



**HAL**  
open science

## Ant-Cam network: Tracking across cameras using SVT parameters

Lobna Ben Khelifa, Jean-Charles Quinton, François Berry

► **To cite this version:**

Lobna Ben Khelifa, Jean-Charles Quinton, François Berry. Ant-Cam network: Tracking across cameras using SVT parameters. International Conference on Distributed Smart Cameras, Sep 2017, Stanford, United States. pp.138-143, 10.1145/3131885.3131930 . hal-01966729

**HAL Id: hal-01966729**

**<https://hal.science/hal-01966729v1>**

Submitted on 29 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Ant-Cam network: Tracking across cameras using SVT parameters

Lobna Ben Khelifa\*  
Pascal Institute  
Clermont University / CNRS  
Clermont-Ferrand, France

Jean-Charles Quinton  
Laboratoire Jean Kuntzmann  
Univ. Grenoble Alpes / CNRS  
Grenoble, France

François Berry  
Pascal Institute  
Clermont University / CNRS  
Clermont-Ferrand, France

## ABSTRACT

Over the last years, the rapid growth of distributed smart cameras has triggered the search for new approaches of smartness of cameras to have better results. As communication among camera entities is becoming more and more complex and new ways of modeling communication have been proposed. These new ways have been taking inspiration from different fields such as socio-economic approach or game theory. Moreover, one of the major problems of the camera network is re-identification. However, in most cases, the interaction between cameras presupposes that the latter are able to perform perfect and unambiguous detections, thus limiting the decision tasks to the Markovian model. Within this paper, we present a new approach of interaction between cameras based on a non-Markovian model. To resolve this issue, we can exploit other types of information rather than visual information to improve re-identification. This information is Spatial, Visual and Temporal (SVT). Temporal information holds the time needed to go from one camera to another, while spatial information contains the path followed by the target which is a key point for the decision-making process. This offers the possibility for the network to learn regularities and then reach a steady state.

## Keywords

Visual sensor network, smart network, smart camera, distributed problem solving

## 1. INTRODUCTION

Smart Camera Networks (SCNs), with their continuous evolution in the recent years, have attracted researchers' attention at the crossroads of various fields: 1) Image sensors and signal processing relate to the architecture of the camera, and aim to process rich information within an embedded system. These fields define the hardware parameters

\*email: lobna.benkhelifa@gmail.com

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

ICDSC 2017, September 5–7, 2017, Stanford, CA, USA

© 2017 ACM. ISBN 978-1-4503-5487-5/17/09...\$15.00

DOI: <https://doi.org/10.1145/3131885.3131930>

of the camera, such as zoom or focus capabilities. 2) Computer vision focuses on the smartness of the camera, what we may call software parameters related to object detection and scene understanding. 3) Sensor networks not only manage the wired or wireless communication between cameras (with signals not limited to simple binary events), but also rely on distributed computing to take advantage of heterogeneity in smart cameras. They pool the distributed resources and capabilities to better perform the tasks.

The SCNs classically rely on high-end cameras. Nevertheless, fully exploiting the interactions between several cameras makes it possible to use very low-specification cameras (e.g. low resolution and limited processing capability). In the animal kingdom, ants reach high levels of [social] organization thanks to their collaborative capabilities. They also achieve wonders at the large scale which are hardly conceivable when considering the individual size and strength of ants. Inspired from the world of ants, we develop networks based on a communicating low-specification camera, named Ant-Cam [1]. We focus in this paper on the cameras interactions and their collaborative performances, in order to solve complex problems that extend far beyond the scope of any individual camera.

A rich set of mathematical tools for decision making can be applied to model the interactions between cameras [9]. The associated approaches can be labeled as socio-economic [5], game-theoretic [19], stochastic [22] or optimization oriented (e.g., particle swarm in [14]). Added to that, they can rely on other forms of meta-heuristics such as the genetic algorithms [4]. The survey in [15] offered a deeper understanding of the existing models used for the SCNs. Some models would increase flexibility even more by exploiting self-reconfigurable [6] and self-calibrating cameras to maximize their performances. The resulting network might adjust camera parameters such as position [18], orientation, pan-tilt-zoom [14], or select where processing should occur. Indeed, a reduced (dynamical) set of cameras with dedicated processing might be sufficient and could better achieve tracking [5], coverage [14], path planning [18] or target detection under various visibility conditions and satisfy given quality requirements [16]. Yet, it would offer heterogeneity in the network and more efficiency while the cameras could learn how to be different [8], [7].

Most of the proposed approaches require some prior knowledge of the network such as positions, Fields Of View (FOVs). In this work, we present a model following the principle of smart dust, where Ant-Cams are scattered in the environment without any a priori knowledge about it. In addition,

some networks use supervised learning to classify usual and unusual behaviors so as easily detect targets and anomalies [13]. However, we opt for unsupervised learning [2], [3].

The main goal of this work is to learn and exploit the regularities in the correlated activity of cameras. The system should be able to build precise predictions based on two components: a model (of the spatio-temporal behavior of expected/previous targets) and observations (related to the current target).

To identify and track a target through a network, the observations inform about the path followed by the target. The system should pick out as much information as possible to better predict the next state (hence not being limited to the Markov assumption). The model in turn depends on the network structure and the assumptions about the trajectories or targets. Thereby, it should be built from observations, indicating which cameras should observe the target at future times, allowing to provide a multi-camera behavior analysis [11]. The survey [21] tackles different issues and aspects of re-identification challenges.

Our model includes two principal parts: The first one is the cognitive knowledge which permits prediction decision using the the detected information and that received from other cameras. The second one is the regulation of that cognition using the feedback received from the other cameras after each event. This cognition control allows the network to be self-monitoring and then to have its own self-regulation process.

Accordingly, the camera is not just a member of the network learning the parameters that may influence its performance, but it extends its capacity to perform things based on the interaction with the other cameras. Consequently, this model enables the cameras to do some tasks more automatically and then to go further with the control of knowledge. This knowledge can be consolidated by the other cameras presented in the network. [?] proposes a Network Consistent Re-identification (NCR) framework which improve the camera pairwise re-identification performance between camera pairs, this performance has been evaluated in [10]. Keep exchanging signatures with the neighborhood until finding the valid matches and then improve next reidentifications is proposed in to improve this performance [17]. Moreover, targets can hardly be viewed in a similar pose by two cameras, simple comparison between the two views can not lead to accurate results. [12] proposes to find the optimal correspondence between images patches using a sparse-based local matching technique.

In our previous work [1], we presented a transition probability for each path independently as we used a non-Markovian model; i.e, the probability of transition was represented by a 3D tensor (2D for the cameras, and 1D for the path where each path was indexed). In this current work, we significantly improve our preliminary work as follows:

- We propose an improved model and an easier computing method without needing to adjust the tensor size each time we have another path. As a result, there is no need to adjust all probabilities presented in a network.
- We take into account the fact that the previous re-identifications are not pinpointed, so the probability of the path is straightened by the preciseness of the previous re-identifications.

The remainder of the paper is organized as follows. First, we describe in section 2 the environment of the Ant-Cam model and introduce the model. Then in section 3, we explain the re-identification method. Next, we present the correlation between the cameras in the network and how they respond to an event in section 4. After that, we present in section 5 the experimental evaluation results and we discuss possible improvements. Finally, a conclusion is drawn in section 6.

## 2. ANT-CAM MODEL

We consider a set of cameras scattered in the environment without any prior knowledge about their positions, FOVs or neighbors. Like ants, the Ant-Cams will communicate depending on their needs (to obtain information in return) and experience (to pro-actively share information that may help others). For instance, when a camera detects a target and anticipates its local trajectory, it broadcasts or selectively warns other cameras, so that they can be operational and in turn better detect the target. Let  $\{c_i\}$  be a fixed set of cameras. Let  $o$  be an object of interest following a trajectory in a network of cameras, thus generating a sequence of events for a subset of cameras. We note  $o_i^t$  the actual event of the object  $o$  passing in front of the camera  $c_i$  at time  $t$ . For each camera  $\{c_i\}$ , the input at time  $t$  is noted  $z_i^t$ , which may be composed of an arbitrary number of components (e.g. color, speed, distance...). We define the entire set of inputs for all cameras at time  $t$  as  $z^t$  and inputs for all times up to time  $t$  (i.e.  $t, tdt, \dots, 0$ ) as  $Z^t$ . Similarly to  $z$ , we finally note  $\hat{z}$  the model of the object  $o$ , represented by the theoretical observation(s) from the perspective of the network cameras. The representation of any target traversing or moving within the network can therefore be approximated by:

$$o^t \equiv \{\hat{z}^t, \hat{z}^{t-dt}, \dots, \hat{z}^0\} \quad (1)$$

where each component represents the theoretical observation of the target in previous instants, and where a sequence of such observations may be required to discriminate between objects (adopting a non-Markovian behavior). To predict a target trajectory, the goal of the network is then to determine the next probable camera(s) that should detect the object ( $o^{t+dt}$ ), based on previous observations  $\{z^t, z^{t-dt}, \dots, z^0\}$ .

The role of the camera is not only limited the use of its own resources to perform object detection, but also to the prediction of where the object may go. The content and quality of the prediction itself depends on the target model and possible information received from other cameras. Appraising this prediction and its efficiency is central in the context of this work. Interactions are required to learn and adjust the predictions to the observed dynamics. For this purpose, the cameras keep exchanging information even after the target moves away, sending or integrating feedbacks on the quality of the prediction. Such signals are needed at the learning stage, but may also help to plan the network activity and to estimate online the prediction quality.

All connections are illustrated in Fig. 1, and we can further refine the different input events received by each Ant-Cam as follows:

- *Observation*: An external event (I3 on Fig. 1) will be

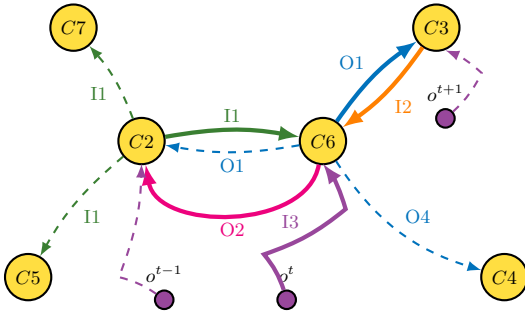


Figure 1: Illustration of possible links between cameras

generated when a target is detected by the Ant-Cam.

- *Prediction*: An internal event (I1) results from the activity in previous nodes (in terms of object trajectories) and predicts a future event to be observed by the current camera.
- *Acknowledgment*: An internal event (I2) from the following nodes confirms that a predicted event was received.

The output events follow the same logic, but there is no external event (since no action is now taken on the environment at the camera level):

- *Prediction*: An internal event (O1) sent to neighbors predicts that a specific event should be received at a given time.
- *Acknowledgment*: An internal event (O2) sent back to the previous node confirms a predicted event was indeed observed. It also conveys information to help improve the prediction of future similar events.

Those events help the camera to create its own identity in the network and get the capacity to be self-critical and self-regulated.

### 3. RE-IDENTIFICATION METHOD

For a predator attack, aggression or food detection, ants try to warn the colony using different kinds of signals (chemical, auditory...) in order to specify the needs so that others can prepare themselves for help. This collaboration allows solving complex problems. For this reason, the ant world is always a source of inspiration and subject of scientific studies. We present here a network model, fully inspired from the ant society. From a set of limited performance cameras, we want to achieve a high level of efficiency. Following the ant metaphor, the Ant-Cams communicate using pheromones which are discharged into the network and propagated to the other Ant-Cams, so they can warn the colony. This pheromones contain different kinds of information : visual  $\nu_r$ , temporal  $\tau_m$  and spatial  $p_{path}$ . Basically, we model and study the camera responses to those events. For instance, after receiving the pheromones, the camera can process an observation event after a given delay. Once the camera receives the event, it will acknowledge and potentially correct the prediction to the previous node, while continuing spreading prediction events to subsequent cameras. This leads to :

- *Pre-event connectivity*
- *Event connectivity*
- *Post-event connectivity*

The Pre-event connectivity is the reception of the prediction I1 in Fig. 1. It contains the three parameters presented above : the visual features  $\nu_r$ , the temporal information  $\tau_m$  and the spatial information  $p_{path}$ . This will be the pheromone broadcast from

The Event connectivity is triggered by target detection. The camera receives the pheromones containing three kinds of information, and then compares the pheromones received with the information extracted from the detection and gets the re-identification confidence:

**Visual features:** They correspond to feature vectors for an observation  $z_i$ . They prototype the current object and contain the target color, the velocity and the category. The visual confidence  $\phi_v$  is then estimated. It represents the similarity between the observation  $z_i^t : \nu_r$  and the prediction received from the other camera j representing  $z_j^{t-dt} : \nu'_r$ . This confidence is defined by:

$$\phi_v = f_v(\nu'_r, \nu_r) = \exp \frac{-|\nu'_r - \nu_r|}{\alpha^2} \quad (2)$$

where  $\alpha$  is a scaling factor.

**Temporal information:** It contains information about detection time. Typically,  $\tau_m$  is the time needed for a target to move from one camera to another. The temporal confidence  $\phi_t$  represents the similarity between the predicted time  $\tau_m$  and the observed time  $\tau_r$  to move from one camera to another. It is defined by:

$$\phi_t = f_t(\tau_m, \tau_r) = \exp \frac{-|\tau_m - \tau_r|}{\alpha^2} \quad (3)$$

**Spatial information:** It is related to the path through the network. The spatial confidence  $p(path)$  is defined by the probability that the target gets through the path s. This information is defined by:

$$\phi_s = f_s(p(s)) \quad (4)$$

Let  $C_5 > C_2 > C_6 > C_4$  and  $C_7 > C_2 > C_6 > C_3$  be the most used paths in the network. A target presented in  $C_6$  will require an estimation of the next state ( $C_3$  or  $C_4$ ), which cannot simply be deduced from the current camera or even when considering the previous camera ( $C_2$  in both cases). Whereas, considering the sequences of previous cameras will disambiguate the trajectories.

This parameter is changed compared to our previous work. In the past, it was limited to an index number of the path s followed by the target. As a consequence, we had to adjust the matrix dimension each time a new path was followed, which might complicate the calculation. To confirm the reception of the target in the event connectivity step, we condition the system to decide whether it is the same or not. For this, we fix a similarity threshold where only the similarity confidences equal to or above it are considered the same.

As presented in the previous section,  $\hat{z}^t$  is the observation of the object o at an instant t by the n cameras presented in the network, which can be presented on  $\hat{z}^t = \{ z_1^t, \dots, z_n^t \}$ .

Therefore, those components are binary values representing whether the object is detected or not by a camera. Thus,  $\hat{z}^t$  will straightforwardly correspond to the camera detecting that object. Accordingly, (1) becomes :

$$o^t \equiv \{z_i^t, z_j^{t-dt}, \dots, z_k^0\} \quad (5)$$

Furthermore, to be more precise, we choose not to limit the observation to binary values here, but we extend it to the probability of this observation. As a result, each observation  $z_i^t$  is represented by  $p(o_i^t|z_i^t)$ .

Those parameters are used to estimate how much does the detected object correspond to the received prediction. This estimation is evaluated by :

$$p(o_i^t|Z^t) = \frac{p(z^t|o_i^t, Z^{t-dt})p(o_i^t|Z^{t-dt})}{p(z^t|Z^{t-dt})} \quad (6)$$

where  $p(z^t|o_i^t, Z^{t-dt})$  is the observation likelihood. It is the probability of getting the camera input if we know that the object is present in front of it at time  $t$ . We assume that the current and past observations are independent from each other and conditioned on the location. In other words, the appearance of the object from the perspective of the camera should be independent from the path followed by the object. Hence, it becomes equivalent to  $p(z^t|o_i^t)$  and evaluated relying on the similarity defined in the feature domain  $\phi_v$  defined by Ea. 2.

$p(o_i^t|Z^{t-dt})$  is the prior belief about the location of the object in the network, conditioned by all previous observations on all cameras. This term can be evaluated from  $p(o_j^{t-dt}|Z^{t-dt})$ , where  $dt$  is the time of the previous event in the camera network. For this, we need to consider the set of observed/memorized paths in the camera network, with associated delays (limited to a fixed horizon, for example. sequences of four nodes in the network). We evaluate this term using the transition matrix [1] with the adjusted probabilities having the temporal similarity  $\phi_t$  (Eq. 3) and the spatial information  $\phi_s$  (Eq. 4).

$p(z^t|Z^{t-dt})$  is a normalizing term. The latter is independent from the camera at all times. Subsequently, We obtain:

$$p(o_i^t|Z^t) \propto p(z^t|o_i^t)p(o_i^t|Z^{t-dt}) \quad (7)$$

## 4. TRANSITION BETWEEN CAMERAS

Ant colonies are described as superorganisms because ants appear to operate as a unified entity, collectively working together to support the colony. In this respect, without any central coordination or a prior knowledge about the network topology, each camera interacts with the neighbors to build up its knowledge, and to successfully contribute to the self-organization of the network and achieve their objectives. This adaptation is done continuously using a different interaction and accompanying environmental dynamics. Once the target is re-identified as described in the previous section, a feedback will be sent to confirm or not the event. This is what we defined as a post-event connectivity. It will enable the adjustment of the probabilities between the cameras once a feedback confirming the reception of the target is received. After each observation in any camera, the model of the target object, transitions and delays can be updated. Let us consider a set of cameras presented in Fig. 2 and an object observed by  $j$  and then observed by  $i$ . Thus, the probability of this transition is estimated by  $j$  after receiving the

feedback from  $i$ :

$$p(i|s = jklm) = \frac{p(s|i) * p(i)}{p(s)} \quad (8)$$

where  $s$  is the path followed by the target before arriving to  $j$ , and  $p(s)$  is a tracking confidence highlighting the certitude that the target gets through the path  $s$ . In fact, it is obtained by the cameras building  $s$  and inversely normalized to  $[0,1]$ . The number of cameras taken into account to build the path should be fixed at the beginning.  $p(i)$  here corresponds to  $p(o_i^t|Z^t)$  presented above and  $p(s|i)$  is posteriori probability of coming from a path  $s$  when arriving at  $i$ . It corresponds to the network history and structure and to the previous targets following the same path.

This connectivity offers the possibility to learn about the spatial relationship between cameras while tracking. Having such information gives the opportunity to build the physical layer graph, L2 in Fig. 2, while the communication layer graph L1 can be found out once the system starts tracking. Having the second layer L2 will reduce the communication cost without significantly sacrificing the network performance, as the cameras will not need to broadcast the information to the whole network any longer.

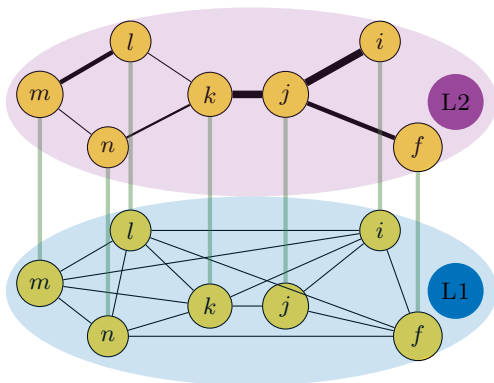


Figure 2: Instance of network: Connectivity graph based on events (L2) and communication graph allowed by technology used for it(Wifi, LoRa..)

## 5. EVALUATION

In this section, we evaluate our model using a simulation platform. The system is composed of a set of nodes able to accomplish two tasks : (i) detecting and excerpting the necessary information and (ii) dispatching the pheromones to other cameras. This evaluation evinces how the system can coordinate the interactions of nodes after each event pattern. This simulation exhibits the coordination and collaboration of the nodes to reach a stable state despite the environment difficulties or some system failures such as losses of messages, detection and re-identification problems.

The simulation is performed using the discrete event simulator OMNeT++ [20]. We simulate sensor nodes, without any specifications about their capacities. For instance, those nodes can be camera nodes able to detect targets. However, what we want to highlight with this simulation is the capacity of using different parameters to re-identify the targets in addition to the comparison of visual features. It demonstrates how the system can reach a stable state

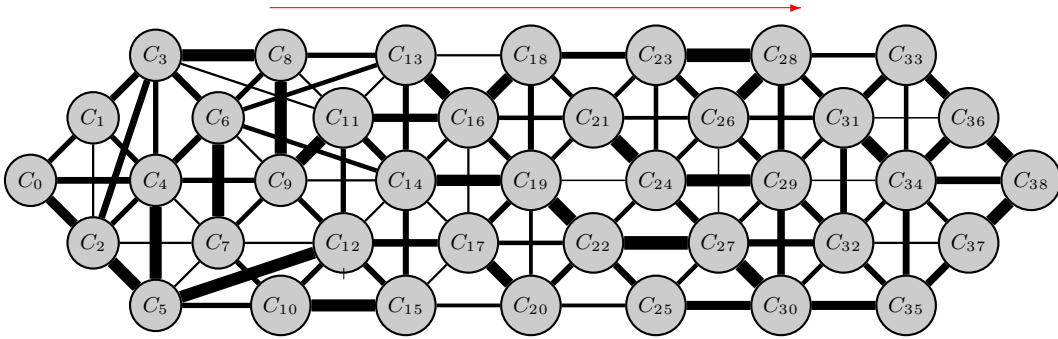


Figure 3: Simulation of 39 nodes with omnetpp

and how it can consider different parameters to take the re-identification decision.

Indeed, the visual confidence is generated using a random generator following a uniform distribution. This is the same for  $\tau_m$ ,  $\tau_r$  and the decision of the next destination. Some hypotheses are assumed:

- We suppose that the target is moving with a constant speed, and also we suppose that we have the same category of targets (pedestrians, cars...). Actually, this case will be added as future work. It will be based on a simple classification of targets, which will add one dimension on our tensor. The transition will depend in this case on the path and on the type of the target.
- If the camera notes that the prediction does not correspond to the detected target, it will suppose that it is a new target appearing in the network.
- If a camera does not receive any feedback, it will consider that it is located in the final destination of the target.

Initially, all the probabilities are fixed to  $\frac{1}{N}$ , where  $N$  is the number of nodes on the network.

A large network is implemented to show how a real network used with a lot of cameras can be implemented in the real world (Fig. 3). The different lines between the nodes represent the possible transitions, and their thickness varies depending on the importance of interactions between cameras. Nevertheless, the interaction between two cameras depend on the path followed before coming there. Hence, we choose to represent just one of them as it is not possible to present all of them. In addition, those lines do not represent the physical communication. In other words, two nodes can communicate due to the technology used, but will not be necessarily a transitional link in case they are not a target destination.

As it is mentioned in the previous section, we consider that the camera has no knowledge about the environment. Once it detects a target, it will broadcast the information to the cameras in its range. The number of cameras to be taken into account in the construction of the path is fixed at three. In this case, we have a system represented by the tensor  $(39*39*39*39*39*1)$ . Accordingly, we show the results on the node  $C_{21}$ . Figures 4a, 4b and 4c present the transition probabilities between  $C_{21}$  and  $C_{22}$ ; it is estimated based on the model presented and the values randomly generated by the simulator. We clearly notice that the probability depends on the followed path. Although we can not present

all possible paths, the latter figures highlight how much it is important to consider a non-Markovian model; it yields much more precision compared to the Markovian model. In spite of the fact that the trajectories are randomly generated, the target movements are still deterministic. Typically, we see that the probabilities converge to a stable state after 100 events (represented by  $N$  in Fig. 4). However, it varies depending on the past followed path. Typically, with Markovian models, the probability of moving from  $C_{21}$  to  $C_{22}$  is the average of the different values presented here and is not reliable information about a network.

As expected, the network takes much more time to reach a stable state. With a lot of path possibilities and an important choice of destinations, the probability converges slowly.

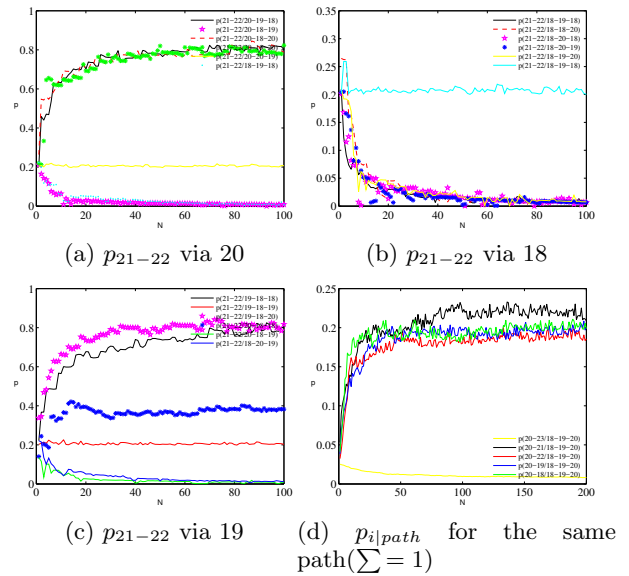


Figure 4: Probability in node  $C_{21}$

Besides the evolution of probabilities, we estimate the variation of the temporal pheromones  $\tau_m$  to be interesting (Fig. 5), as it represents the delay estimated to reach a camera. Considering that we have the same type of targets moving at a constant speed, the system is able to extract stable delay expectations. Unlike the probability, the evolution of temporal information does not depend on the followed path and is the same between two cameras. This temporal information can give an idea about the distance between the

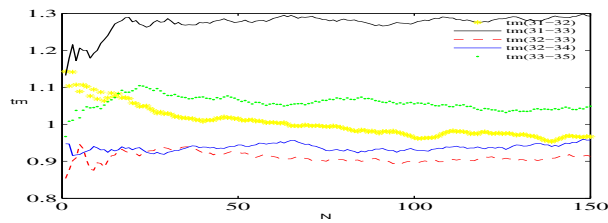


Figure 5: Evolution of the estimated time

cameras. This distance is relative and not physical. In this respect, we can deduce that the distance between 31 and 32 is one and a half more than the distance between 32 and 33.

## 6. CONCLUSION

Nowadays, researches give accurate results in terms of identification and tracking using single camera. However, once it comes to distributed cameras, other problems appear. It is not anymore about detection and classification, it is about sharing information and deciding whether it is the same or not. The main idea of the paper is that re-identification can not be limited to the visual comparison between two images captured from two cameras. Other parameters can be useful in this case, and can highly improve the results. Moreover, this model can fit very-low-specifications cameras. The system can reach a stable state despite the lack of information given to the cameras at the beginning. Once reached, it will be able to detect anomalies and find useful information such as the distance between cameras.

## 7. ACKNOWLEDGMENTS

This research has been financed by the the European Union under the European Regional Development Fund (ERDF).

## 8. REFERENCES

- [1] L. Ben Khelifa, L. Maggiani, J. C. Quinton, and F. Berry. Ant-Cams Network: a cooperative network model for silly cameras. In *International Conference on Distributed Smart Cameras 2016*, 2016.
- [2] C.-t. Chu and J.-n. Hwang. Fully Unsupervised Learning of Camera Link Models for Tracking Humans Across Nonoverlapping Cameras. 24(6):979–994, 2014.
- [3] C.-t. Chu, J.-n. Hwang, J.-y. Yu, and K.-z. Lee. Tracking Across Nonoverlapping Cameras Based On The Unsupervised Learning Of Camera Link Models.
- [4] C. E. Deptt. Optimal Sensor Placement for Surveillance of Large Spaces. 2009.
- [5] L. Esterle, P. R. Lewis, X. I. N. Yao, and B. Rinner. Socio-Economic Vision Graph Generation and handoff in Distributed Smart Camera Networks. *ACM Transactions on Sensor Networks*, 0(3):1–24, 2014.
- [6] S. Juan C, M. Christian, S. Karen, F. Gian Luca, and C. Andrea. Self-reconfigurable smart camera networks. pages 67–73. Computer, Vol.47, no.5, May 2014.
- [7] P. R. Lewis, L. Esterle, A. Chandra, B. Rinner, and X. Yao. Learning to be Different : Heterogeneity and Efficiency in Distributed Smart Camera Networks.
- [8] P. R. Lewis, L. Esterle, A. Chandra, B. Rinner, and X. Yao. Learning to be Different : Heterogeneity and Efficiency in Distributed Smart Camera Networks. 2013.
- [9] L. Maggiani, L. Ben Khelifa, J. C. Quinton, M. Petracca, P. Pagano, and F. Berry. Distributed coordination model for smart sensing applications. In *International Conference on Distributed Smart Cameras 2016*, 2016.
- [10] N. Martinel, S. Member, G. L. Foresti, S. Member, and C. Micheloni. Person Reidentification in a Distributed Camera Network Framework. pages 1–12, 2016.
- [11] N. Martinel, S. Member, C. Micheloni, C. Piciarelli, G. L. Foresti, and S. Member. Camera Selection for Adaptive Human – Computer Interface. 44(5):653–664, 2014.
- [12] N. Martinel and C. Micheloni. Sparse Matching of Random Patches for Person Re-Identification. 2014.
- [13] E. Mohamed, D. Francis, P. Wilfried, and A. Hamid. Sleep Analysis for Elderly Care Using a Low-Resolution Visual Sensor Network. In *6th International Workshop, HBU 2015*, Osaka, Japan, Sept. 2015. Springer International Publishing.
- [14] Y. Morsly, N. Aouf, M. S. Djouadi, and M. Richardson. Particle swarm optimization inspired probability algorithm for optimal camera network placement. *IEEE Sensors Journal*, 12(5):1402–1412, 2012.
- [15] C. Piciarelli, L. Esterle, A. Khan, B. Rinner, and G. L. Foresti. Dynamic Reconfiguration in Camera Networks: A Short Survey. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016.
- [16] J. V. C. I. R., J. García, N. Martinel, A. Gardel, I. Bravo, G. Luca, and C. Micheloni. Modeling feature distances by orientation driven classifiers for person. 38:115–129, 2016.
- [17] P. Re-identification, C. Micheloni, and C. Piciarelli. Distributed Signature Fusion for.
- [18] J. R. Riehl, G. E. Collins, and J. P. Hespanha. Cooperative search by UAV teams: A model predictive approach using dynamic graphs. *IEEE Transactions on Aerospace and Electronic Systems*, 47(4):2637–2656, 2011.
- [19] B. Song, C. Soto, A. K. Roy-Chowdhury, and J. a. Farrell. Decentralized camera network control using game theory. *Second ACM/IEEE International Conference on Distributed Smart Cameras*, pages 1–8, 2008.
- [20] A. Varga and H. Rudolf. An overview of the omnet++ simulation environment. 2008.
- [21] R. Vezzani, D. Baltieri, and R. Cucchiara. People Reidentification in Surveillance and Forensics : A Survey. 46(2), 2013.
- [22] C. Yu, G. Sharma, and R. Ny. Sensor Scheduling For Lifetime Maximization in User-Centric Image Sensor Networks. pages 1–12.