Cognitive micro-Agents: individual and collective perception in microrobotic swarm

S. Kornienko, O. Kornienko, C. Constantinescu, M. Pradier and P. Levi,
Institute of Parallel and Distributed Systems, University of Stuttgart,
Universitätsstr. 38, D-70569 Stuttgart, Germany

Abstract

In this paper we present the research results in the field of perception for real microrobotic swarm. The proposed hardware and software solution uses IR-based reflective measurement for individual perception and the Dampster-Shafer evidential reasoning for hypothesis refinement in collective perception. Especial attention is paid to a reliable identification of encountered geometries and a reduction of local communication. Based on the experimental results we make a conclusion about cognitive capabilities of individual microrobots and the whole swarm.

1 Introduction

Miniaturization represents now a very important trend in many areas of research. Molecular-scale or nanotechnological devices jumped from science-fiction novels to research papers. Even the today's technology allows creating complete autonomous systems, such as robots, in the size of 1 mm^3. As demonstrated by a progress in the I-Swarm project [I-Swarm, 2003 2007], the swarm of thousand such microrobots gets reality as well as come into the reality impressive applications of this technology.

The scaling down of the hardware influences almost all important parameters of microrobots, as e.g. running time, communication distance and channel capacity, computational power, movement and so on. However we ask ourselves about *"intelligence" of such a microrobot; is it also scaled down so that we get finally some "stupid moving thing"* [Kornienko *et al.*, 2004] ? Since many years there exists in the scientific literature the opinion that "artificial intelligence" for very small systems drifts towards "collective artificial intelligence", like those in social insects [Bonabeau *et al.*, 1999]. For collective systems the "individual intelligence" gets some pre-intelligence form. The question is *which minimal degree of individual intelligence does allow growing "collective intelligence"* ?

In this paper we consider such an aspect of cognitive intelligence as perception. In a microrobotic swarm the size of a robot is essentially smaller than the size of most environmental objects. The recognition of these objects is primarily done in collective way. However here we encounter the same question about "individual aspects" of collective perception. Is a microrobot able to provide enough sensory information for the collective perception ? Which sensing and processing steps should be done individually and which collectively ?

For answering these questions we designed and prototyped a sensor system for our own test microrobot. This is is actually larger as envisioned in I-Swarm project however is very cheap and easy to reproduce without specific equipment. Based on this prototype we can investigate questions about "individual/collective intelligence" so that the results, e.g. principles, methods, algorithms can be later implemented in the $1mm^3$ robot. The size of the sensor system is $23\times23\times5$mm. It uses the Megabitty board ($23\times23\times2$mm) with Atmel AVR Mega 8 microcontroller, having 8 kB ROM and 1 kB RAM [Megabitty, 2005]. Besides perception, the board supports 6-directional robot-robot and host-robot communication, with the average communication radius 0-140mm (with special solution for deadlock reduction) and a maximum of 300mm. The sensors are also used for proximity sensing in navigation. The communication subsystem for a large microrobotic swarm is described in [Kornienko *et al.*, 2005]. In this paper we present the development of the perception system for the sensor board and the problems of individual and collective perception in microrobotics.

The rest of paper is organized as follows. In the next two sections the problem of individual perception and the development of IR-perception system are described. Then, we discuss the nonlinearities of this perception and the algorithms of feature extraction and surfaces classification. The last two sections are devoted to the problem of collective classification and preliminary experiments.

2 Problems of individual perception in microrobotic swarms

As mentioned before, the recognition of large objects by small microrobots is primarily performed in a collective way. However the prerequisite for collective perception is the surface identification and classification that is performed by each microrobot. We name further this process as individual perception. From the collective perception point of view the following types of surfaces are required to be identified:

1) *infinite-size surfaces* (from a robot's viewpoint), as huge objects or borders;

2) *finite-size surfaces* (a microrobot has to calculate the visible size of a surface) which are classified, at least, into small, medium and large;
3) *convex and concave corners*;
4) *2-side and 3-side concave surfaces*;
5) *one-surface/many-surfaces geometry*.
Additionally, the microrobots have to be able to perform the following activities:

1) *detection of holes (gangways) in surfaces*;
2) *classification of the perceived surfaces into defined classes and providing a probability of correct classification*;
3) *recognition of robot's own position in relation to a corner* (left/right from a corner) or even its own slope to a surface.

When each robot identifies the surface in its own sensing areal, further collective processing consists in fusing individual observations into many hypotheses and collective identification of most probable hypothesis about the observed object (see also [Ye *et al.*, 2002]).

Returning to the issue of individual perception, we identified the following implementation possibilities:

1) **vision-based** way by e.g. using some small micro(faced)-cameras;
2) **reflection-based** way by using laser or infra-red light, ultra-sound etc.;
3) **wavelength-based** way such as color sensing;
4) by using **specific** chemical, temperature, vibration, magnetic and so on sensors (we do not consider them here).

The vision-based way represents the most information intensive mode. However its application in microrobotics has several difficulties caused by very limited computational capabilities and small memory. Algorithms of image processing are difficult to be implemented in this hardware. Moreover due to very small size we prefer to use the same sensors for navigation (proximity sensing and obstacle detection) and communication (robot-robot and host-robot) purposes as well. Finally, the geometrical features from deep images are essentially more useful for collective perception than edges and regions from camera's grey-value images. Thus, the vision-based as well as wavelength-based ways, although they have found a large application in mini- and usual robotics, unfortunately are less useful here. The reflection-based perception uses the principle of sending and receiving a signal, that can be also used for navigation and communication.

Considering different alternatives for reflection-based perception we focus primarily on laser, electro-magnetic/inductive and infra-red systems. Ultra-sound systems do not satisfy the size limitation. Though the laser provides the most exact measurement and long range, there are several technical difficulties to use it with the microrobot. So, choosing between electro-magnetic/inductive and infra-red systems, we prefer the last ones due to their simplicity, relative long working range and small energy consumption.

Generally, the IR-systems are recently dominant in so-called small-distance-domain, as e.g. for communication between laptops, hand-held devices, remote control and others. The IR-solution is not new in robotic domain, see e.g. [Kube, 1996], [Suzuki *et al.*, 1995]. There are many approved schemes or even industrial sensors for IR-communication. However the fusion of perception and communication using IR-devices does not find too many applications, perhaps because of a high nonlinearity of IR-based perception and availability of more appropriate solutions in the domain of usual robotics. Therefore the microrobotic domain of integrated IR-solution (perception, communication, navigation) is more or less unexploited.

The IR-based perception consists on sending an IR radiation beam and receiving the reflected light. The intensity of this light contains information about the geometry of reflecting surface (primarily a distance between IR-recever/emitter and surface). As mentioned, the IR-based perception is highly nonlinear. The most large influence on accuracy of perception exerts the resolution of the distance sensor. In the center of radiation ray, the intensity of IR radiation is highest. Closely to the bounds of this ray, this intensity becomes gradually degraded (Figure 1). The main component of a reflecting light consists of the energy of the central radiation stream. However low-intensity "secondary streams" spread the reflecting light so that object's edges and gaps between objects get non-recognizable. With a poor resolution of distance sensor, small geometrical elements cannot be perceived and so cannot be used as features for recognition. Therefore

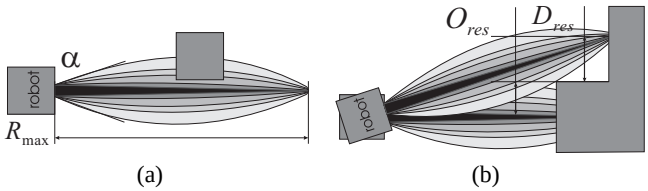


Figure 1: *Perception by using the IR beam, R_{max} recognition distance, D_{res}, O_{res} distance/object resolution.* **(a)** *Thickness of radiation beam and influence on the size measurement;* **(b)** *Nonlinearity in the identification of many-surfaces geometry.*

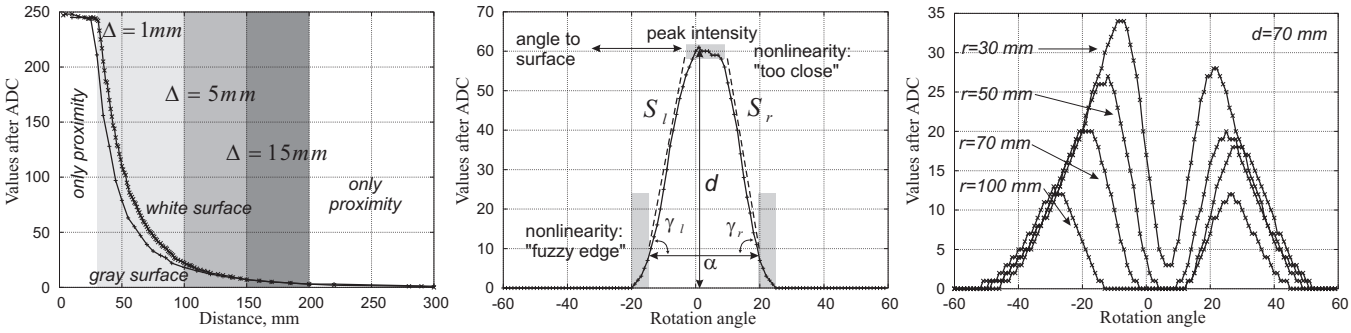
for perception are suitable only such IR-emitters that have an as small as possible opening angle of the beam.

Secondly, the accuracy of measurement depends on the distance to a reflecting surface[1]. In Figure 2(a) we demonstrate this effect for the developed sensor system. Nonlinear accuracy essentially influences the further recognition of features.

The reflecting light is also very sensitive to the color of reflection object. In Figure 2(a) we show the distance measuring values for white and gray objects. Further in experiments we use only white color objects. The distance measuring also depends on the object's slope to a radiation ray. In Section 4 we discuss in detail these nonlinearities and suggest some approaches to absorb them.

Since we did not found a suitable integrated IR-solution for the microrobot, we decided to develop our own required hardware and the corresponding processing algorithms. In the next sections we describe them.

[1]The dependence between reflecting light and distance is also nonlinear however this problem can be easily solved by a look-up table or some approximation functions.



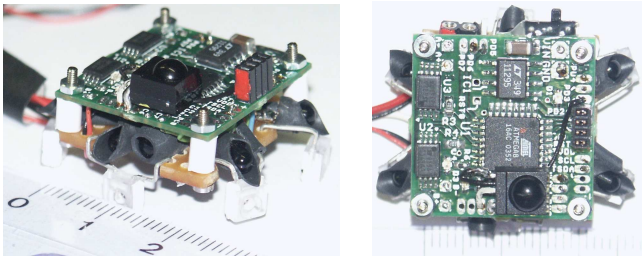
(a) (b) (c)

Figure 2: **(a)** *Dependency between ADC values of emitter voltage on phototransistor and the distance to reflecting object. Shown are values for the white reflecting object (white paper) and the grey reflecting object (grey cardboard);* **(b)** *The used features of IR-diagrams relevant for identifying the surfaces;* **(c)** *The "thickness effect" of radiation beam by scanning a gap with different size r. The distance between a microrobot and the gap is 70 mm.*

3 Development of the IR-based perception system

The main requirement on the IR-perception is given by as small as possible opening angle of the radiation ray. Additionally, IR-emitter has to provide a high energy beam, being able to get good deep images. Finally, IR-emitter and receiver should be able to work in a communication mode.

The perception system of the microrobot is a part of IR-system used for proximity sensing, obstacle detection, distance measurement and communication, as well (Figure 3). For the perception and objects recognition we use only the



(a) (b)

Figure 3: **(a)** *The megabitty board and the sensors board used in the prototype of a microrobot;* **(b)** *The 6-directional sensor system for directional communication and proximity sensing.*

distance measuring sensor, so that only this sensor is further considered. This sensor consists of a receiver with a wide opening angle (used also for communication and proximity sensing) and an emitter with as small as possible beam angle (used for perception and long-range communication). We utilize the Si phototransistor TEFT4300 (60^o, peak sensitivity 950 nm) and the high power GaAs/GaAlAs emitter TSAL6100 (radiant intensity >80 mW/sr, 20^o, the real opening angle is of 18-22^o, 950 nm). This combination is a result of many experiments with different sensors (over 30 pairs), with integrated receiver/emitter like SFH9201, as well as non-integrated ones. The TEFT4300-TSAL6100 pair demonstrated the best spectral coupling, the longest sensing distance and the acceptable nonlinearity of sensing. Although the IR-emitter is relatively large for the microrobot (8xϕ5 mm), the specific construction of the chassis allows to hide it inside the robot.

Since IR-emitter and receiver are non-integrated and are placed side by side in the chassis, they have to be optically isolated. The optical isolation of the emitter allows also reducing the opening angle of the beam up to 10-15^o (it reduces also a perception distance). However the main problem here is to provide similar optical characteristics of isolation for a large number of different microrobots in a swarm (to avoid later the problem of individual calibration of each microrobot).

The principle of object recognition is the following. As soon as a robot detects (by means of proximity sensors) an obstacle in front of itself, it switches on the high power IR-emitter and after 1ms delay (needed to get reliable reflecting light) measures voltage on the emitter of phototransistor. The dependence between emitter voltage (after ADC) and the distance to an object is shown in Figure 2(a). Generally, this sensor perceives distances up to 300 mm. However accuracy of measurement is different. For the pair *distance-accuracy* where Δ is the accuracy, we obtained the following values: 30-100 mm $\rightarrow$ Δ=1 mm, 100-150 mm $\rightarrow$ Δ=3-5 mm, 150-200 mm $\rightarrow$ Δ=10-15 mm and after 200 mm $\rightarrow$ Δ=30-50 mm. Therefore, the reasonable measuring distance for object recognition lies within 30 mm-100 mm (with the accuracy of 1-2 mm).

Not only the resolution of the IR-sensor is important for scanning the objects. During scanning, a microrobot turns on some degrees. The more exact is this turning, the more precise is the spatial resolution of sensor data. Microrobot does not possess any devices allowing to measure positions and orientation of chassis or wheels. Therefore there is only one way to rotate a robot, namely to turn the motors on and after some delay turn them off. This delay has to be so chosen, that a robot rotates on some fixed degree. The motors are controlled through the H-bridge SI9988, that can change a polarity of supplying current. Choosing normal polarity for one motor and inverse polarity for the second motor, the ro-

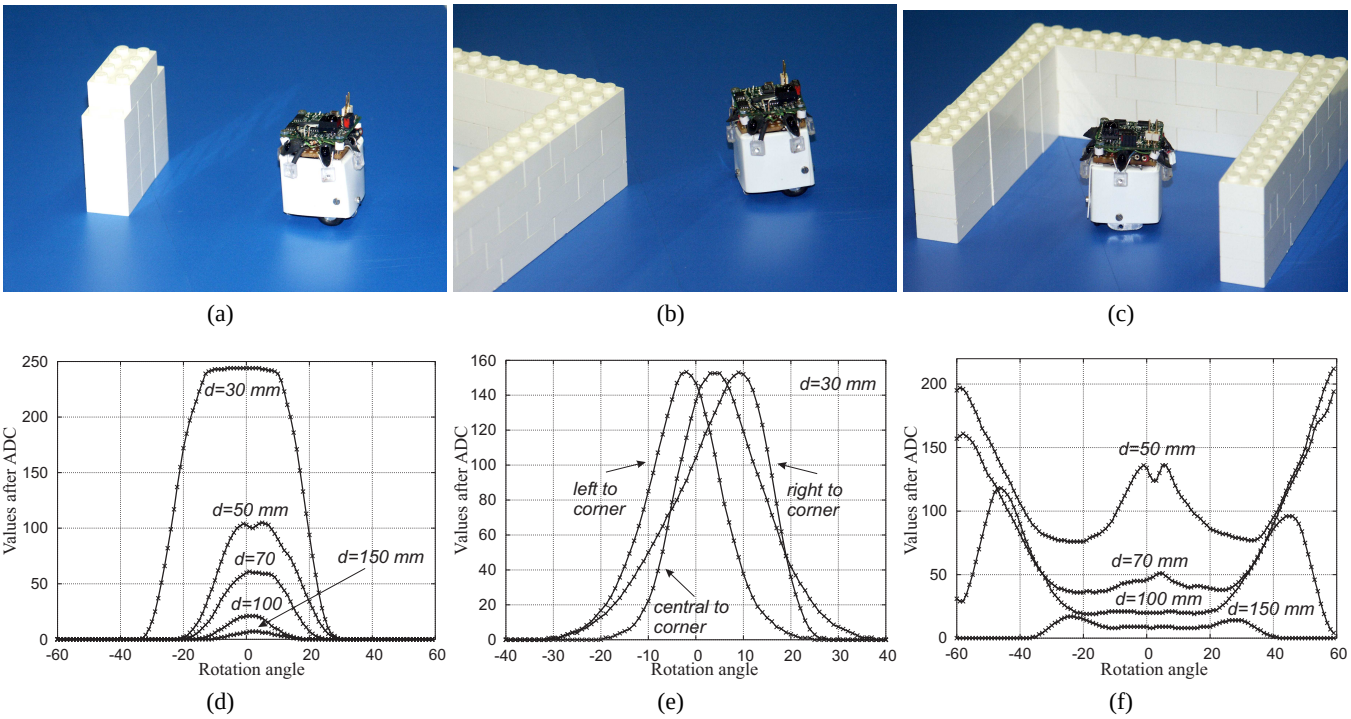


Figure 4: *"Jasmine", the prototype of the microrobot, scans different surfaces, where d is the distance to surfaces.* **(a)** *Scanning of the finite-size surface, object 48 mm;* **(b)** *Scanning of the convex surface;* **(c)** *Scanning of the 3-concave surface;* **(d)** *The IR-diagram for finite-size surface;* **(e)** *The IR-diagram for convex surface;* **(f)** *The IR-diagram for 3-concave sides surface of 95×95×95 mm;*

bot can rotate without changing its own position. In this way we get relatively shift-errorless deep images. After some tests we achieved the resolution and accuracy of rotation 1^o (taking into account different friction between weels/chassis and floor surface).

In our experiments, when a robot detects an obstacle on the distance of 70 mm $\pm$ 10 mm, it stops and then rotates 60^o left. After that it scans the obstacle with the distance sensor by rotating 120 ^{o} right. During this scanning it writes the obtained values of distances each 1 degree into an integer array. In this way we have 120 values describing a visible geometry of the encountered obstacle. In Figure 4 and 5 we demonstrate some geometries of encountered obstacles and the scanned surfaces.

4 Features extraction from IR-deep images

After performing the first experiments, we faced the following challenge: which features of the obtained IR-diagrams are relevant for identifying the geometry of the surfaces ? By analyzing the IR-diagrams in Figure 4 and 5, we find the following features as representative and useful in the IR-based individual perception (Figure 2(b)):

1. *The angle* α, which represents the scanning angle between the first visible edge and the last visible edge of the surface;

2. *The peak intensity of the diagram,* I_{max}. This corresponds to the maximal intensity of reflecting light and, in turn, to the minimal distance d between the surface and the microrobot. For the most types of surfaces (beside convex corners) this minimal distance is measured as a perpendicular to a surface. This feature allows calculating the visible size of a surface by using trigonometric relations;

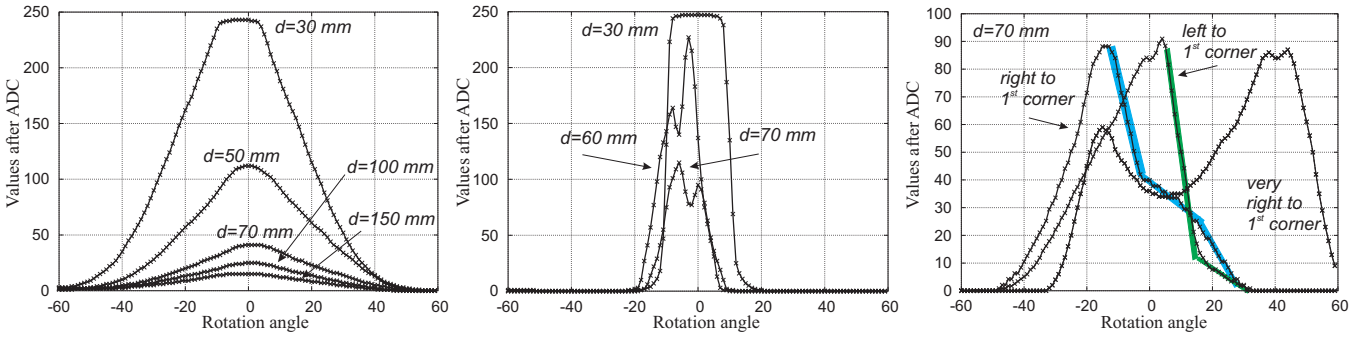
3. *The left and right slopes*, denoted as γ_l and γ_r are useful for identifying the size-type of the surface (unlimited, big, medium, small). They are calculated as slopes of the approximation lines S_l, S_r. The slope denotes also the "degree of a distance decreasing" and enable us to identify the so-called "convex surfaces" that cannot be recognized in the trigonometrical way;

4. *The position of the "center" of the IR-diagram,* P_{imax} in relation to the scanning angle ("0", origin point on the X axis). Displacement of the center points to a slope between the front of robot and surface. In this way we can identify a directional orientation of the microrobot.

Now we formalize the nonlinearities mentioned in Section 2 and present their impact on the corresponding features:

1. *Nonlinear thickness* of the IR radiation ray and so different distribution between high-energy beam and low-energy beam. The first effect of this nonlinearity consists in spread edges (Figure 2(b)). This nonlinear effect can be absorbed by calibration. The second effect is shown in Figure 2(c). At scanning many-surfaces geometry (a gap between objects) a robot cannot reliable differentiate between 2-concave surfaces and surfaces that belong to different objects;

2. *Nonlinear measurement for small distances*. As known from other IR-distance measurement systems (e.g. [Caprari and Siegwart, 2003]), the maximal intensity of measurement



(a) (b) (c)

Figure 5: *IR-diagrams for different types of surfaces, d is a distance to surfaces.* **(a)** *"Infinite-size" surfaces with flat geometry;* **(b)** *Convex round (external diameter 125 mm) surface;* **(c)** *Many-surfaces geometry (1^{st} convex corner 122×60 mm and 2^{nd} concave corner 60×95 mm), robot positioned 70 mm before the middle part.*

lies in 10-25% before the front of IR-receiver, after that the intensity goes down (therefore small distances cannot be measured by these systems at all). Due to the specific restriction and the application of high-power GaAs/GaAlAs emitter, we removed this effect. However the surfaces that lie less then 40 mm away from a robot are represented only by values 245-250. In this way, for close measurement (30 mm) we get a flat horizontal diagram. Another undesired effect in small-average distances (40-70 mm) consists in a spontaneous decreasing of peak intensity (this is observable in all IR-diagrams in Figures 4 and 5). We cannot identify the nature of this nonlinearity and assume multiple IR-reflections as a reason for them;

3. *Nonlinear accuracy* of distance measurement. This requires nonlinear correction (it is done as a look-up table) of trigonometric relation in dependence of distance. However this nonlinearity is very "tricky". Even when a robot starts a measurement in the "good" area of 40-120 mm, a part of geometry can lie over 150 or 200 mm away. The effect of this nonlinearity appears in unreliable identification of many-surfaces geometry (Figure 5(c) "left to 1st. corner");

4. *Nonlinear rotation* of the robot. This can lead to different left γ_l and right γ_r slopes even for symmetric surfaces. The most easiest solution here is to calibrate γ_l and γ_r;

5. *Nonlinearity in measuring convex surfaces*. The identification of all types of convex geometries is performed by γ_l and γ_r. The difference between slopes for e.g. round objects (Figure 5(b)), convex corners (Figure 4(b)) and finite-size flat objects (Figure 4(a)) is small, moreover due to a nonlinear intensity diagram, these slopes change with distances. This problem has some basic character and we hardly belief that with all nonlinearities of IR-perception we are able to reliable identify the type of convex surfaces.

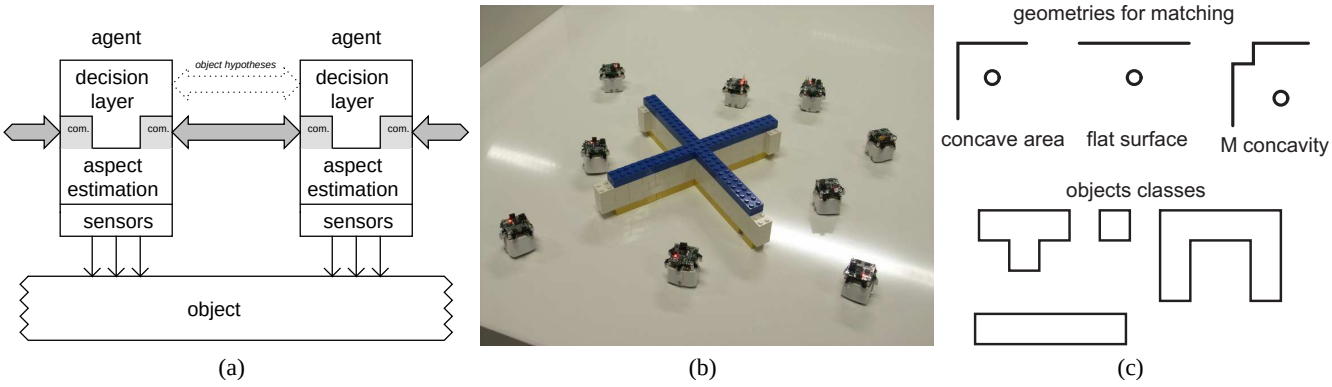
The main problem of these nonlinearities represents the necessity to maintain many look-up tables for corrections. This, in turn, is limited by a small memory of Atmel microcontroller. The assumption is that this problem can be solved in collective way. We can reduce the accuracy of individual recognition (so that to satisfy all hardware constraints) till such a degree which still allows a reliable collective recognition. Now, based on the discussed features and nonlinearities, we can briefly analyze the types of surfaces.

1. Surfaces with flat geometry. The flat type of geometry is primarily characterized by only one peak value on the IR-diagram. Finite-size surfaces are also characterized by large left and right slopes and scanning angle $\alpha \ll 120^\circ$, Figure 4(a). The size L_{vis} can be calculated as $2d\tan(\alpha/2)$, taking into account the "fuzzi edge" nonlinearity.

"Infinite-size" surfaces (Figure 5(a)) have small slopes of IR-diagrams and $\alpha \sim 120^\circ$. To absorb the nonlinearity of slopes for small and large distance, we apply the polygonal approximation [Pitas, 1993] and use in calculation the relation $\gamma_{\{r,l\}}/S_{\{r,l\}}$ instead of simple $\gamma_{\{r,l\}}$, where $S_{\{r,l\}}$ is the length of approximating line. In the performed experiments the probability of correct identification is very high and the accuracy of size calculation is of 5 mm (15 mm in the worst case).

2. Surfaces with convex geometry. Surfaces with convex geometry possess also only one peak value, however larger slopes then flat geometries. This type of geometry has to be identified before the calculation of size, which has no sense in this case. There are several types of convex geometry: convex corners and convex round surfaces (Figure 4(b)), convex many-surface geometry (can be recognized only collectively)(Figure 4(f)). We identify this geometry by $\gamma_{\{r,l\}}/S_{\{r,l\}}$ in the IR-diagrams. The difference between them points to a position in relation to a corner (left to a corner, right to a corner). The probability of correct identification of convex round geometry is very high, however convex corners are often classified as flat geometry. One approach to avoid this problem is the so-called "active exploration" (simple move towards the surface and scan again induces the appearance of a large "flat region" in the peak intensity which points to the flat type of geometry).

3. Many-surfaces and concave geometries. Concave geometries manifest primarily as multiple peaks in IR-diagrams. Based on the number of peaks we can differentiate between 2-concave (concave corners) and 3-concave sides geometry (Figure 4(c)). Concave many-surfaces geometries (Figure 4(b)) can be also classified by one robot. They have one peak value, however multiple left or right slopes. Many-



(a) (b) (c)

Figure 6: **(a)** *Distributed architecture for collective observation;* **(b)** *Spatial distribution of the robots around the observed object;* **(c)** *Geometries for matching and objects classes.*

surfaces geometry can also be composed from surfaces that belong to different objects. Generally, concave geometries can be identified with high reliability, however some fine differentiation between them is not always possible.

4. Estimation of probability. Since the robot cannot reliable classify the type of surfaces, it calculates a probability of correct classification. The calculation is done in the following way. We measure the possible values of α, d, $\gamma_{\{r,l\}}/S_{\{r,l\}}$, P_{imax} and estimate L_{vis} for all types of surfaces. The robot uses last square metrics to calculate the relation between the measured values and these presaved types. For collective perception a robot sends all possible classifications that have the probability over 30%.

Through the presented features of the IR-deep image we tried to classify several surfaces and to identify the classification probability as well, as base steps or components required for the individual perception.

5 Collective Perception

The described in the previous sections individual perception provides the sensor input for the collective perception. The approach for collective perception presented here proposes that each robot *talks* to its neighbors to exchange information about the surrounded object. In this task we limit ourselves only to the problem of *collective classification* [Pradier, 2005]. The robot possesses the objects models and have only to order the collective sensor input to one of the presaved model.

The distributed architecture for collective perception is shown in Figure 6(a). There is no privileged agent with a special role: all robots perform the same operations. The suggested method is homogeneous, i.e. all robots act the same and there is no need for a leader. Due to the homogeneous architecture the approach is robust, scalable, moreover new robots can join the team dynamically without any need to readjust any task assignment. Figure 6(b) shows how robots are deployed during collective observation. There are two possible implementations for the propagation of hypotheses: a single agent collects the information needed to identify an object by moving around it and performing the sensing operations; a single agent acquires local evidences and propagates hypotheses for the further fusion.

5.1 Object model

Given the limitations on the sensing capabilities of the robots, object classes can only be defined in terms of their geometries, as mentioned in Section 4. Figure 6(c) shows the 2D geometries of the four object classes which will be used subsequently. Once robots are situated around the object, they can estimate the local properties of the object as seen from their current positions called *viewpoints*. The actual measurement obtained from a viewpoint v can be noted as $S(v)$; $S : V \mapsto \text{feature vector}$ and represents the output of the distance sensors. Given an object class, it is possible to establish the expected sensor outputs for a number of views. A number of viewpoints $n_V^{K_i}$ for each object class K_i are chosen, along a trajectory situated in the center of the measurement domain, and noted as $\mathfrak{V}^{\mathfrak{K}_i} = \left(v_n^{K_i}\right)$. The corresponding expected measurements for objects of class K_i are $S\left(\mathfrak{V}^{\mathfrak{K}_i}\right) = \left(S\left(v_1^{K_i}\right), \ldots, S\left(v_{n_V^{K_i}}^{K_i}\right)\right)$. Therefore, the object model for a class K_i incorporates an ordered sequence of views for different *successive* positions around objects of that class. The starting position is arbitrary: only the ordering is relevant. The direction — clockwise or counterclockwise — can be chosen arbitrarily, but must be the same for all object models.

Additionally, object models include information about the reachability of different viewpoints, taking into account both geometrical constraints and the limitations imposed by the communication capabilities of the robots. It is noted as $W^{K_i} = \left\{ \left(v_j^{K_i}, v_k^{K_i}\right) \middle| \; v_k^{K_i} \text{ reachable from } v_j^{K_i} \right\}$. Finally, the corresponding distances between viewpoints in W^{K_i} are added to the object model, as $d_{\mathfrak{V}} : \mathfrak{V}^{\mathfrak{K}_i} \times \mathfrak{V}^{\mathfrak{K}_i} \to \mathbb{R}$.

The set of all canonical measurements — corresponding to sets of observable features, called *aspects* — in the model is noted $A = \left\{S\left(v_i^{K_j}\right)\right\}$ and its cardinality can be reduced by clustering the expected measurements. In that case, a sequence of canonical views could match several $(identity, position)$ pairs.

The goal of collective classification in a swarm of robots is

to estimate that class K_i the object being observed belongs to. When n_r robots are situated in an area surrounding the object (measurement domain) in positions $w_1, \ldots, w_{n_r}$, they are ordered implicitly depending on their position around the object as the perimeter of the latter is explored in a given trigonometric direction. Given these positions, the robots will measure $(S(w_1), \ldots, S(w_{n_r}))$. The proposed collaborative classification method will try to estimate the corresponding canonical viewpoints $\left(v^{K_n}_{n(1)}, \ldots, v^{K_n}_{n(n_r)}\right)$ given the above measurements; the end result, namely the class K_n the object belongs to, is implicit. *This means that not only the class of the object, but also the relative positioning of each robot can be obtained.*

Without any other a-priori information, and based only on the features observed by a robot, the latter can already generate an hypothesis regarding its current viewpoint, and implicitly which object it is observing, if the matched view is only present in that object model. If the observed features match closely the features corresponding to a view that is unique to an object class, the latter can be retained as a likely hypothesis for the whole object.

5.2 Hypotheses fusion

By observing the object from a given position, a robot can only generate local, basic hypotheses. Generally this is not enough to determine the class the object belongs to. The information obtained from different measurements should be fused via exchange of hypotheses between different robots. Amongst the many fusion processes introduced in the literature [Abidi and González, 1992; Hall, 1992; Klein, 1999], the Dempster-Shafer (DS) formalism [Hutchinson and Kak, 1992] was retained because it does not require a-priori class probabilities and is able to capture the notion of uncertainty. The often-cited drawback of the DS method is that its complexity grows exponentially with the cardinality of the primitive hypothesis set. However, due to the way hypotheses are generated from the object models, the complexity can be proven to be polynomial [Hutchinson and Kak, 1992].

Dempster-Shafer (DS) evidential reasoning [Shafer, 1976] is an extension to Bayesian inference that allows each source of information to contribute only to the evidence it has gathered, without overcommitting or trying to make hasty choices based on incomplete information. The Dempster-Shafer approach allows to express the lack of information by separating belief for a proposition from its mere plausibility, assigning probability masses to sets of propositions in such a way that the latter is free to move to any subset.

Probability mass assignment. Information sources can distribute probability masses among subsets of Θ, where Θ is the set of all statements about the possible outcomes of a random experiment. It is represented by the *frame of discernment* (FOD). The FOD is a set of mutually exclusive and exhaustive statements named *singletons*. When a probability mass is assigned to a set of singletons, it is free to move to any subset. Consequently, assignment of probability mass to Θ represents ignorance, since the probability mass can move to any element of Θ. When a source of evidence cannot differentiate between two propositions, it can assign a probability mass to a set including both.

The *probability mass assignment* function associates a probability mass to the sets in the power-set 2^Θ of Θ; it is therefore a function $m : 2^\Theta \longrightarrow \mathbb{R}$ verifying the following properties $m(\emptyset) = 0$, $0 \leq m(X) \leq 1$, $\sum_{x \in 2^\Theta} m(x) = 1$. The subsets $\{x_i\}$ of Θ such that $m(x_i) > 0$ are called *focal elements*; the union of those subsets is termed *core* of the probability assignment m.

Dempster's orthogonal sum. Two different sources of information will yield different mass distributions m_1 and m_2. Dempster's *rule of combination*, or *orthogonal sum*, can combine them if they are relative to the same FOD Θ, according to $m = m_1 \oplus m_2$, $m(X) = K \sum_{X_1 \cap X_2 = X} m_1(X_1)\, m_2(X_2)$. K is a normalization term defined as $K = \dfrac{1}{1 - \sum_{X_1 \cap X_2 = \emptyset} m_1(X_1)\, m_2(X_2)}$, which normalizes the new probability masses so that their sum is still unity. It can be seen as a measure of the degree of conflict between the two sources of information. When $\sum_{X_1 \cap X_2 = \emptyset} m_1(X_1)\, m_2(X_2) = 1$, the information is completely inconsistent and it is impossible to integrate it: the orthogonal sum is then undefined.

Hypothesis refinement. General, non-basic hypotheses are noted $H^{level} = \{(a_k, \ldots) | \, a_k \in A\}$. It is important to note that a_k could correspond to the output from several canonical viewpoints. The set of all possible hypotheses is noted H. Clearly the sequences of canonical measurements can only correspond to valid view sequences in some object model; impossible sequences, such as those having views that cannot belong to the same object, will not be generated.

In general, a robot will propagate its current beliefs about the object to the "next" neighboring robot along the perimeter of the object — initially $m(H^0)$. When this information is sent, the receiving robot can access the following:
- belief of the previous robot $m_1(H^n)$;
- distance to the robot whose message is being received d_{pre};
- its own beliefs about the observing part of the object $m_2(H^0)$.

The Dempster-Shafer combination rule for two hypothesis sets in a compatible frame of discernment $m(H_n) = \dfrac{\sum_{H_i \cap H_j = H_n} m(H_i)\, m(H_j)}{1 - \sum_{H_i \cap H_j = \varnothing} m(H_i)\, m(H_j)}$ is slightly modified to use the information about the relative positions of the robots as follows. Given an hypothesis set H^n, the refined hypotheses will be $H^{n+1} = \{U(h \oplus a) | \, h \in H^n, a \in A, h \oplus a \in H\}$, where the last condition means that the new view sequence must be possible for at least one object class. The operation $\oplus : H \times A \longmapsto H$ is defined as $h \oplus x = (h_1, \ldots, h_n, \ldots, a_{m_1}, \ldots, a_{m_p}, x)$, where the views $a_{m_1}, \ldots, a_{m_n}$ are a "filler", and x is the view that is to be added to the sequence. An additional restriction can be imposed to the $\oplus$ operation, namely that the filler has to be no longer than some arbitrary number of viewpoints k with $p < k$ in the above expression. The output of the function $U(h)$ is defined as the shortest hypothesis equivalent to h, that is, an hypothesis that corresponds to the same $(\mathrm{object}, \mathrm{offset})$ matches.

Algorithm 1 Hypothesis refinement (pseudocode).

```
function new_hypotheses(incoming_message)
 in_hypotheses = decode(incoming_message)
 new_hypotheses = {}
 for hypothesis in in_hypotheses do
  for feature in initial_feature_estimates do
   if exists_successor(hypothesis, feature)
    for succ in successors(hypothesis, feature)
     pm = hypothesis.pm * feature.probability *
      distance_factor(incoming_message.distance, succ.dist)
     if succ not in new_hypotheses
      add succ to new_hypotheses with p.m. pm
     else
     add pm to the probability mass of succ in new_hypotheses
     end endfor endif endfor endfor
 trim_hypotheses(new_hypotheses)
 total_pm = sum of all prob. masses in new_hypotheses
 for hypothesis in new_hypotheses
  hypothesis.pm /= total_pm endfor
 return new_hypotheses
endfunction
```

The new probability mass assignment is calculated with $m'\left(h_i^{n+1}\right) = \sum_{h^n \oplus x = h_i^{n+1}} m_1\left(h^n\right) m_2\left(x\right) \xi\left(d_{pre}, d_{model}\right)$, $m\left(h_i^{n+1}\right) = \frac{m'\left(h_i^{n+1}\right)}{\sum_k m'\left(h_k^{n+1}\right)}$, where an additional normalization is required due to the usage of the distance term $\xi\left(d_{pre}, d_{model}\right)$. The latter reuses the known distances between the last canonical viewpoint of h^n and the viewpoint that is chosen to match x. ξ is taken as the normal distribution $\xi\left(d_{pre}, d_{model}\right) = \frac{1}{\sqrt{2\pi}\gamma d_{model}} e^{-\frac{\left(d_{pre} - d_{model}\right)^2}{2\gamma^2 d_{model}^2}}$ whose standard deviation depends on the expected distance, to cope with the increasing inaccuracy as the latter grows; in practice, values around $\gamma \sim 0.5$ yield good results. The overall process is described in Algorithm 1.

5.3 Hypothesis encoding and compression

Once a number of robots have acquired information about the object they are observing, hypotheses can be refined through exchanges. The associated communication cost is proportional to the volume of data being communicated. It is possible to bound the cost of the communication associated to collective classification as follows. It can be seen that there can only be at most $n_V = \sum_i n_V^{K_i}$ hypotheses being considered at any point in time, representing the number of differentiable object identities and poses. The information about the hypothesis to be transmitted can be encoded either by explicit encoding on a per-hypothesis basis, or by factoring out information common to multiple hypotheses and using implicit information (like ordering) across message fragments.

Per-hypothesis encoding. A unique identifier for each hypothesis can be encoded using only $\left\lceil \log_2 \sum_i n_v^{K_i} \right\rceil$ bits. Due to memory constraints, hypotheses can be encoded alternatively as $h = \langle i, l, o \rangle$, where K_i is an object model, l is the number of aspects of the hypothesis and o is the offset in the canonical views sequence. Thus, each hypothesis can be encoded in at most $\lceil \log_2 |K_i| \rceil + \left\lceil 2 \log_2 \max_i n_V^{K_i} \right\rceil$ bits. The amount of information transmitted for k hypothesis is directly proportional to the latter, resulting in $k\left(\left\lceil \log_2 \sum_i n_v^{K_i} \right\rceil + c_{pm}\right)$ bits normally, where c_{pm} is the amount of bits needed to encode the probability mass itself.

Implicit encoding. Only *hypothesis selector* can be sent, indicating which hypotheses are actually transmitted and a sequence of probability masses. It consists of n_V bits: the n-th bit specifies if the probability mass of the hypothesis whose identifier is n is attached to the message. In order to transmit k hypotheses $\sum_i n_v^{K_i} + k c_{pm}$ bits are needed. Therefore this encoding approach is only practical when $k \gg 1$, that is, when a large number of probability masses are to be transmitted, so the overhead is amortized.

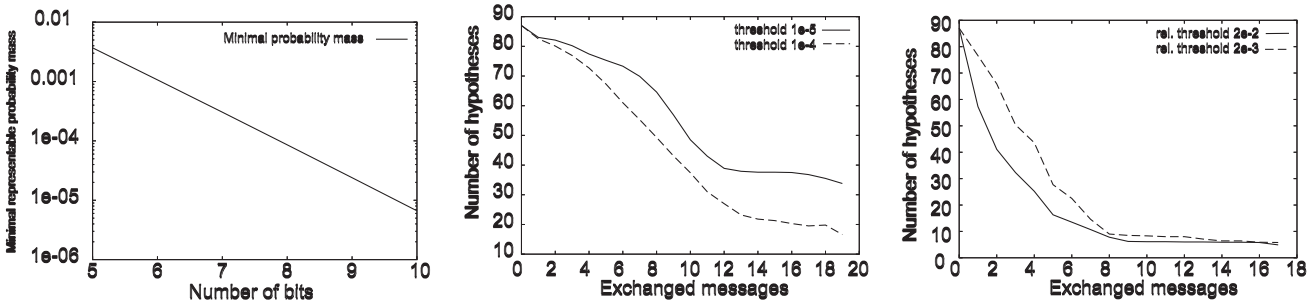
Linear encoding. If a simple linear, fixed-point scheme is employed, and the resolution is chosen to be a fraction of the average probability mass $\frac{\overline{m}}{k}$, as many as $\left\lceil \log_2 \left(\frac{k}{\overline{m}}\right) \right\rceil$ bits would be needed. For a reasonable value of $k = 10$ and $n_V = 60$, a fixed-point encoding would require $\lceil \log_2 (10 \times 60) \rceil = 10$ bits per probability mass.

Dynamic range compression. ITU-T G.711 [ITU, 1988] introduces two *compression* algorithms based on the following key idea: the signal is compressed according to a logarithmic expression. The simplest one, μ-law, applies the transform $y = \operatorname{sign}(x) \frac{\ln(1+\mu|x|)}{\ln|1+\mu|}, -1 < x < 1$ where μ is chosen according to the desired output resolution; for 8 bits, $\mu = 255$. The similarity with probability mass encoding is striking. Indeed, based on the μ-law expression, probability values can be encoded using $E(m) = 2^n \frac{\ln(1+(2^n-1)m)}{n \ln(2)}$ so that the encoded probability mass fits in n bits, and the distortion ratio is minimal. Figure 7(a) shows the minimal representable probability mass for different encoding lengths.

Hypothesis set compression. Regardless of the method used to encode the hypothesis set, the cost, in terms of amount of information to be transmitted, grows with the number of hypotheses propagated. It is thus desirable to minimize the cardinality of the hypothesis set before transmission. This can be performed either by *trimming* (discarding hypotheses whose probability mass is comparatively or in absolute terms small) or by *coalescing* (grouping several hypotheses into one corresponding to the union of the corresponding propositions).

Trimming. The cardinality of an hypothesis set can be reduced by simply ignoring unlikely hypotheses. The simplest way is retaining only hypotheses whose probability mass is higher than some absolute threshold. Figure 7(b) shows the cardinality of the hypothesis set when the latter is trimmed according to different absolute thresholds. The hypothesis can also be made smaller by removing all the hypotheses whose probability mass is below a threshold relative to the most likely hypothesis, i.e. those that satisfy $pm < r \times \max_i pm_i$, where r is the relative threshold and pm_i are the probability masses. The performance of this method is illustrated in Figure 7(c).

Coalescing. It is possible to further minimize the cost of transmitting an hypothesis set by transferring only some hypotheses. Those not specified explicitly can be coalesced into



(a) (b) (c)

Figure 7: **(a)** *Minimal representable probability mass using dynamic range compression;* **(b)** *Number of hypotheses with reduction based on absolute threshold;* **(c)** *Number of hypotheses with reduction based on relative threshold.*

a more general hypothesis carrying the sum of the probability masses using $m(X) = \sum_{Y \subseteq X} m(Y)$. This scheme also makes the overall communication more robust, since it can be interrupted without adverse consequences at any point during the transmission of the probability masses.

Termination. When information fusion is successful, the whole group of robots will converge as a whole towards a common decision regarding the nature of the object. The final decision of each robot can be taken as the hypothesis with the highest associated probability mass. It is therefore necessary to know when a given hypothesis set can be considered as "refined enough". The key idea is that hypothesis refinement can be considered finished when enough evidence has been collected, i.e. the ambiguity of a set of hypotheses is larger than a given threshold. [Hutchinson and Kak, 1992] defines the ambiguity of an hypothesis set, closely related to the concept of entropy in information theory, as follows: $A(\Omega) = -K \sum_{\theta \in \Omega} p(\theta) \log p(\theta)$, $p(\theta) = \frac{\sum_{\theta \in H} m(H)}{|H|}$. This definition takes into account the fact that an hypothesis might correspond to several individual statements or singletons. It can be seen than the ambiguity measure of the probability mass assignment $\{\Omega \Longrightarrow 1\}$, i.e. complete ignorance, corresponds to the entropy of an equiprobable distribution over $|\Omega|$ possible outcomes.

6 Preliminary experiments and discussion

Preliminary experiments have been performed with 10 prototypes of the microrobots Jasmine in the field of individual and collective perception. In experiments we measured the feature extraction and surface's recognition, as described in Section 4 and collective hypothesis refinement, as described in Section 5. The robots are placed in the situations like those depicted in Figures 4, 5. Table 1 contains the probability mass assignments for the three stored patterns "flat surface", "concave area" and "M concavity", represented in Figure 6 (c). The calculated probabilities from experimental scans confirm the results predicted by the simulation. The collective classification process was tested in hybrid approach, where the real scan data are taken from the microrobots, however the hypothesis fusion was performed in the host computer. The reason is a lack of bidirectional communication in the prototypes, that is currently under improvement. Figure 8(a) shows

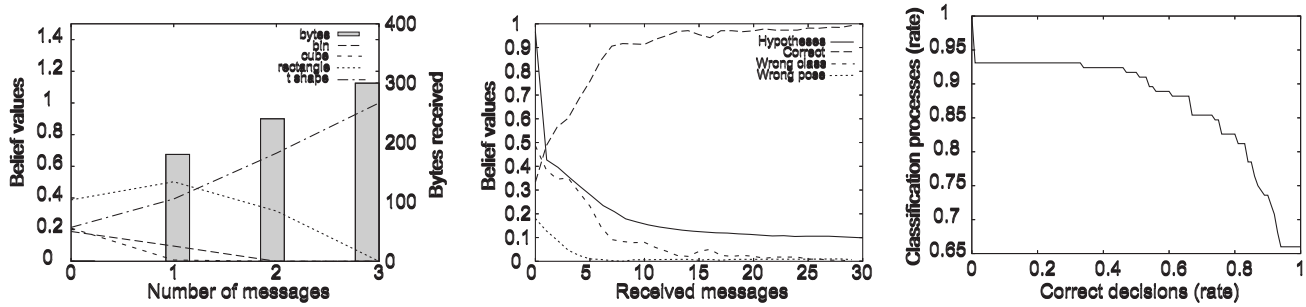
Feature	Distances			Probability masses		
	Flat	Concv.	M	Flat	Concv.	M
Concavity	1020	543	1096	0.26	0.49	0.24
	765	872	1359	0.41	0.36	0.23
	664	764	1251	0.42	0.36	0.22
	1275	861	995	0.27	0.39	0.34
	702	215	1105	0.20	0.67	0.13
	1020	1020	1418	0.37	0.37	0.26
Flat surface	258	812	1864	0.69	0.22	0.10
	259	954	1846	0.71	0.19	0.10
	510	872	1862	0.54	0.31	0.15
M concavity	1785	1785	1646	0.32	0.32	0.25
	1530	1343	789	0.25	0.28	0.48
	1436	1288	1190	0.30	0.34	0.36
	1444	1331	895	0.27	0.29	0.44
	1530	1376	1053	0.28	0.31	0.41
	1624	1570	1312	0.31	0.32	0.38
	1457	1294	861	0.26	0.29	0.44
	1275	1061	559	0.22	0.27	0.51

Table 1: Probability mass assignments according to Jasmine's scan data.

the belief of a robot after its initial estimation, which is based only on the information obtained via distance sensors, and after reception of messages from other robots. The belief values converge quickly towards the correct value.

Figure 8(b) illustrates the evolution of robots placed around a "T shaped" object. The curves "correct", "wrong class" and "wrong pose" indicate respectively the fraction of robots that took the correct decision, those which made a mistake in the class of the object, and finally those which were able to determine the class of the object correctly but could not estimate their relative positions accurately. The graphs corresponds to an average value for several successful processes.

Figure 8(c) shows the success rate for different convergence rates. It can be interpreted as follows: a pair $\left(\frac{x}{100}, \frac{y}{100}\right)$ in the curve means that in y percent of the runs the rate of correct decisions remained stable at x percent or higher after thirty message exchanges. We can therefore see that in around 66% of the processes all robots took the right decision regarding the object identity and their relative position



(a) (b) (c)

Figure 8: **(a)** *Evolution of the classifying estimations of a robot. The belief value evolves as the robot obtains information from its peers, while observing an object of class "T shape";* **(b)** *Convergence in successful collective classification processes (T);* **(c)** *Success rate for different convergence thresholds.*

(the rate for a convergence equals to or greater than 80% exceeds 82%), more than one half of the robots reached correct decisions regarding both object identity and position in over 90% of the classification operations. The group of microrobots converged towards a wrong decision regarding the identity of the object in around 5% of the classification processes. Around 10% of the classification processes end up with less than one robot out of ten with correct identity but wrong positional decisions. Around 15% of the classification processes failed to converge to either a correct decision within a 20% rate or to an erroneous decision.

Summary. In this paper we addressed the specific problem of perception in a swarm of microrobots. We investigated the process of individual perception by designing and implementing the IR sensory system. We researched also the problems related to IR-based perception and developed/tested the hardware and the corresponding algorithms allowing sensing and classifying the geometry of the surfaces. The collective classification was performed by fusing local hypotheses by using a formalism based on the Dempster-Shafer evidential reasoning. Communication needs were analyzed. Experiments demonstrated that, the size of robot is scaled down (over 20 times in comparison with the middle-size league in RoboCup), however the microrobot still possesses cognitive features. However we also observe that the smaller the size (the more reduced capabilities) of a separate robot is, the more functionalities can be achieved only in collective way.

References

[Abidi and González, 1992] Mongi A. Abidi and Rafael C. González, editors. *Data fusion in robotics and machine intelligence*. Academic Press Professional, Inc., 1992.

[Bonabeau *et al.*, 1999] E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm intelligence: from natural to artificial systems*. Oxford University Press, New York, 1999.

[Caprari and Siegwart, 2003] G. Caprari and R. Siegwart. Design and control of the mobile micro robot alice. In *Proc. of the 2nd Int. Symposium on Autonomous Minirobots for Research and Edutainment, AMiRE'2003*, pages 23–32, 2003.

[Hall, 1992] David Lee Hall. *Mathematical Techniques in Multisensor Data Fusion*. Artech House, Inc., 1992.

[Hutchinson and Kak, 1992] S. A. Hutchinson and A. C. Kak. Multisensor strategies using dempster-shafer belief accumulation. In *Data fusion in robotics and machine intelligence*, chapter 4, pages 165–209. Academic Press Professional, Inc., 1992.

[I-Swarm, 2003 2007] I-Swarm. *I-Swarm: Intelligent Small World Autonomous Robots for Micro-manipulation, 6th Framework Programme Project No FP6-2002-IST-1*. European Communities, 2003-2007.

[ITU, 1988] ITU. G.711 pulse code modulation (pcm) of voice frequencies. nov 1988.

[Klein, 1999] Lawrence A. Klein. *Sensor and Data Fusion Concepts and Applications*. Society of Photo-Optical Instrumentation Engineers (SPIE), 1999.

[Kornienko *et al.*, 2004] S. Kornienko, O. Kornienko, and P. Levi. Generation of desired emergent behavior in swarm of microrobots. In *Proc. of the 16th European Conf. on AI (ECAI 2004), Valencia, Spain*, 2004.

[Kornienko *et al.*, 2005] S. Kornienko, O. Kornienko, and P. Levi. Collective ai: context awareness via communication. In *Proc. of the IJCAI 2005, UK*, 2005.

[Kube, 1996] C.R. Kube. A minimal infrared obstacle detection scheme. *The Journal for Robot Builders*, 2(2):15–20, 1996.

[Megabitty, 2005] Megabitty. see http://groups.yahoo.com/group/megabitty/. 2005.

[Pitas, 1993] I. Pitas. *Digital Image Processing Algorithms*. Prentice Hall, 1993.

[Pradier, 2005] M. Pradier. *Collective Classification in a Swarm of Microrobots*. Master Thesis, University of Stuttgart, Germany, 2005.

[Shafer, 1976] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[Suzuki *et al.*, 1995] S. Suzuki, H. Asama, A. Uegaki, S. Kotosaka, T. Fujita, A. Matsumoto, H. Kaetsu, and I. Endo. An infra-red sensory system with local communication for cooperative multiple mobile robots. In *Proc. of International Conference on Intelligent Robots*, pages 220–225, 1995.

[Ye *et al.*, 2002] Y. Ye, S. Boies, J. Liu, and X. Yi. Collective perception in massive, open, and heterogeneous multi-agent environment. In *Proc. of the AAMAS'02*, pages 1175–1182, Bologna, Italy, 2002.