Distributed Agreement Based ML Approximation

Mohamad Mohamadi¹, Hamid Parvin*¹², Es’hagh Faraji¹², Sajad Parvin¹

¹) Department of Computer Engineering, Nourabad Mamasani Branch, Islamic Azad University, Nourabad Mamasani, Iran
²) Young Researchers and Elite Club, Nourabad Mamasani Branch, Islamic Azad University, Nourabad Mamasani, Iran

mohammadi.oty@gmail.com; parvinhamid@gmail.com; eshagh1363@gmail.com; sajad.parvin@gmail.com

Abstract

In this manuscript we suggest a fast adaptive distributed method for maximum likelihood approximation (MLA) in multiple view object localization problem. For this purpose, we use “up to scale” property of projective geometry and by defining coefficients for convergence criterion, we increase the convergence speed of the consensus algorithm. We try to present a mathematical model for the problem. We use two types of error function. The proposed method uses maximum likelihood for obtaining its best parameters. Our approach utilizes “up to scale” property in projective geometry to reach the consensus quickly. The difference between nodes’ values and meanwhile consensus values are evaluated by two error functions. To estimate consensus value in the second error function, we used local weighted average of each node. At the last of the paper, we prove our claims by experimental results.

Keywords: Maximum Likelihood Approximation, Data Fusion, Consensus Algorithm, Homography

1. Introduction

Increasing interest in wireless sensor networks, the use of the multiple-view structure in the object tracking applications rises [11, 7, 3, and 22]. Projective geometry is used as a mathematical tool for the multiple-view tracking systems. Data of a scene from all the views should be illustrated in the same coordination, namely global coordination. Relations between cameras' coordination and any arbitrary global coordination are described by homography which is a popular nonlinear transformation (in Euclidean geometry) in multi-view schemes. The use of homography implies that the cameras’ data are represented in the projective geometry rather than Euclidean geometry (or any other geometry) [5]. The most important property of the projective geometry is their ability to preserve data insensitive to scale. Though the relation between each camera coordination and global coordination in Euclidean geometry is nonlinear, this relation in the projective geometry has a linear form. With this linear model maximum likelihood approximation is a good choice to approximate the objects' coordination in multi-view tracking systems.

Distributed data fusion is another important issue in the sensor networks applications [21, 14]. It increases the system’s ability in dealing with the failure in any part of the network. Consensus algorithm is an efficient solution in distributed approaches in many
applications such as beam forming [23], spectral sensing in cognitive radio [24, 1, 13, and 18], target tracking [25, 9, and 19] and adaptive filters [20]. Because of the improvements and the extension ideas in sensor and multi-agent networks, consensus algorithm is also used in data fusion [21, 4, and 12] in the recent years. In this procedure each node communicates only with its neighbors and after several iterations, nodes reach the consensus in the whole network. Actually, this consensus value (values) can play the role of an auxiliary variable or a cost function in the network and help(s) us reaching the final purpose [10]. Consensus algorithm is organized in Euclidean geometry and all iterations and its convergence criterion are described in this geometry [17, 15]. In [16] a new criterion for convergence in the Riemannian manifold is introduced. Though the projective geometry is a special type of Grassmann geometry [2] and Grassmann geometry itself is a kind of Riemannian geometry. In this work we don’t use Riemannian manifold criterion for consensus because the convergence of Riemannian consensus criterion isn’t guaranteed in every situations. Instead, we propose a new convergence criterion in the projective geometry.

In this paper, we propose a consensus based MLA for the object localization in the projective geometry. For this purpose, we apply a modification in the Euclidean consensus criteria and use this new criterion for consensus in the projective geometry. This paper is organized as follows. In the next section, a short explanation of the consensus procedure comes. Then the description of the distributed MLA algorithm which is presented in [21] is developed. The problem statement and our proposed procedure are introduced in section 4. Finally, in section 5, with the numerical results we show our proposed scheme performance with respect to the traditional consensus algorithms. Then the paper is terminated by a conclusion.

2. Consensus Procedure

Consider a connected sensor network modeled by a digraph \( G = (V, E, A) \) with \( N \) edges, where \( V = \{v_1, v_2, \ldots, v_N\} \) is the set of nodes, \( E \subseteq V \times V \) is the set of edges so that \((v_i, v_j) \in E\) if there is an edge between the \( i^{th} \) and the \( j^{th} \) nodes and \( A \) is the digraph adjacency matrix. Let \( x_i(k) \) as the vector value of the \( i^{th} \) node at the moment \( k \). If all sensors’ values are represented as \( X(k) = [x_1(k), \ldots, x_N(k)] \), consensus algorithm will prepare as iterative procedure forcing all sensors’ values tend to an consensus value (vector in here), named consensus value [10]. The maximum, the minimum consensus and the weighted average consensus are several usual consensus values which are used in many applications [8]. Different iterative procedures could be used, depending on the definition of the consensus value. The algorithm’s policy to reach the final consensus value is named the consensus protocol [8]. However, the computation of the weighted average is the most important consensus algorithm. By the above definition about the network topology and nodes’ value, the weighted average consensus is defined as below:

\[
\lim_{k \to +\infty} x_i(k) = w^T X(0)
\]  

(1)
Where \( \mathbf{w} \) is a \( N \times 1 \) vector indicates the weight of each sensor in obtaining process of the final consensus vector. Equation (1) indicates that all nodes reach to the identity vector in convergence. The consensus protocol in this case is:

\[
\mathbf{X}(k) = \mathbf{D}\mathbf{X}(k-1) = \mathbf{D}^k \mathbf{X}(0)
\]  

(2)

Where \( \mathbf{D} \) is an \( N \times N \) matrix named consensus matrix. The component \( d_{ij} \) of \( \mathbf{D} \) is equal to zero if \((v_i, v_j) \notin E\). To reach the consensus, the consensus matrix and its corresponding graph must have the specific properties. This matrix must be nonnegative matrix (have no negative entry) and right stochastic. Also, the corresponding graph on which the consensus matrix is defined must be connected. Nonnegativity and being right stochastic of \( \mathbf{D} \) guarantee that all eigen values of \( \mathbf{D} \) rely inside unit circle based on Gershgorin’s Disc theorem [6]. Moreover, the second property implies that 1 will be an eigenvalue of \( \mathbf{D} \) and its associated eigenvector is such that all its entries are one. The third property satisfies that this eigenvalue is first order geometry [10]. In [21, 8, 20] several methods for construction of the consensus matrix are represented. In all of the mentioned methods in [21, 8, 20] network’s digraph must be a connected graph. If \( \mathbf{D} \) is a double stochastic matrix, the weighted average in (1) will tend to the geometric average.

**Theorem 1** The consensus protocol in (2) induces to the weighted average consensus defined in (1) [10].

**Proof:** By substituting singular value decomposition of \( \mathbf{D} \):

\[
\mathbf{D} = \mathbf{V}\mathbf{L}\mathbf{V}^{-1}
\]

Where \( \mathbf{V} \) is a \( N \times N \) unitary matrix, its columns are eigenvectors of the consensus matrix and \( \mathbf{L} \) is a diagonal matrix with eigenvalues of \( \mathbf{D} \) corresponding to \( \mathbf{V} \) matrix columns as its entries, we can rewrite (2) as:

\[
\mathbf{X}(k) = \mathbf{D}^k \mathbf{X}(0) = \mathbf{V}\mathbf{L}^k \mathbf{V}^{-1}\mathbf{X}(0)
\]

As all eigenvalues of \( \mathbf{D} \) rely inside unit circle and 1 is the first order eigenvalue, all the diagonal entries of \( \mathbf{L} \) are tended to zero except one. Without loss of generality, assume \( \mathbf{v}_N \) is the eigenvector corresponding to 1 as its eigenvalue, so:

\[
\mathbf{D}^k = \mathbf{v}_N^T \mathbf{u}_N^T = \mathbf{1}_N \mathbf{u}_N^T = \begin{bmatrix} \mathbf{u}_N & \ldots & \mathbf{u}_N \end{bmatrix}^T
\]

Where \( \mathbf{1}_N \) is a vector with all entries equal by one and \( \mathbf{u}_N^T \) is the \( N^{th} \) row of \( \mathbf{V}^{-1} \). Thus, the weight vector in (1) is equal with \( \mathbf{u}_N^T \).

3. **Distributed MLA in Sensor Networks**

With the same assumption for the network in the previous section, suppose that the \( i^{th} \) node measures a vector with dimension \( m_i \) for unknown parameter \( \theta \in \mathbb{R}^m \) as:
\[ y_i = A_i \theta + \zeta_i \]  

where \( A_i \in \mathbb{R}^{m \times m} \) is a known matrix and \( \zeta_i \) is the zero mean white Gaussian noise vector with \( \mathcal{S}_i \) as its covariance matrix. The maximum likelihood estimation of \( \theta \) is obtained by:

\[ \hat{\theta}_{ML} = S^{-1}q \]  

where:

\[ S = \sum_{i=1}^{N} A_i^T \mathcal{S}_i^{-1} A_i \]  

\[ q = \sum_{i=1}^{N} A_i^T \mathcal{S}_i^{-1} y_i \]  

By defining \( \mathcal{S}_i (0) = A_i^T \mathcal{S}_i^{-1} A_i \) and \( q_i (0) = A_i^T \mathcal{S}_i^{-1} y_i \), distributed consensus based solution for MLA is obtained by these iterative equations [21]:

\[ S_i (k) = \sum_{j=1}^{N} d_{ij} S_j (k - 1) \]  

\[ q_i (k) = \sum_{j=1}^{N} d_{ij} q_j (k - 1) \]  

\[ \hat{\theta}_{ML,i} (k) = S_i^{-1} (k) q_i (k) \]  

where \( \hat{\theta}_{ML,i} \) is the MLA of the \( i^{th} \) node and \( d_{ij} \) are entries of matrix \( D \) as the consensus matrix. Considering the results in the previous section:

\[ \lim_{k \to +\infty} S_i (k) = \frac{1}{N} \sum_{i=1}^{N} A_i^T \mathcal{S}_i^{-1} A_i = \frac{1}{N} S \]  

\[ \lim_{k \to +\infty} q_i (k) = \frac{1}{N} \sum_{i=1}^{N} A_i^T \mathcal{S}_i^{-1} y_i = \frac{1}{N} q \]  

\[ \lim_{k \to +\infty} \hat{\theta}_{ML,i} (k) = \hat{\theta}_{ML} \]  

So, implementation of two consensus algorithm for obtaining \( S \) and \( q \) are induced to distributed solution for the MLA. In the next section we provide a modification in this procedure in order to increase the speed of convergence.

4. Object Localization

A. Problem Statement
Consider a network digraph mentioned in section 2 which monitors a plane for object detection. By assuming an arbitrary object on the plane as \( \mathbf{p} = (x, y, z) \in \mathbb{P}^2 \) in homogeneous coordination (equal to \( (x / z, y / z) \in \mathbb{R}^2 \)), each camera registers this object as \( \tilde{\mathbf{p}}_i = (\tilde{x}_i, \tilde{y}_i, \tilde{z}_i) \in \mathbb{P}^2 \), \( i \in \mathcal{V} \) in its local coordination. Projective geometry \( \mathbb{P}^N \) is equivalent to \( \mathbb{G}_{N,1}^2 \), so the object location can be represented on the Riemannian Geometry. If \( \mathbf{H}_i \in \mathbb{R}^{3 \times 3} \) denotes a homography between reference and the \( i^{th} \) camera coordination \( ^i \mathbb{P} \), then relation between the object coordinate in each camera \( (\tilde{\mathbf{p}}_i) \) and the reference coordinate \( (\mathbf{p}_i) \) is described as \( \mathbf{p}_i = \mathbf{H}_i \tilde{\mathbf{p}}_i \) [5]. Based on the previous section results, Algorithm 1 shows the distributed MLA for object in the plane and obtains \( \mathbf{p}_{ML} \), where \( \varepsilon > 0 \) is an arbitrary positive constant, \( d_{ij} \) are the entry of consensus matrix, \( \mathbf{p}_{ML,i} \) is the MLA of object position in the \( i^{th} \) node and \( \mathbf{M}(k) \) is equal to:

\[
\lim_{k \to +\infty} \mathbf{S}_i(k) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{A}_i^{T} S_i^{-1} A_i = \frac{1}{N} \mathbf{S} \quad (13)
\]

**Algorithm 1** Distributed MLE for Object Localization

1: for \( i = 1 \rightarrow N \) do
2: \( \mathbf{S}_i(0) = \mathbf{H}_i^T \Sigma_i^{-1} \mathbf{H}_i \)
3: \( \mathbf{q}_i(0) = \mathbf{H}_i^T \Sigma_i^{-1} \tilde{\mathbf{p}}_i(0) \)
4: end for
5: while \( \| \mathbf{M}^T(k) - \mathbf{M}^T(k-1) \|^2 > \varepsilon \) do
6: for \( i = 1 \rightarrow N \) do
7: \( \mathbf{S}_i(t) = \sum_{j=1}^{N} d_{ij} \mathbf{S}_j(t-1) \)
8: \( \mathbf{q}_i(t) = \sum_{j=1}^{N} d_{ij} \mathbf{q}_j(t-1) \)
9: \( \mathbf{p}_{ML,i}(t) = \mathbf{S}_i^{-1}(t) \mathbf{q}_i(t) \)
10: end for
11: end while

**B. The proposed Approach**

In this subsection, we introduce an approach to increase the speed of the consensus convergence. In our proposed procedure, we don’t use Riemannian consensus because Riemannian consensus is converged to Frechet mean and its convergency is not guaranteed [16]. Instead, we propose a modified Euclidian consensus on Projective geometry. Our algorithm uses this property of homogenous geometry that \( (x, y, z) \) and \( (ax, \alpha y, \alpha z) \), where \( \alpha \) is a nonzero coefficient, show the same point in \( \mathbb{R}^2 \) geometry (equal to \( (x / z, y / z) \)). As mentioned in the introduction, the consensus algorithm and convergence to consensus in the network are based in the Euclidean geometry, while our problem is stated in the projective geometry. Therefore, because of "up to scale" property of projective geometry, there are infinity options to represent object location in

---

1. Projective Geometry
2. Grassmann Geometry
3. It depends on calibration matrix and the position of each camera [5]
the projective geometry. For example, suppose the final object location in the plane is \((x, y, z)\). In this situation there is no difference whether the consensus algorithm reaches to \((x, y, z)\) or \((ax, ay, az)\), as both of these points show the same point in the Euclidean geometry.

Based on the above mentioned point, let an object localization problem with \(p_c\) be as the final consensus point. In each iteration, the consensus algorithm induces to each node’s value \(p_i(k)\) tended to \(p_c\) \((p_i(k) \rightarrow p_c, i \in \{1, 2, ..., N\}\). In this work, as data represented in the homogenous coordinates, at each moment, tending \(p_i(k)\) to any multiplied value of \(p_c\) causes the same point in \(R^2\). Moreover, tending any multiplied value of \(p_i(k)\) to any multiplied value of \(p_c(k)\) has the same results. So, in our approach, the consensus procedure can be summarized in an attempt to establish a relationship \(\gamma_i(k)p_i(k) \rightarrow \gamma_i(k)p_c\), or \(\frac{\gamma_i}{\gamma_i(k)}p_i(k) \rightarrow p_c\). Based on what was mentioned in the above, we can compute coefficients \(\gamma_i(k)\) and \(\gamma_i(k), i \in \{1, 2, ..., N\}\), so that algorithm converges to the final values faster. For better imagination, consider a problem in \(P^1\) geometry by a set of two tuple vectors as \(p_i(k), i \in \{1, 2, 3\}\) initialized in the below:

\[
p_{1}(0) = \begin{bmatrix} 6 \\ 7 \end{bmatrix}, p_{2}(0) = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, p_{3}(0) = \begin{bmatrix} 4 \\ 13 \end{bmatrix}
\]

Let consensus matrix be:

\[
D = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}
\]

Regarding to this configuration, the final consensus vector equals to \(p_c = \begin{bmatrix} 4 & 8 \end{bmatrix}^T\). Figure 1 shows the iterative changes in \(p_i(k)\) for nodes 2 and 3 under the consensus algorithm. As shown in Figure 1 at first iteration \(0.8p_3(1) \approx p_c\) and in second iteration \(0.48p_2(2) \approx 0.42p_c\). In other words, \(p_3\) reaches the final consensus value after first iteration and \(p_2\) reaches the final consensus after two iterations. All of the lying points on solid line are equal in \(R\) geometry and only have a difference in an scale factor. More consideration on Figure 1 shows that initial value of \(p_2\) is equal to \(p_c\). So, \(p_2\) does not need to be changed in iterative consensus algorithm and it can be introduced as the final consensus value. Therefore, we can propose Algorithm 2 for object localization. The only difference between two algorithm is in the convergence condition. In the next subsection we introduce an approach for obtaining \(a\) and \(g\) values.
Fig. 1. Up-to scale property in $P^1$. $p_3$ reaches consensus, almost, at $k = 1$ by setting $\alpha_3(1) = 0.8$ and $\gamma_3(1) = 1$. Also $p_2$ gets consensus, approximately, at $k = 2$ by setting $\alpha_2(2) = 0.48$ and $\gamma_2(2) = 0.42$.

Algorithm 2 Fast Distributed MLE for Object Localization

1: for $i = 1 \rightarrow N$ do
2: $\alpha_i(0) = 1, \gamma_i(0) = 1$
3: $S_i(0) = H_i^T \Sigma_i^{-1} H_i$
4: $q_i(0) = H_i^T \Sigma_i^{-1} \hat{p}_i(0)$
5: end for
6: $\beta(0) = \left[ \frac{\alpha_1(0)}{\gamma_1(0)} \frac{\alpha_2(0)}{\gamma_2(0)} \cdots \frac{\alpha_N(0)}{\gamma_N(0)} \right]^T$
7: while $\|M^T(k)\beta(k) - M^T(k-1)\beta(k-1)\|^2 > \epsilon$ do
8: for $i = 1 \rightarrow N$ do
9: $S_i(k) = \sum_{j=1}^{N} d_{ij} S_j(k-1)$
10: $q_i(k) = \sum_{j=1}^{N} d_{ij} q_j(k-1)$
11: $p_{M,i}(k) = S_i^{-1}(k)q_i(k)$
12: Compute $\beta(k)$
13: end for
14: end while

C. Obtaining $\alpha, \gamma$ Values

In order to obtain $a, g$ values, two global error functions are considered as follows:

$$E(k) = \sum_{i,j=1}^{N} \left\| \frac{\alpha_i(k)}{\gamma_i(k)} p_i(k) - \frac{\alpha_j(k)}{\gamma_j(k)} p_j(k) \right\|^2$$  \hspace{1cm} (14)
\( \Delta(k) = \sum_{i=1}^{N} \left[ \frac{\alpha_i(k)}{\gamma_i(k)} p_i(k) - p_c \right]^2 \)  

(15)

\( \varphi(k) = E(k) + \Delta(k) \)  

(16)

\( E(k) \) satisfies the consensus between weighted nodes’ values and \( \Delta(k) \) induces convergence to the final consensus vector. The optimum proportion of \( \beta_i(k) = \alpha_i(k) / \gamma_i(k) \) is computed when \( \varphi(k) \) is minimized. Therefore, we obtain it by computing the following equations linearly:

\[
\frac{\partial \varphi(k)}{\partial \beta_i(k)} = 0 \quad i \in \{1, 2, \ldots, N\}
\]

(17)

However equations (14) and (15) cannot be calculated in a distributed way. Moreover, the final consensus values are interested and in each iteration, these values are unknown. Therefore we develop the following node’s error function.

\[
E_i(k) = \sum_{j \in N_i} \left[ \frac{\alpha_i(k)}{\gamma_i(k)} p_i(k) - \frac{\alpha_j(k)}{\gamma_j(k)} p_j(k) \right]^2
\]

(18)

\[
E_i(k) = \sum_{j \in N_i} \left[ \frac{\alpha_i(k)}{\gamma_i(k)} p_i(k) - \frac{\alpha_j(k)}{\gamma_j(k)} p_j(k) \right]^2
\]

(19)

\[
\Delta_i(k) = \left[ \frac{\alpha_i(k)}{\gamma_i(k)} p_i(k) - \hat{p}_{c,i}(k) \right]^2
\]

(20)

where \( N_i \) is the set of \( i^{th} \) node’s neighbors.

In (19) \( \hat{p}_{c,i}(k) \) is an approximation of \( p_c \) in the \( i^{th} \) node. An estimation value for \( \hat{p}_{c,i}(k) \) is demonstrated in the next section.

Now, we calculate \( \frac{\partial \varphi(k)}{\partial \beta_i(k)} = 0 \) to obtain proportion \( \alpha_i(k) / \gamma_i(k) \). After some calculations, we have:

\[
\frac{\alpha_i(k)}{\gamma_i(k)} = \frac{1}{|N_i|} \sum_{j \in N_i} \frac{\alpha_j(k)}{\gamma_j(k)} \frac{p_j^T(k) p_j(k) + p_i^T(k) \hat{p}_{c,i}}{2 p_i^T(k) \hat{p}_{c,i}}
\]

(21)

The first term in the numerator of (21) is induced by \( E_i \) and the second term is induced by \( \Delta_i \). By using (21) in each iteration we can obtain appropriate coefficients to reach the consensus. In the next section, simulation results of the convergence procedure would be shown.
5. Experimental Results

In this section, we present simulation results of our proposed approach. For this purpose, we construct a static random connected network. Also, we select maximum-degree weighted matrix [21] as the consensus matrix. We use random 3×3 matrices for simulating calibration matrix of cameras and add independent white Gaussian noise to the measurements. The results are obtained by averaging over 100 epoches by different random initial values. To approximate \( p_c \), different strategies such as local averaging of each node and local weighted averaging of each node (based on the consensus matrix’s rows values) can be used. Simulation results verified the performance of these choices. In this paper, we use the latter. So, the estimated value is obtained as below:

\[
\hat{p}_{c,j}(k) = \sum_{j=1}^{N} d_{ij} \hat{p}_j(k)
\]  

(22)
Fig. 4. Number of iterations versus different network sizes and \( \frac{\alpha}{\gamma} \) in the 19th node for \( q \)

Fig. 5. \( \frac{\alpha}{\gamma} \) in the 19th node for \( q \)

Obviously, when \( k \) increases, the above equation tends to \( p_c \). Figure. 2 and Figure. 4 show the number of iterations which needed to reach consensus for \( S \) and \( q \) in different sizes of camera networks. The ideal fast approach uses real value of the final consensus vector(\( p_c \)). Therefore, it has the best results, but it can’t be realized. As shown in the figures, our proposed method (fast distributed approach) has significant improvement with respect to the standard consensus algorithm. Finally in Figure. 3 and Figure. 5, the ratio of \( \alpha \) and \( \gamma \) in the 19th node for a 20 cameras network is shown. Our approach has a fluctuational behavior in the first iterations, as we used approximated function instead of \( p_c \); however, by passing the time and tending \( \hat{p}_{ci}(k) \) to \( p_c \), the distributed approach ratio is converged to the ideal one.
6. Conclusion

In this paper, we introduce a fast and distributed maximum likelihood estimation using the consensus algorithm. Our approach utilizes "up to scale" property in projective geometry to reach the consensus quickly. The difference between nodes’ values and meanwhile the difference between nodes’ values and consensus values are evaluated by two error functions. To estimate consensus value in the second error function, we used local weighted average of each node. Experimental results show that this estimation can improve the convergence speed.

References


