Pingliang XU, Yaqi CUI, Wei XIONG, 2025. MH-T2TA: a multiple-hypothesis algorithm for multi-sensor track-to-track association with an intelligent track score. *Frontiers of Information Technology & Electronic Engineering*, 26(11):2231-2253. https://doi.org/10.1631/FITEE.2400340

MH-T2TA: a multiple-hypothesis algorithm for multi-sensor track-to-track association with an intelligent track score

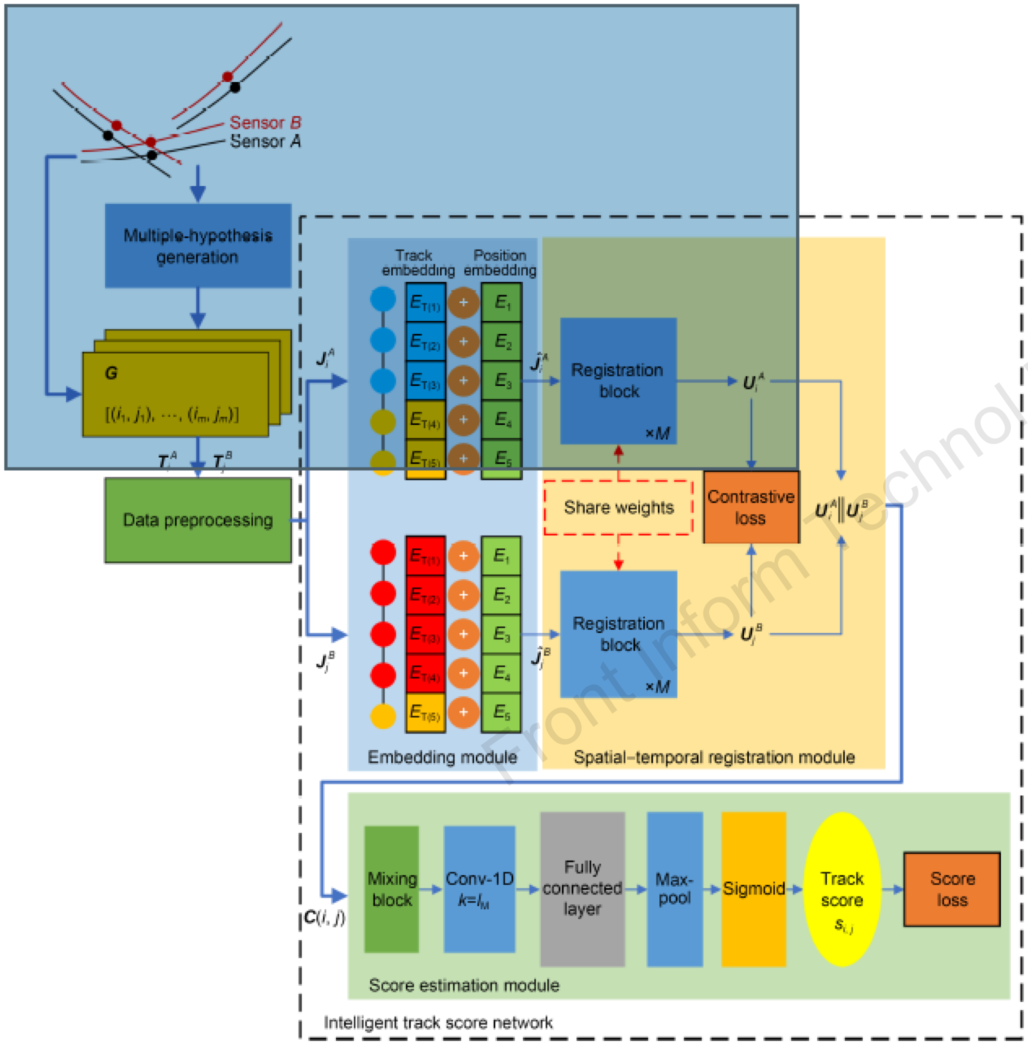
Key words: Track-to-track association; Multiple-hypothesis algorithm; Track score; Neural networks

Corresponding author: Yaqi CUI
E-mail: cui_yaqi@126.com
ORCID: https://orcid.org/0000-0002-4408-9962

Motivation

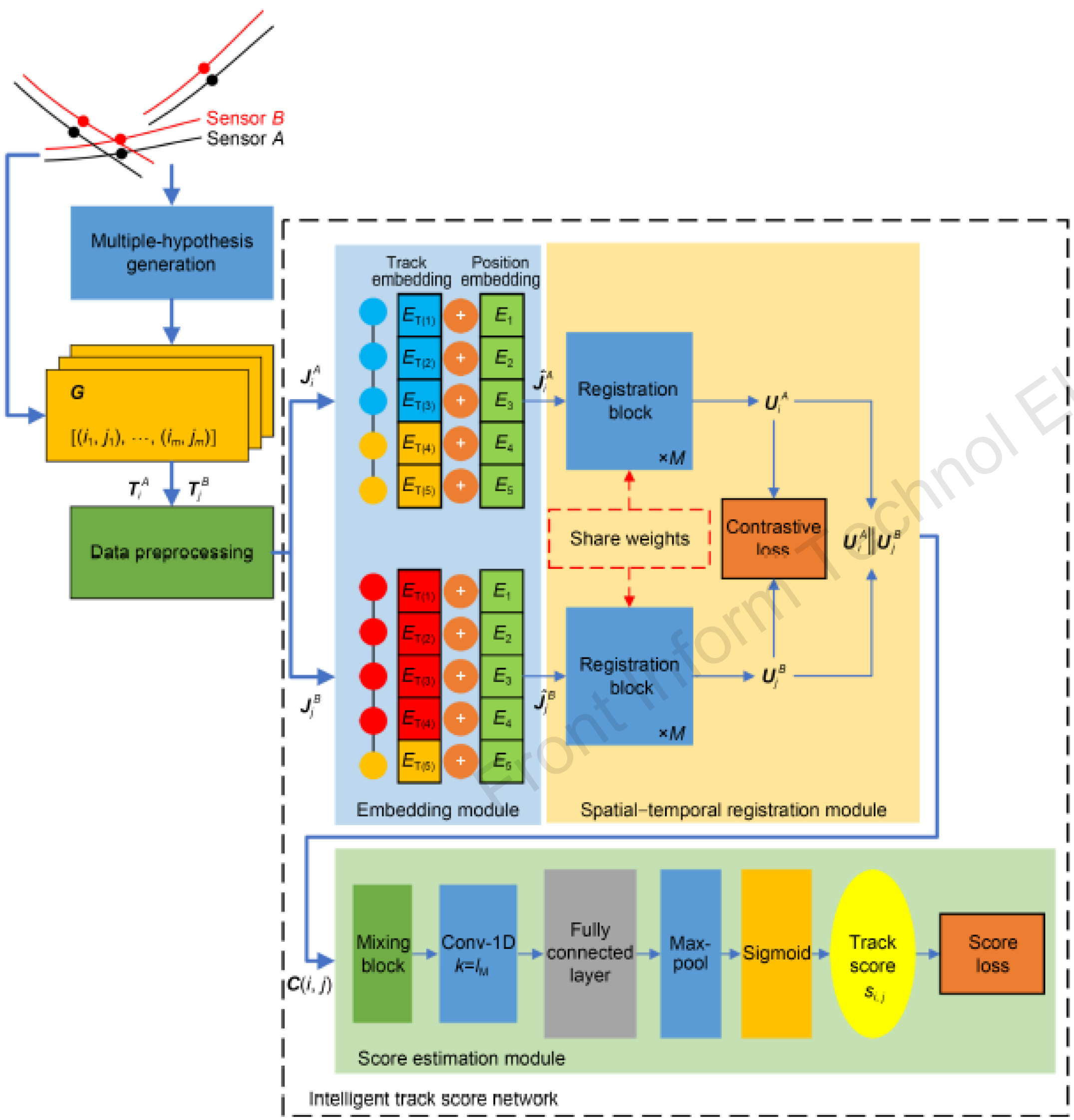
- It is difficult to obtain the optimal association result using heuristic association algorithms, but a multiple-hypothesis algorithm can solve this problem. Traditional track-to-track association (T2TA) methods depend heavily on prior information and assumed motion models, but a heuristic algorithm can solve this problem. The impact of various uncertainties results in inaccurate T2TA, but a heuristic algorithm can solve this problem.
- We extend the widely used multiple-hypothesis algorithm in MHT to T2TA problems and combine the multiple-hypothesis algorithm with deep learning. This can alleviate the dependencies on prior information and assumed motion models, reduce the impact of various uncertainties, and obtain optimal association results.
- We propose an attention-based network and a contrastive learning architecture to achieve the spatial–temporal registration of tracks from different sensors. A spatial–temporal mixing block is proposed to estimate track scores by sufficiently mixing spatial–temporal features.

Schematic of MH-T2TA



- MH-T2TA consists of two parts: multiple-hypothesis generation and an intelligent track score network.
- The network consists of three modules: an embedding module, a spatial–temporal registration module, and a score estimation module.
- Two branches in the spatial–temporal registration module share network parameters and weights.

Schematic of MH-T2TA



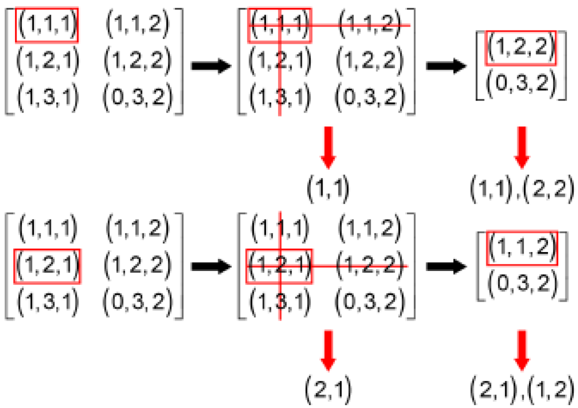
- MH-T2TA consists of two parts: multiple-hypothesis generation and an intelligent track score network.
- The network consists of three modules: an embedding module, a spatial–temporal registration module, and a score estimation module.
- Two branches in the spatial–temporal registration module share network parameters and weights.

Multiple-hypothesis generation

- The aim of multiple-hypothesis generation is to generate all possible track association matrices in one association cluster, with no association conflicts in each track association matrix. It includes gating and coarse association, association cluster, and hypothesis generation.

$$
\begin{cases}
G_{\mathrm{D}}=\sqrt{\left(\dfrac{1}{l_A}\sum_{k=1}^{l_A}x_i^{k,A}-\dfrac{1}{l_B}\sum_{k=1}^{l_B}x_j^{k,B}\right)^2+\left(\dfrac{1}{l_A}\sum_{k=1}^{l_A}y_i^{k,A}-\dfrac{1}{l_B}\sum_{k=1}^{l_B}y_j^{k,B}\right)^2}\le 0.1,\\
G_{\mathrm{S}}=\left|\dfrac{\sqrt{\left(x_i^{l_A,A}-x_i^{1,A}\right)^2+\left(y_i^{l_A,A}-y_i^{1,A}\right)^2}}{t_i^{l_A,A}-t_i^{1,A}}-\dfrac{\sqrt{\left(x_j^{l_B,B}-x_j^{1,B}\right)^2+\left(y_j^{l_B,B}-y_j^{1,B}\right)^2}}{t_j^{l_B,B}-t_j^{1,B}}\right|\le 0.1,\\
G_{\mathrm{H}}=\left|h_i^A-h_j^B\right|\le 60,\\
G_{\mathrm{T}}=\max\left(t_i^{1,A},t_j^{1,B}\right)-\min\left(t_i^{l_A,A},t_j^{l_B,B}\right)\le 0.
\end{cases}
$$

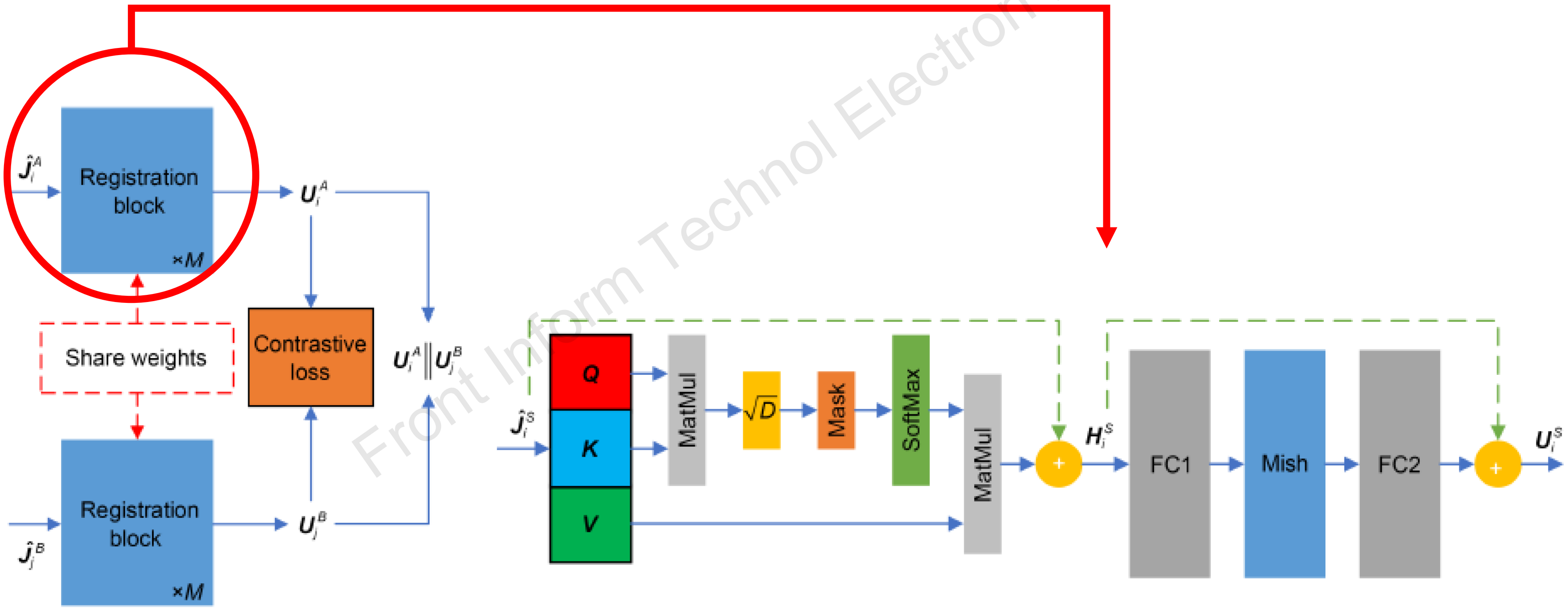
	1	2	3
1	1	1	0
2	1	1	0
3	0	0	1



Gating and coarse association | Association cluster | Hypothesis generation

Spatial–temporal registration module

- The aim of the spatial–temporal registration module is to eliminate the effects of errors and inconsistent update periods and obtain the unified registration track.



Contrastive learning

Attention mechanism

Loss function

- The different sensor contrastive loss L_{dc}
  - For different sensors, it is used to make the tracks from the same target close to each other, and the tracks from different targets move away from each other.

$$\boxed{\begin{aligned} L_{dc} &= \frac{1}{2} s_{i,j} D_d^{\ 2} + \frac{1}{2}\left(1 - s_{i,j}\right)\left[max\left(0, m - D_d\right)\right]^2 \\ D_d &= \left\| \boldsymbol{U}_i^A - \boldsymbol{U}_j^B \right\|_{\mathrm{F}} \end{aligned}}$$

- The same sensor contrastive loss L_{sc}
  - For the same sensor, it is used to make the tracks from different targets move away from each other.

$$\boxed{\begin{aligned} L_{sc} &= \left[max\left(0, m - D_s\right)\right]^2 \\ D_s &= \left\| \boldsymbol{U}_i^S - \boldsymbol{U}_j^S \right\|_{\mathrm{F}} \end{aligned}}$$

- The score loss L_s
  - It is used to make the track score estimated by the intelligent track score network as similar as possible to the real track score.

$$\boxed{L_s = \mathrm{MSE}\left(\hat{s}_{i,j}, s_{i,j}\right) = \frac{1}{2}\left(\hat{s}_{i,j} - s_{i,j}\right)^2}$$

$$\boxed{L = L_{dc} + L_{sc} + L_s}$$

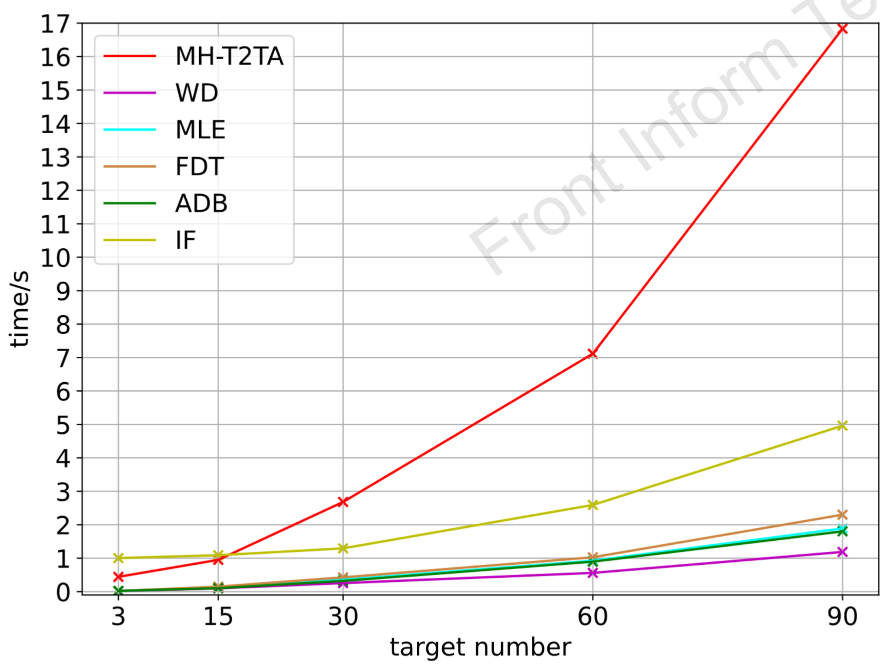
Contrastive experiments

Table 3 Association results of contrastive experiments

Method	AF1*	F1				
		3	15	30	60	90
WD (Kanyuck and Singer, 1970)	0.4949	1	0.8036	0.8069	0.6452	0.4311
MLE (Sun et al., 2023)	0.6858	1	0.9463	0.9144	0.7994	0.4799
REF (Qi et al., 2018)	0.7779	1	0.8704	0.7626	0.7896	0.7524
FDT (Du W et al., 2013)	0.6155	1	0.8490	0.8518	0.6670	0.4506
ADB (Xu and Fang, 2021)	0.7254	1	0.9891	0.9740	0.7765	0.5554
IF (Jin et al., 2023)	0.4665	1	0.9454	0.8228	0.4658	0.2507
HD (Hausdorff, 1914)	0.6059	1	0.8490	0.8118	0.6575	0.4492
FD (Alt and Godau, 1995)	0.6055	1	0.8490	0.8069	0.6642	0.4456
LCSS (Vlachos et al., 2002)	0.8257	1	0.9463	0.9208	0.8508	0.7513
ERP (Chen and Ng, 2004)	0.5502	1	0.9278	0.8404	0.5151	0.3989
EDR (Chen et al., 2005)	0.7609	1	0.9546	0.9306	0.7779	0.6528
MH-T2TA (ours)	0.8888	1	0.9848	1	0.9451	0.7946

* AF1 is the weighted mean F1 value of all fix testing scenarios

- MH-T2TA achieves the best AF1 and the best association performance.



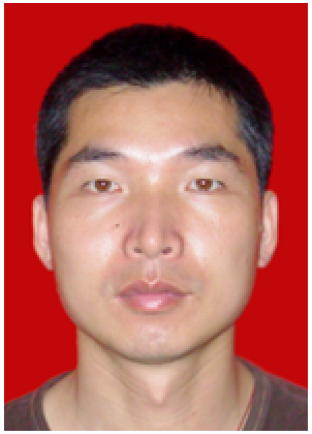
- However, as the number of targets increases, the excessive number of tracks within a single cluster in dense scenarios results in numerous association hypotheses for each cluster, leading to a significant increase in the time required to generate these association hypotheses.

Conclusions

- Aiming at the ubiquitous problems of errors and inconsistent update periods in the track data, we designed a spatial–temporal registration module based on self-attention and a contrastive learning architecture to eliminate errors and unify the distributions of asynchronous tracks.
- Focusing on the suboptimal association results and the dependencies on prior information and the assumed motion model, we combined the multiple-hypothesis algorithm with deep learning to construct an intelligent track score network for estimating the track score of a pair of tracks.
- How to further reduce the time required for hypothesis generation to enhance association efficiency and how to improve the interpretability of the intelligent track score network are key points for future research.

Pingliang XU received B.S. and M.S. degrees from Naval Aviation University, Yantai, China, in 2019 and 2021 respectively, and is currently pursuing a Ph.D. degree in information and communication engineering at Naval Aviation University. His research interests include track association, information fusion, and deep learning.

Yaqi CUI received B.S., M.S., and Ph.D. degrees in information and communication engineering from Naval Aviation University, Yantai, China, in 2008, 2011, and 2014, respectively. Now, he is an associate professor at the Naval Aviation University. His research interests include information fusion, machine learning, and deep learning with their applications in information fusion.



Wei XIONG received B.S., M.S., and Ph.D. degrees from Naval Aviation University, Yantai, China, in 1998, 2001, and 2005, respectively. From 2007 to 2009, he was a Postdoctoral Researcher at the Department of Electronic Information Engineering, Tsinghua University, Beijing. He is currently a Full Professor at the Naval Aviation University, where he teaches random signal processing and information fusion. He is one of the Founders and Directors of the Institute of Information Fusion, Naval Aviation University. He is a Member and Director General of the Information Fusion Branch of the Chinese Society of Aeronautics and Astronautics. His research interests include pattern recognition, remote sensing, and multi-sensor information fusion.