the local density of sampling points:

$$h_{\rho}(y) = \sup_{w \in B(y,\rho)} \min_{\mathbf{x} \in M^{(j)}} \left\| \mathbf{w} - \mathbf{x} \right\|_2 \qquad (3)$$

while $F_{\Phi}(\cdot)$ (the *power function*) is a known function that depends only on the specific RBF choice (gaussian, multiquadric, etc.); some common expressions of the power function are reported in [18].

Under some additional technical assumptions (decay to zero of the RBF Fourier transform and $S(\mathbf{x})$ smoother than $\theta(\mathbf{x})$), the j-th vehicle determines its next exploring radius ρ_{k+1}^{j} as the *local fill distance* that allows to satisfy the mission error requirements through Equation (2).

B. Distributed dynamic programming: backward phase

The next sampling point for each vehicle is chosen as the solution of a *minimum-cost path* problem by adding to the network the virtual nodes S and E, whose arches have no cost, see figure 1. Starting from the last node of the network, i.e. the vehicle number n, the following procedure is repeated until the first node (the vehicle number 1) is reached:

- The $(j+1)$-th vehicle determines its exploring radius $\rho_{k+1}^{(j+1)}$, as described in the previous subsection, and sends this information to the vehicle number j.

- The j-th vehicle computes the Cost Matrix C_j^{j+1}: each element c_{ab} of the Cost Matrix is related to the choice of the point $p_b^{(j+1)}$ taken by the vehicle $(j+1)$ with respect to the choice of the point $p_a^{(j)}$ taken by the vehicle number j. The value of the cost is given by the following rule:

$$
c_{ab} = \begin{cases} +\infty & p_b^{j+1} \notin A \\ +\infty & \left\| p_a^j - p_b^{j+1} \right\|_2 > D_{\max} \ , \\ \varphi\left(p_a^j, p_b^{j+1} \right) & otherwise \end{cases}
\tag{4}
$$

where A is the geographical domain of interest, $D_{\max}$ the maximum distance between two topologically adjacent vehicles that allows to maintain the network connection, $\varphi\left(p_a^j, p_b^{j+1} \right)$ is the value of a suitable scalar potential field generated at the point p_a^j by the past measurements points ($x_i \in M$) and by p_b^{j+1} as a new, hypothetical, measurement point for the vehicle $(j+1)$-th.

- The potential field is defined as:

$$
\varphi\left(p_a^j, p_b^{j+1} \right) = q\left(\frac{1}{\left\| p_a^j - p_b^{j+1} \right\|_2^n} + \sum_{x_i \in M} \frac{1}{\left\| p_a^j - x_i \right\|_2^n} \right) - \frac{Q}{\left\| p_a^j - x_{fd} \right\|_2^n} \ ,
\tag{5}
$$

where q, Q, n, are positive numbers and $x_{fd} \in A$ is the point, in the region A, where the fill distance is reached. The scalar potential field defined in (5) allows to spread the vehicles over the area taking in to account the past sampling points (considering them as electric charges, $n=2$) and the zones where the density of measurements is lower (the point $x_{fd} \in A$ attracts the mobile agents). The proposed cost function (eq. 4) allows to spread the vehicle over the region and, at the same time, to maintain the communication among the mobile agents.

- Let p_s be a generic element of $P_m^{(j)}$, i.e., the set of the candidate sampling points for the vehicle j, and let us recursively define

$$
f_j\left(p_s \right) = \min_{p_z \in P_m^{(j+1)}} \left[c_{sz} + f_{j+1}\left(p_z \right) \right]
\tag{6}
$$

as the minimum const related to the choice of $p_s^{(j)}$ as a new measurement point for the j-th vehicle, where c_{sz} is an element of the cost matrix C_j^{j+1}.

The cost function is evaluated for each feasible sampling point of every node, beginning from the last node E, $f(E) = 0$, and going back until the node S. The virtual nodes S and E do not contribute to the cost.

C. Distributed dynamic programming: forward phase.

The *minimum-cost path* from S to E indicates the optimal sequence of the next sampling points for the AUV team, with respect to the cost function defined by equations 4 and 5. Starting from the virtual node S, the sampling point for the first vehicle is chosen as the first step of the *minimum-cost path*; then the (j-1)-th vehicle sends the coordinate of the measurement point to the j-th vehicle; the procedure is then iterated along the serial structure until the last mobile agents of the network receives the coordinate of its sampling point from the (n-1)-th vehicle.

D. Moving toward the sampling points

After the forward phase has been completed, all the vehicles move toward their next sampling point, perform and share their measurements, update the information set I and the estimate of the oceanographic quantity θ. The exploration ends when the equation 2 is satisfied over the whole experimental area. In the next section a particular application of the adaptive sampling will be described through simulative results.

IV. SIMULATION RESULTS

The algorithm described in the previous section has been tested in a simulative scenario where the oceanographic data are field data from past oceanographic cruises. The map to be estimated is the ocean temperature in a shallow water region of 70m depth and 5x5 km width, characterized by a warmer water mass at the centre of the area. Three cooperating vehicles are supposed to perform the mission and it is assumed that the vehicles are equipped with a TD (Temperature Depth) probe, sampling the water column at one sample per meter at each required (x, y) position and a network device that allows communication between two vehicles, if their distance is lower than 1 Km. Empirical Orthogonal Functions (EOFs) are used to code the vertical temperature profile, as commonly done in oceanography.

In figure 2a the behaviour of the team during the exploration is shown: even though eq.5 aims to spread the vehicles over the region, eqs. 4 and 6 constraint them to move in team, in order to keep the network links. Figure 2b shows the paths followed by the three AUVs and the location of the sampling points: a higher number of measurements is taken where the variability of the temperature field is greater, i.e., where the temperature gradient is higher. In figure 3a and 3b are depicted the estimated temperature field at 25m depth and the related approximation error, respectively.

V. CONCLUSION

An algorithm for adaptive on-line planning of oceanographic missions to be performed in cooperation by a team of AUVs has been presented. Two mission goals have been considered: the main one is expressed in terms of accuracy in the reconstruction of the environmental field to be sampled, the second takes in to account the need of maintaining the network links

between the vehicles. Adaptive cooperative behaviour is achieved by the team in terms of local evaluation of the sampled field smoothness, and selection of the next sampling point in order to reach the desired accuracy and satisfy the network constraints; smoothness evaluation and accuracy estimation have been proposed in terms of analytical formulation related to field estimation by RBFs, while a distributed dynamic programming formulation has been proposed to satisfy the second requirement.

Current work is progressing toward the extension of this approach to the case of a generic articulated communication chain structure, to the case in which ordering of the vehicles may be interchanged, and to the investigation of the degradation of performance when connectivity is lost: note that even in presence of loss of connectivity, for instance in one intermediate node, the two resulting chains can still proceed to apply the on-line planning algorithm, although with a limited information set updates.

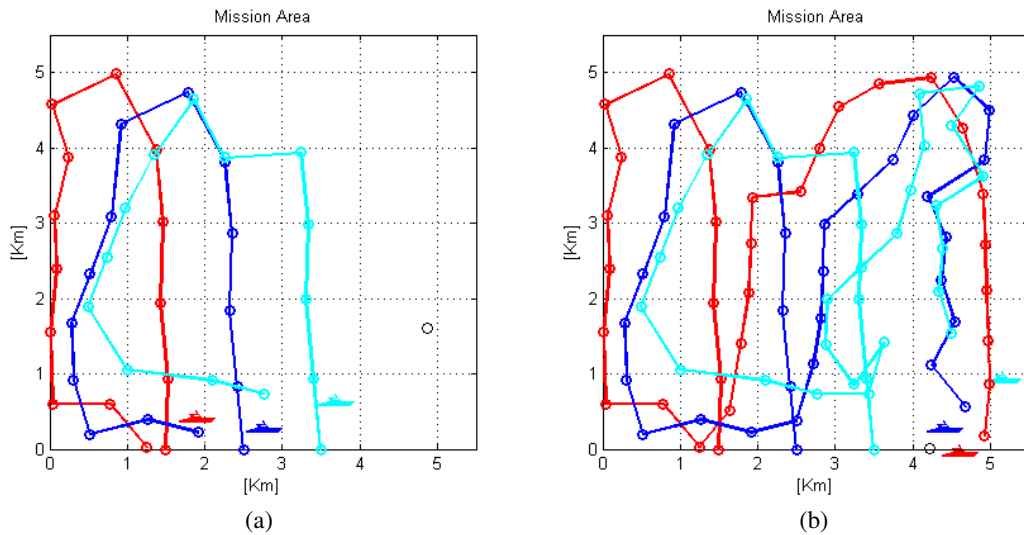(a)   (b)

Fig. 2. (a) Paths in the mission area during the exploration. The vehicles adapt their exploring radius to guarantee the accuracy on the reconstruction of the environmental field, while they move in formation in order to keep the network links. (b) Paths in the mission area and measuring points selected by the vehicles during the exploration.
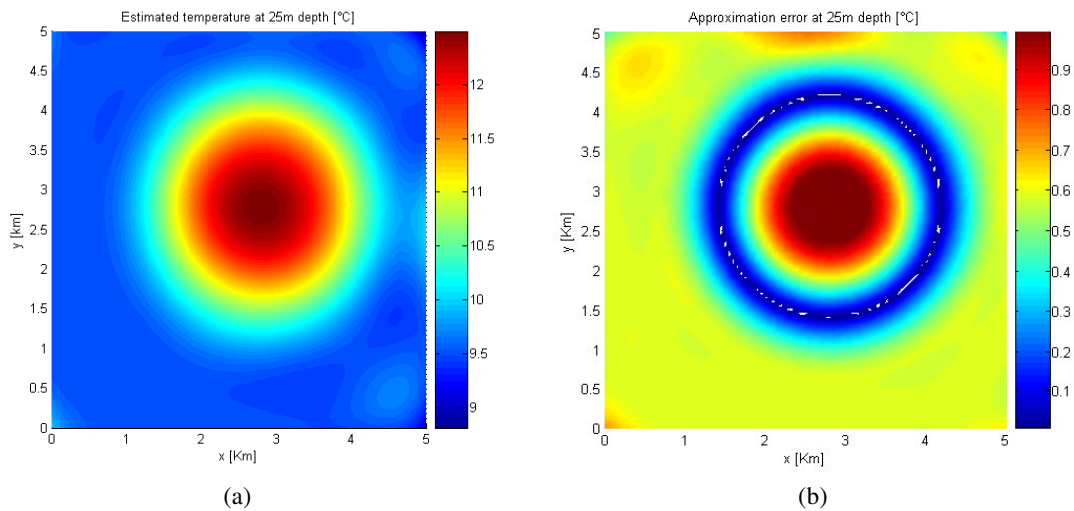

(a)   (b)

Fig.3. (a) Estimated temperature field at 25m depth. (b) Approximation error at 25m depth.

VI. REFERENCES

[1] M. Veloso, "Autonomous robot soccer teams," in *The bridge*, vol. 33, part.1, National Academy of Engineering, Ed. Spring, 2003, pp. 8–12.

[2] W. Burgard, M. Moors, C. Stachniss, and F.E. Schneider, "Coordinated multi-robot exploration," *IEEE Journal of Robotics*, vol. 21, no. 3, pp. 376–386, June 2005.

[3] A. Danesi, D. Fontanelli, P. Murrieri, V. G. Scordio, and A. Bicchi, "Cooperative Simultaneous Localization and Map Building for Servoing," in *Proc. Multiagent Robotic Systems Trends and Industrial Application*, Padova, July 2003.

[4] A. Girard, A. Howell, and K. Hedrick, "Border patrol and surveillance missions using multiple unmanned air vehicles," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, vol. 1, pp. 620–625, December 2004.

[5] D. Cecchi, A. Caiti, S. Fioravanti, F. Baralli, and E.Bovio, "Target detection using multiple autonomous underwater vehicles," in *Proc. IARP Int. Workshop on Underwater Robots IWUR 05*, Genova, November 2005.

[6] V. Kumar, N. E. Leonard, and A.S. Morse, *Cooperative Control*, LNCIS 309. Springer-Verlag, Berlin 2005.

[7] G.T. Tanner, J. Pappas, and V. Kumar, "Leader-to-formation stability," *IEEE Trans.Robotics and Automation*, vol. 20, no.3, pp. 443-455, June 2004.

[8] T. Balch, T., R.C. Arkin, "Behavior-based formation control for multi-robot teams," *IEEE Trans. Robotics and Automation*, vol. 14, no. 6, pp. 926-939, December 1998.

[9] T. Ikeda, T. Mita., and Y. Masaki, "Formation control of underwater autonomous vehicles," in *Proc. 16th IFAC World Congress on Automatic Control*, Praga, July 2005.

[10] F. Arrichiello, S. Chiaverini, and T.I. Fossen, "Formation control of Marine Surface Vessels using the Null-Space-based Behavioral Control," in K.Y. Pettersen, J.T. Gravdahl, and H. Nijmeier, *Group Coordination and Cooperative Control*, Springer-Verlag's Lecture Notes in Control and Information Sciences series, Springer-Verlag, Heidelberg, pp. 1-19, 2006.

[11] E. Børhaug, A. Pavlov, R. Ghabcheloo, K. Y. Pettersen, A. Pascoal, and C. Silvestre, "Formation Control of Underactuated Marine Vehicles with Communication Constraints," in *Proc. 7th IFAC Conference on Manoeuvring and Control of Marine Craft*, Lisboa, 2006.

[12] R. Ghabcheloo, A. P. Aguiar, A. Pascoal, and C. Silvestre, "Coordinated Path-Following Control of Multiple Auvs in The Presence of Communication Failures And Time Delays," in *Proc. 7th IFAC Conference on Manoeuvring and Control of Marine Craft (MCMC 2006)*, Lisboa, 2006.

[13] P. Ögren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: adaptive gradient climbing in a distributed environment," *IEEE Trans. Automatic Control*, vol. 49, no. 8, pp. 1292-1302, August 2004.

[14] J. G. Bellingam, Y. Zhang, "Observing processes that vary in time and space with heterogeneous mobile networks," in *Proc. IARP Int. Workshop on Underwater Robots IWUR 05*,Genova, November 2005.

[15] A. Alvarez, A. Caffaz, A. Caiti, G. Casalino, E. Clerici, F. Giorgi, L. Gualdesi, A. Turetta, and R. Viviani, "Folaga: a very low cost autonomous underwater vehicle for coastal oceanography," in *Proc. 16th IFAC World Congress on Automatic Control*, Praga, July 2005.

[16] A. Caiti, A. Munafò, and R. Viviani, "Adaptive On-Line Planning Of Environmental Sampling Missions With A Team Of Cooperating Autonomous Underwater Vehicles," *International Journal of Control*, to be published.

[17] A. Caiti, A. Munafò, and R. Viviani, "RBF-Based Adaptive On-Line Planning of Auv Teams Environmental Missions," in *Proc. 7th IFAC Conference on Manoeuvring and Control of Marine Craft (MCMC 2006)*, Lisboa, 2006.

[18] R. Schaback (1997), *Reconstruction of multivariate functions from scattered data*, Available: http://www.num.math.uni-goettingen.de/schaback/research/group.html