# Consensus in Visual Sensor Networks Consisting of Calibrated and Uncalibrated Cameras

Melanie Schranz and Bernhard Rinner
Institute of Networked and Embedded Systems
Alpen-Adria Universität Klagenfurt, Austria
{firstname}.{lastname}@aau.at

***Abstract*—Fusing multiple views of different cameras in a visual sensor network (VSN) require the calculation of a consensus. As a result we are able to overcome the uncertainties from the tracking algorithm and physical constraints in the observed environment. This paper addresses the topic of reaching a consensus within VSN of calibrated or uncalibrated cameras. In contrast to several proposed techniques, where the consensus is calculated on *all* cameras simultaneously, our approach calculates the consensus on a single camera. Furthermore, the responsibility for this calculation is migrated to the camera actually detecting the object with a better subjective view than its neighbors. This saves resources and reduces the communication overhead within the network.**

## I. Introduction

A main task of VSNs is the surveillance of large areas to identify suspicious changes in the environment or to classify human behavior. Typically, these networks consist of a combination of cameras with overlapping and non-overlapping field of views (FOVs) as well as observable and non-observable areas of the contributing cameras. To perform a multi-view tracking, VSNs require a collaborative way to agree upon the different outputs of each camera. Therefore, the cameras act as autonomous agents, analyze the retrieved data locally and exchange only processed data. This data is necessary to form a joint decision and finally, reach a global state, a consensus for all participating cameras in a VSN as done in [1], [2], [3], [4]. Reaching a consensus is defined as reaching an agreement regarding certain features. These features are retrieved as results of the tracking algorithm. They could represent appearance (color, size) or spatio-temporal (time, location) observations.

Tracking algorithms offer the information needed by the consensus algorithm. Depending on different parameters such as distance, orientation or occlusions of the tracked object they have a different performance. However, this is not reflected in state-of-the-art consensus algorithms which assume perfect tracking and do not cope with errors. Moreover, the neighborhood is assumed to be known [1].

Contrary to state-of-the-art consensus algorithms, we do not employ all available cameras in the network, but rather transfer the algorithm from one camera to another dependent on the tracking result. The camera to perform the algorithm is selected by comparing the utility of its tracking result with the others in the neighborhood. A similar approach to migrate the leadership is presented in [5].

The remainder of this paper is organized as follows: Section II briefly describes the problem covered in this paper and specifies the considered VSN. Section III explains the main parts of the consensus algorithm. Furthermore, it describes how to enhance this algorithm in a VSN of calibrated (III-A) and uncalibrated cameras (III-B). Finally, Section IV summarizes the paper and gives an outlook on future work.

## II. Problem Formulation

In this paper we consider a VSN with a fixed set of calibrated or uncalibrated cameras $C$ partially with overlapping FOVs. In addition, full observability is assumed. Their task is to monitor the environment and to identify a set of objects $O$ moving in their FOV. The communication between these cameras is assumed to be lossless. The VSN performs a distributed tracking algorithm, meaning, each camera is able to detect and track dedicated objects within its FOV. Furthermore, it is assumed that the tracker classifies the objects and assigns each tracked object an utility $\alpha$ of the tracking result for each camera, as done in [6]. Furthermore, the VSN reaches a consensus on position and velocity in a network of calibrated cameras and a consensus on classification in a network of uncalibrated cameras. To reduce the communication overhead, we do not calculate the consensus on each camera individually. In each step, a single camera is selected automatically based on the utility of the tracking result. This camera collects all necessary data and calculates the consensus. The main advantage of migrating the consensus algorithm to a single camera is a decrease of communication overhead. Compared to state-of-the-art consensus algorithms, where data is exchanged to *all* participating cameras, our approach employs only a single camera to calculate the consensus. Furthermore, decreasing the overall communication overhead can be realized by exchanging messages only within the neighborhood. In a calibrated network the neighborhood can be derived easily; in an uncalibrated network we identify the neighborhood using a vision graph as proposed in [6].

## III. Reaching a Consensus

Generally, a consensus describes an agreement over all the states $x_i \in R$ of the cameras yielding to a single, in this case, one-dimensional value

$$x_1 = x_2 = \ldots = x_{consenus}. \quad (1)$$

The consensus calculation for the proposed VSN uses the Kalman Consensus Algorithm introduced in [1]. For its computation it requires additionally an information vector $u$, an information matrix $U$ and the estimated state of the tracked object $\overline{x}$. In the calibrated as well as in the uncalibrated network the utility $\alpha_j$ is calculated for each time step on each camera for a certain object as proposed by [6]. This utility is used by the consensus algorithm to weight the tracking result. Therefore, a message $m_j$ transmitted from each camera to the camera currently holding the consensus responsibility needs to have following content

$$m_j = (u_j, U_j, \overline{x}_j, \alpha_j). \tag{2}$$

As in [1] the sensor data is fused for further consensus calculation. In our approach the fusion is extended by weighting the information matrix with the utility $\alpha_j$.

### A. Consensus in a VSN of calibrated cameras

In a VSN of calibrated cameras the neighborhood can be easily determined. In addition to tracking objects, the calibrated camera of a VSN is able to map the object's image plane position to the ground plane. Therefore, we can reach a consensus, e.g., on the object's ground plane position.

The cameras in the neighborhood transmit position together with utility to the camera having the consensus responsibility as in Equation 2. The current state describes the position $\overline{x}_j = (x_j, y_j)$ and could be further enhanced by additionally transmitting the velocity of the object $v_j$.

The camera responsible for the consensus calculation has then all the necessary data to reach a consensus for the location of the object on the ground plane of all cameras in its neighborhood. These steps are summarized in Algorithm 1.

### B. Consensus in a VSN of uncalibrated cameras

In a VSN of uncalibrated cameras the neighborhood is usually not known. To overcome this problem, we use the approach of a vision graph as proposed by [6], [7]. The control strategy as in [6] manages tracking responsibilities and learns the neighborhood relations. Instead of the tracking responsibility, our approach transfers the responsibility to calculate the consensus to the camera having the best utility of the tracked object.

For a VSN of uncalibrated cameras we cannot reach a consensus on the ground plane position of the object. When transmitting the message in Equation 2, we could use for the state $\overline{x}_j$ a parameter, e.g., indicating the class of an object. This parameter could be used to calculate a consensus on the object classification. As for the calibrated network, this camera has then all necessary data for calculating a consensus on the tracked object.

Furthermore, compared to [1], [2], [4], where the neighborhood is assumed to be known, our approach determines the neighborhood along with the consensus calculation using the utility of the tracking result. As described in [6], a link weighted according to the utility is created between the camera calculated the consensus and the camera having the best utility of the object in the next step. We change the proposed unidirectional link to a bidirectional link. Additionally, links should be created to cameras, where the utility was not high enough to win, but is still satisfactory for our application. The strength of the links follows the pheromone evaporation rule and is dependent on the network and the tracked object's characteristics as described in [6].

Since the neighborhood is calculated out of the data already used for the consensus calculation, there is no additional communication necessary.

**Algorithm 1** Reaching a consensus and migrating its responsibility in a VSN

Camera $c_i \in C$ owns the object $o_k$
**FOR** each object $o_k$

1) *Send auction initiation to neighbors* $c_j \in C$
2) *Receive data from all* $c_j$*, where* $\alpha_j > 0$ *as in Eq. 2*
3) *Weight tracking results*
4) *Calculate consensus*
5) *Migrate responsibility of calculating the consensus from* $c_i$ *to camera with best* $\alpha_j$

## IV. Summary and Outlook

In this paper we presented a new approach for calculating the consensus in a VSN with calibrated and uncalibrated cameras. A future step is to validate the proposed algorithm with simulations and further, with a physical test bed.

The addressed VSN is able to fully observe the area of interest. A further step is to extend the transferring behavior of the algorithm to VSNs that also deal with non-observable areas, so called gaps.

## References


[1] B. Song, C. Ding, A. Kamal, J. Farrell, and A. Roy-Chowdhury, "Distributed camera networks," *Signal Processing Magazine, IEEE*, vol. 28, no. 3, pp. 20 –31, May 2011.

[2] A. Jorstad, P. Burlina, I.-J. Wang, D. Lucarelli, and D. DeMenthon, "Model-based pose estimation by consensus," in *Proceedings of the International Conference onIntelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, Dec. 2008, pp. 569 –574.

[3] R. Tron, A. Terzis, and R. Vidal, "Distributed consensus algorithms for image-based localization in camera sensor networks," in *Distributed Video Sensor Networks*, B. Bhanu, C. V. Ravishankar, A. K. Roy-Chowdhury, H. Aghajan, and D. Terzopoulos, Eds. Springer London, 2011, pp. 289–302.

[4] E. Kokiopoulou and P. Frossard, "Distributed classification of multiple observation sets by consensus," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 104 –114, Jan. 2011.

[5] F. Qureshi and D. Terzopoulos, "Distributed coalition formation in visual sensor networks: A virtual vision approach," in *Distributed Computing in Sensor Systems*, ser. Lecture Notes in Computer Science, J. Aspnes, C. Scheideler, A. Arora, and S. Madden, Eds. Springer Berlin Heidelberg, 2007, vol. 4549, pp. 1–20.

[6] L. Esterle, P. Lewis, M. Bogdanski, B. Rinner, and X. Yao, "A socio-economic approach to online vision graph generation and handover in distributed smart camera networks," in *Proceedings of the ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC)*, Aug. 2011, pp. 1 –6.

[7] H. Kim, J. Romberg, and W. Wolf, "Multi-camera tracking on a graph using markov chain monte carlo," in *Proceedings of the ACM/IEEE International Conference on Distributed Smart Cameras*, Sept. 2009, pp. 1 –8.