Map Fusion for Precise Yet Efficient Collaborative SLAM

Luigi Maciel Ribeiro[1], Nadia Nedjahr[2], Paulo Victor R. Carvalho[1]

[1] *Institute of Computing, Federal University of Rio de Janeiro (UFRJ)*
Av. Athos da Silveira Ramos, 274 - Cidade Universitári, Rio de Janeiro, 21941-916, RJ, Brazil
luigimaciel@dcc.ufrj.br, paulov@ien.gov.br
[2] *Faculty of Engineering, State University of Rio de Janeiro (UERJ)*
R. São Francisco Xavier, 524 - 5° andar - Maracanã, Rio de Janeiro, 20550-011, RJ, Brazil
nadia@eng.uerj.br

Abstract. Collaborative Simultaneous Localization and Mapping (C-SLAM) is an active research area in robotics that aims to enable the collaboration of multiple robots in constructing a shared map and simultaneously estimating their positions. However, Map fusion poses a significant challenge, especially when involving a large group of robots. It aims at obtaining an accurate global representation of the environment. This paper proposes an novel approach using the Fourier Transform, the Pearson correlation coefficient and Particle Swam Optimization to address the map fusion problem. Efficiently merging maps into a global representation requires careful consideration of spatial relationships and alignment of these maps. The Fourier Transform analyzes spectral features in each robot's measurements, extracting insights about spatial distribution. The Pearson correlation coefficient evaluates spectral similarity between different map sections, facilitating region pairing for successful fusion. The search for optimal fusion parameters is performed using the Particle Swarm Optimization Algorithm. These distinct regions guide the fusion process, optimizing global map creation. Instead of a complete map fusion, selective fusion of sections increases the likelihood of success. Experiments involving five robots in a simulated environment validate the proposed approach, demonstrating the capability of optimized map fusion to provide a more accurate and comprehensive representation of the environment. This enhancement should contribute to refine further the C-SLAM.

Keywords: C-SLAM, Map fusion, Fourier Transform, Pearson Correlation, Multi-robots

1 Introduction

Collaborative Simultaneous Localization and Mapping (C-SLAM) emerges as a prominent research area in contemporary robotics, aiming to enable cooperation among multiple robots in constructing shared maps and simultaneously estimating their positions, Dörr et al. [1], Saeedi et al. [2]. Within this context, the fusion of individual maps to create a precise global representation of the environment stands out as a crucial challenge. This fusion not only facilitates collaboration among robots but also offers substantial advantages in scenarios involving complex operations and dynamic environments.

The foundational premise of C-SLAM rests on data and resource sharing, with the central goal of effectively constructing a comprehensive map of the environment. In environments challenging for a single robot to explore, such as industrial complexes, the applicability of C-SLAM becomes pertinent, simplifying mapping processes in demanding contexts.

The fusion of maps, an essential component of C-SLAM, involves integrating individual maps generated by each robot into a cohesive global map, Ma et al. [3]. This procedure is indispensable to promote a meaningful collaboration between the robots operating jointly in a shared space, as observed in search and rescue operations or complex industrial settings. However, the task of map fusion encounters technical hurdles. The initiation process of C-SLAM often takes place without well-defined initial references, resulting in maps that, while seemingly coherent, may exhibit translation and rotation variations due to undefined references. Furthermore, many real-world environments feature repetitive structures, which can introduce additional challenges to map fusion, leading to misalignments and ambiguities.

This paper proposes an innovative approach as a response to these challenges. The use of the Discrete Fourier Transform assumes a fundamental role, enabling spectral analysis of measurements obtained by the robots. This analysis provides an in-depth understanding of the mapped environment, addressing issues such as misalignments and local minima that impact the accuracy of map fusion. At the heart of this contribution is overcoming fundamental obstacles in map fusion, with direct implications for the effectiveness of robotic collaboration and obtaining reliable and accurate global maps. This approach not only offers a technical solution but also outlines a path for future investigations and practical applications. In essence, this research presents a noteworthy advancement in the field of C-SLAM, carrying significant implications both within the realm of robotics and beyond.

The structure of this paper comprises five sections. Initially, in Section 2, we survey main recent works in the domain of map fusion. Subsequently, in Section 3, we elaborate on the proposed method for map fusion. Following that, in Section 4, we showcase the implementation results. Finally, in Section 5, we draw conclusions and outline avenues for future research.

2 Relates Works

In Marjovi et al. [4], the innovative concept of robotic clusters is introduced to enhance the resolution of complex problems. This concept proposes the formation of a robotic cluster, consisting of a individual robots group capable of sharing their processing resources. This allows robots to tackle challenging tasks by sharing their processing units. The paper explores the concept, requirements, characteristics, and architecture of robotic clusters in detail. Furthermore, the problem of 'topological map fusion' is addressed as a case study to illustrate the implementation and evaluate the functionality of the presented idea. The research also presents a novel parallel algorithm developed to address this specific problem. Experimental results validated that robotic clusters significantly accelerate computations in multi-robot systems. This proposed mechanism has potential applications in various areas of robotics and has the power to enhance the performance of multi-robot systems, especially when solving problems that require substantial processing resources.

In Hörner [5], a new algorithm for fusion 2D maps created by different robots without knowledge of initial relative positions is presented. The algorithm is inspired by computer vision techniques for creating panoramas from individual photos. The presented algorithm uses data represented by grid occupancy maps, allowing good scalability for heterogeneous swarms of multiple robots and enabling the use of the algorithm with different SLAM algorithms. The map fusion algorithm is implemented as a publicly available ROS package and is accepted for distribution by ROS. The algorithm's performance is tested in the ROS environment using the VREP simulator. For evaluation purposes, a ROS package is developed for autonomous exploration of environments in this work.

In Carpin [6], an innovative algorithm for fusing maps generated by multiple robots exploring the same environment is proposed. This algorithm generates a set of candidate transformations necessary for the fusion of two maps, comprising translations and rotations. Each transformation is weighted, allowing the distinction of uncertain situations and enabling the tracking of various scenarios in the face of ambiguities. The transformations are derived from the spectral analysis of information present in the maps. This approach stands out for its deterministic, non-iterative nature, and computational efficiency. The conducted experiments cover publicly accessible datasets, as well as maps produced by two robots exploring both indoor and outdoor environments. Through experimental validation, it is demonstrated that the proposed technique consistently achieves map fusion with notably distinct characteristics.

3 Proposed Method for Map Fusion

Map fusion is the process of combining information from different maps to create a unified and more comprehensive representation of an environment. In the context of Simultaneous Localization and Mapping (SLAM) solutions for multiple robots, map fusion is employed to merge the individual maps created by each robot into a single global map.

Map fusion enables the integration of information from multiple robots to create an accurate global representation of the environment. In applications involving multiple robots, map fusion aims to overcome the challenge of synchronizing individual maps, each constructed from an unknown initial reference. This results in maps with similar topological structures but different translations and rotations. The optimized fusion process occurs in steps, as illustrated in the flowchart presented in Fig. 1.

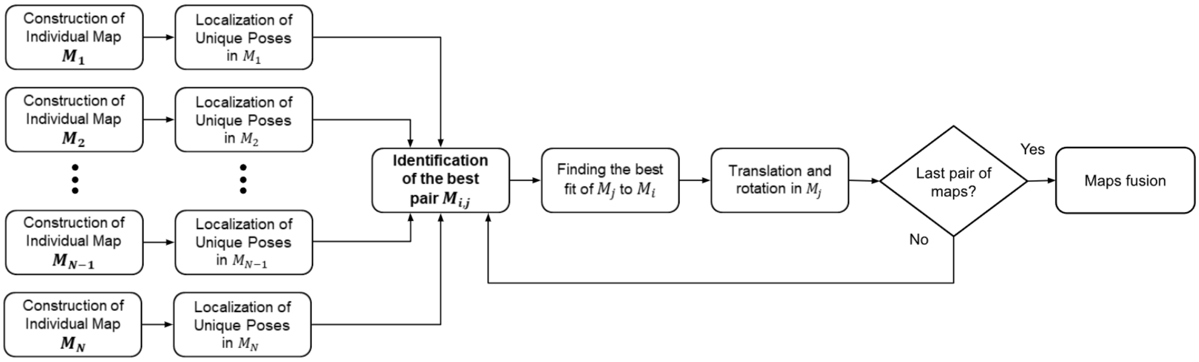


Figure 1. Proposed map fusion process steps

3.1 Construction of Individual Maps

A map is a matrix M with r rows and c columns, where $M(i, j)$ represents the state of the cell. Each cell can take one of three values: "free", "occupied", or "unknown". These values encode the environment's state in the corresponding cell and they are essential for creating an understanding of the surroundings.

The process of map generation involves collecting data from the robot's sensors, such as laser range sensors, depth cameras, or LiDAR. These sensors provide information about the distance and geometry of obstacles in the environment. By processing this data, the robot identifies obstacles and their positions, which are then integrated into the map representation. Contemporary SLAM algorithms utilize occupancy beliefs for each cell, which are converted to the "occupied" state in the grid map. As a result, the generated map becomes a discrete representation of the environment's features and obstacles.

3.2 Localization of Unique Pose

To identify unique regions, we use spectral analysis based on the Discrete Fourier Transform (DFT), as presented by Brigham [7]. Discrete Fourier Transform (DFT) aids in identifying distinct spectral features, enabling the detection of unique poses. This involves calculating the overall map's spectrum and comparing it with the spectra of pose sections.

DFT is a fundamental tool in signal and time series analysis, playing a crucial role in extracting frequency domain information from time domain data. The equation defining DFT is given as eq. (1):

$$DFT(u) = \sum_{x=0}^{N-1} f(x)e^{-i2\pi ux/N}, \tag{1}$$

where, $DFT(u)$ represents the magnitude of the transform for frequency u, N denotes the total number of points in the time series, and $f(x)$ is the sample value at time instant x. The complex exponential $e^{-i2\pi ux/N}$ introduces the frequency component, where i is the imaginary unit and 2π is a constant amplifying the angular change. The variable u controls the frequency of interest in spectral analysis, while x iterates through the time series samples.

In this step, unique poses present in each map are identified. To achieve this crucial goal of our approach, we apply a spectral analysis algorithm based on the DFT, as outlined in algorithm 1. Accurate detection of unique poses is of paramount importance as it significantly enhances the probability of successful map fusion. When examining the map constructed in the simulation, the presence of several similar structures becomes evident. This leads to a substantial impact of repetitive structures on spectral behavior when calculating the overall map's spectrum, considering all ξ poses. Therefore, to identify unique poses, which are those located in regions with atypical features. We divide the analysis into sections containing σ poses each. Then we calculate the spectrum of each section to determine the divergence from the overall spectrum. This is crucial, as the strong influence of typical and repetitive regions on the overall spectrum causes unique poses, with their distinct structures, to exhibit significant divergence from the average spectrum in the analysis sections. This spectral behavior in regions with unique poses stands out clearly in comparison to the overall spectrum.

Algorithm 1 Unique Pose Localization

Calculate the overall spectrum: $o\_spec = |dft(lidar\_r)|$ ▷ eq. (1)
Initialize parameters: $sub\_sec \leftarrow 5$, $u\_sec \leftarrow []$, $pose\_d \leftarrow []$, $L\_D \leftarrow []$
for i in range from sub_sec to $number\_poses - sub\_sec$ **do**
    $s\_spec \leftarrow |dft(lidar\_r[i - sub\_sec : i + sub\_sec, :])|$
    for $spec$ in s_spec **do**
        $dif \leftarrow \sum |spec - o\_spec|$, axis=1
        $L\_D.append(dif)$
    end for
    $pose\_d.append(L\_D)$
end for
    ▷ Identify unique poses based on discrepancies, where the threshold is adjusted to ensure identification of a standardized number of poses.
for $i, pose$ in $enumerate(pose\_d)$ **do**
    $has\_unique \leftarrow False$
    for dif in $pose$ **do**
        if $\sum(dif < \delta) \leq 1$ **then**
            $has\_unique \leftarrow True$
            break
        end if
    end for
    if has_unique **then**
        $u\_sec.append(i)$
    end if
end for
return u_sec

3.3 Identification of Best Pair of Maps

Spectral similarity between different sections of unique poses is evaluated using the Pearson Correlation Coefficient, as introduced by Cohen et al. [8]. The pair of maps with the highest number of poses exhibiting high similarity is chosen as the selected pair in this stage of the process for map fusion. The Pearson Correlation Coefficient, denoted as ρ_{xy}, is a statistical measure assessing the linear relationship between two variables x and y. In this formula, presented in eq. (2):

$$\rho_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{2}$$

where, n denotes the total number of observations in the sample. The variables x_i and y_i represent individual values of observations i in x and y, respectively. The means $\bar{x}$ and $\bar{y}$ denote the arithmetic means of observations in variables x and y, respectively. The calculation of the Pearson Correlation Coefficient involves summing the products of differences between individual observations and their respective means, normalized by the product of square roots of the sums of squares of differences between observations and their respective means in both variables.

After identifying unique poses, it becomes possible to determine the ideal pairs as well as the sequence of map fusion. By examining the similarity of unique poses among all robots, we can identify pairs that exhibit higher degree of agreement, with the corresponding ρ being higher than a predefined threshold θ. Focusing specifically on unique regions, this process of similarity analysis becomes particularly valid, as, at this stage, common regions present in maps from all robots will not interfere with the similarity assessment outcomes.

3.4 Find the Best Fit

The search for the optimal fusion parameters is conducted through the Particle Swarm Optimization (PSO) Algorithm proposed by Kennedy and Eberhart [9]. In this optimization process, the objective function plays a crucial role in guiding the search towards the ideal solution. The adopted objective function is the mean squared error between the masks generated from unique poses, as defined in eq. (3):

$$\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2, \tag{3}$$

where, n represents the total number of data points used in comparing the masks, $\hat{y}_i$ is the predicted (or estimated) value of the mask at position i, and y_i is the actual or observed value of the mask at position i. This equation calculates the average of the squared differences between the predicted values $\hat{y}_i$ and the actual values y_i for all data points, providing a quantitative measure of deviation between masks generated from unique poses. The smaller the value resulting from eq. (3), the closer the masks are, indicating a more accurate alignment between the maps.

The Particle Swarm Optimization (PSO) is employed to optimize the determination of required horizontal translation Δ_x, the required vertical translation Δ_y and the required angle Δ_ϕ for rotation that lead to a more precise map fusion. PSO simulates the behavior of a swarm of particles, where each particle represents a candidate solution in the search space. Through iterations and exploration of solution space, particles collaborate to find parameters that minimize the objective function, enabling convergence towards an optimal fusion alignment.

3.5 Translation and Rotation

With the located parameters, map fusion is performed by applying a translational and rotational transformation based on the parameters identified by PSO, and the process is repeated until all maps are aligned in a common frame of reference. To perform a translation followed by a rotation, we first apply the translation and then the rotation. Thus, to translate and rotate the points of the masks, we use the following eq. (4):

$$\begin{aligned} x' &= (x + \Delta_x) \cdot \cos(\Delta_\phi) - (y + \Delta_y) \cdot \sin(\Delta_\phi) \\ y' &= (x + \Delta_x) \cdot \sin(\Delta_\phi) + (y + \Delta_y) \cdot \cos(\Delta_\phi), \end{aligned} \tag{4}$$

where, x' and y' represent the coordinates of the point after the application of both translation and rotation.

3.6 Maps Fusion

To formally define the problem of map fusion, we consider a set of N individual maps denoted as $M_1, M_2, \ldots, M_N$, each associated with a corresponding robot. Each map M_i is represented as a set of features F_i extracted from robot measurements. The goal of map fusion is to generate a fused global map M_{global} that accurately encapsulates the combined information from all individual maps.

Let M_1 and M_2 be two maps represented as matrices of dimensions $r \times c$, where r is the number of rows and c is the number of columns. Each cell $M_1(i, j)$ and $M_2(i, j)$ represents the state of the pixel at row i and column j of maps M_1 and M_2, respectively. The pixel state in a map can be one of the following: "free", "occupied", or "unknown". Map fusion, denoted as M_{fusion}, is obtained by considering the following rules:

$$M_{\text{fusion}}(i, j) = \begin{cases} \text{"unknown"}, & \text{if } M_1(i, j) = \text{"unknown" and } M_2(i, j) = \text{"unknown"} \\ \text{"occupied"}, & \text{if } M_1(i, j) = \text{"occupied" or } M_2(i, j) = \text{"occupied"} \\ \text{"free"}, & \text{otherwise} \end{cases}$$

where $1 \leq i \leq r$ and $1 \leq j \leq c$.

4 Results

In order to simulate the process of cooperative local mapping involving multiple robots to generate the required datasets, a simulated environment is created. This environment consists of an occupancy grid map, as depicted in Fig. 2. During the navigation stage, each robot records its current pose and LIDAR readings, resulting in the construction of a dataset containing 5000 poses. Using these datasets, it becomes possible to generate occupancy grid maps for each robot, as illustrated in Fig. 3. Once the parameters are successfully located, we apply translation and rotation transformations to the M_2. The outcome of this process is showcased in Fig. 4.

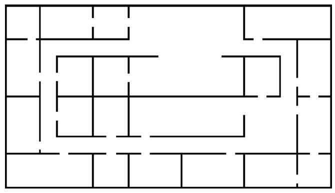
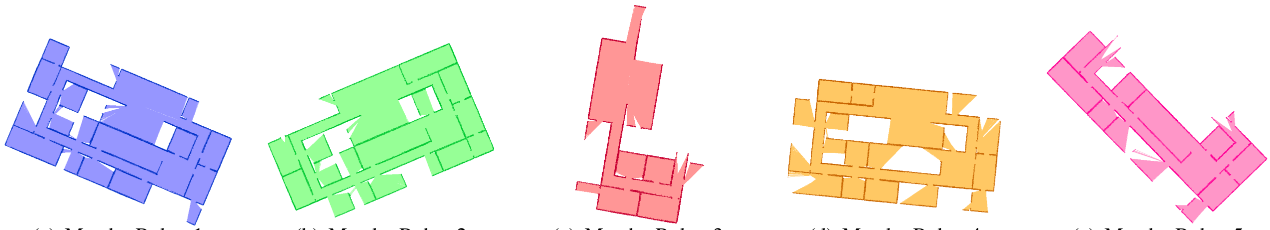
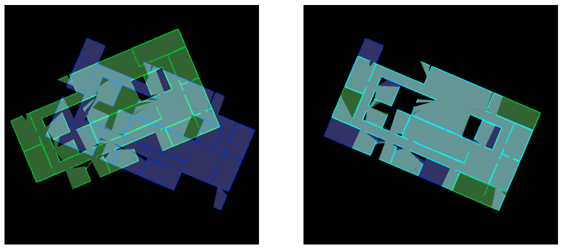


Figure 2. Grid map used in simulation



(a) Map by Robot 1 (b) Map by Robot 2 (c) Map by Robot 3 (d) Map by Robot 4 (e) Map by Robot 5

Figure 3. Maps constructed by robots during the SLAM process



(a) Before alignment (b) After alignment

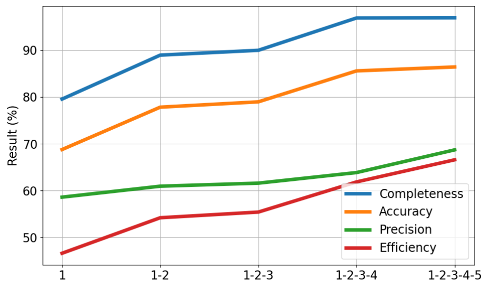
Figure 4. Map fusion of M_1 and M_2

We employ four metrics to evaluate the maps generated through each fusion process. The completeness is calculated as the ratio of the total number of mapped pixels to the number of pixels in the reference map. The accuracy is determined by the ratio of the total number of pixels accurately mapped to the total number of pixels in the reference map. The precision is gauged by the ratio of correctly mapped black (occupied) pixels to the total number of black pixels mapped. Lastly, the efficiency is derived as the product of completeness and precision. The results are presented in Table 1.

Table 1. Fusion results of the 5 maps

Maps	Completeness	Accuracy	Precision	Efficiency
M_1	79.55	68.77	58.60	46.51
$M_{1,2}$	88.94	77.82	60.94	54.06
$M_{1,2,3}$	89.97	78.96	61.59	55.54
$M_{1,2,3,4}$	96.87	85.57	63.85	62.29
$M_{1,2,3,4,5}$	**96.91**	**86.41**	**68.70**	**66.67**

The results demonstrate a consistent improvement across all evaluated parameters, leading to an overall enhanced performance. The graphical representations of the outcomes and the final map resulting from the fusion of the five maps are depicted in Fig. 5. Although a completeness of 100% is not achieved, the outcome came remarkably close to this value. Due to the lack of control or constraints on robot movements, many easily accessible areas were mapped multiple times by the same robot and by different robots. The incorporation of motion control techniques is a prospect that could potentially yield even greater gains.



(a) Line chart showing the results

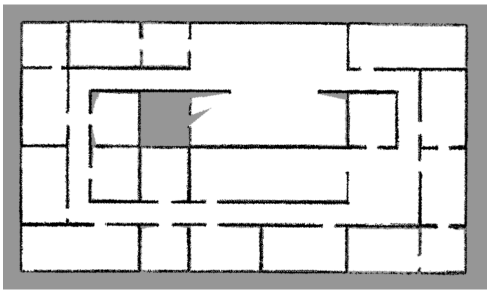

(b) Final map

Figure 5. Performance results obtained by the fusion of the 5 maps

5 Conclusions

In this paper, we have presented an innovative approach that utilizes Fourier Transform, the Pearson correlation coefficient and the Particle Swam Optimization for map fusion. By analyzing spectral features in the measurements of each robot, the Fourier Transform provides valuable insights into spatial distribution, while the Pearson correlation coefficient evaluates spectral similarity between different map sections and Particle Swam Optimization optimizes the search for alignment parameters. This approach enables the identification of specific regions, contributing to the efficient creation of a more accurate global map.

The experiments, conducted in a simulated environment involving five robots, have validated the effectiveness of the proposed approach. The results demonstrate consistent improvements in terms of completeness, accuracy, precision, and efficiency for the resulting maps from each fusion process. While we observe that a completeness of 100% is not achieved, the approach significantly approached this value, highlighting the optimization potential in map fusion. This work is a significant contribution to advancing research in collaborative mapping applications, paving the way for improvements in map functionality and C-SLAM research.

We also identify the opportunity to better utilize resources by restricting robot navigation in specific regions of the environment. As another future work, we intend to consider mapping area limitations for each robot, avoiding redundancy in exploring easily accessible regions and encouraging coverage of not mapped areas. Additionally, we envision the application of this approach in chaotic and dynamic environments, enabling the classification of fixed and dynamic obstacles through the observation of references from different robots.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] S. Dörr, P. Barsch, M. Gruhler, and F. G. Lopez. Cooperative longterm slam for navigating mobile robots in industrial applications. In *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pp. 297–303. IEEE, 2016.
[2] S. Saeedi, M. Trentini, M. Seto, and H. Li. Multiple-robot simultaneous localization and mapping: A review. *Journal of Field Robotics*, vol. 33, n. 1, pp. 3–46, 2016.
[3] L. Ma, J. Zhu, L. Zhu, S. Du, and J. Cui. Merging grid maps of different resolutions by scaling registration. *Robotica*, vol. 34, n. 11, pp. 2516–2531, 2016.
[4] A. Marjovi, S. Choobdar, and L. Marques. Robotic clusters: Multi-robot systems as computer clusters: A topological map merging demonstration. *Robotics and Autonomous Systems*, vol. 60, n. 9, pp. 1191–1204, 2012.
[5] J. Hörner. Map-merging for multi-robot system, 2016.
[6] S. Carpin. Fast and accurate map merging for multi-robot systems. *Autonomous robots*, vol. 25, pp. 305–316, 2008.
[7] E. O. Brigham. *The fast Fourier transform and its applications*. Prentice-Hall, Inc., 1988.
[8] I. Cohen, Y. Huang, J. Chen, J. Benesty, J. Benesty, J. Chen, Y. Huang, and I. Cohen. Pearson correlation coefficient. *Noise reduction in speech processing*, pp. 1–4, 2009.
[9] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks*, volume 4, pp. 1942–1948. IEEE, 1995.