

# Multimodal AI-Mediated Interaction for Cross-Disciplinary Design Collaboration

<b>Submitted to</b>	Lin Li, Program Chair — IEEE MIPR 2026   cathylin@whut.edu.cn
<b>Session Organizer</b>	Prof. Teng-Wen Chang — National Yunlin University of Science and Technology, Taiwan   tengwen@yuntech.edu.tw
<b>Conference</b>	9th IEEE International Conference on Multimedia Information Processing and Retrieval (MIPR 2026)   Bangkok, Thailand   9–11 August 2026
<b>Paper Deadline</b>	<b>May 31st, 2026</b>
<b>Date Submitted</b>	March 27, 2026

## 1. SESSION TITLE AND THEME

**Session Title:** Multimodal AI-Mediated Interaction for Cross-Disciplinary Design Collaboration

**Conference Theme Alignment:** Multimedia Intelligence for Sustainable Futures

**Primary MIPR Tracks:** Generative and Foundation Models in Multimedia | Affective and Perceptual Multimedia | Mobile and Wearable Multimedia | Content Understanding

## 2. SESSION RATIONALE

Design collaboration in complex engineering, urban planning, and product development increasingly involves heterogeneous disciplinary teams working with AI-generated multimedia artifacts — images, physiological biosignals, semantic vectors — as shared mediating objects. Yet existing multimedia systems research treats these modalities separately: generative AI for creative ideation, physiological sensing for affective computing, and semantic processing for coordination tracking are each active subfields with no integrating framework.

This special session addresses precisely this gap. It brings together researchers across three converging areas to establish a unified research agenda for AI-mediated multimodal collaboration:

- Generative visual mediation — how text-to-image systems shape semantic alignment trajectories across iterative collaborative design cycles
- Affective-semantic fusion — how physiological signals (HRV, EDA) from wearable devices complement semantic distance metrics to provide earlier and richer coordination diagnostics
- AI as boundary object — how AI-generated images function not as convergence tools but as negotiation triggers that reveal and productively reshape disciplinary differences

The session directly addresses the MIPR 2026 theme of Multimedia Intelligence for Sustainable Futures by contributing computational frameworks, system architectures, and empirical evidence for how multimodal AI can sustainably support human collaborative processes — not by replacing human judgment but by making coordination dynamics visible and actionable.

A defining feature of this session is its emphasis on trajectory-based assessment: rather than evaluating collaboration at a single endpoint, all contributing papers examine how alignment, ideation, and affective states evolve across iterative interaction cycles. This temporal, process-oriented perspective represents a significant advancement over the dominant static-outcome paradigm in multimedia collaboration research.

### 3. SCOPE AND TOPICS OF INTEREST

The session invites original research contributions on the following topics:

#### 3.1 Generative AI and Visual Mediation in Collaboration

- Text-to-image generation as alignment instruments in multi-stakeholder design
- Image specificity, ambiguity, and their effects on semantic convergence trajectories
- AI-generated images as boundary objects in participatory and cross-disciplinary design
- Trajectory-level analysis of visual mediation effects across iterative cycles

#### 3.2 Multimodal Affective Sensing for Collaborative Systems

- Wearable physiological sensing (HRV, EDA, ECG) for real-time affective state monitoring in team settings
- Fusion of semantic and physiological signals for coordination assessment
- Temporal relationships between affective and cognitive coordination dynamics
- Affective feedback display design for collaborative environments

#### 3.3 Computational Frameworks for Coordination Assessment

- Semantic embedding methods (BERT, SBERT, CLIP) for measuring cross-disciplinary alignment
- Coordination metrics and indices for iterative design collaboration
- Trajectory classification: monotonic convergence, delayed convergence, stable non-convergence
- Decision support systems for early-stage engineering design under semantic heterogeneity

#### 3.4 Human-AI Co-Ideation and Collaborative Creativity

- Human-AI co-creation systems for design ideation and concept exploration
- Tracking the evolution of design ideas across iterative human-AI cycles
- Effects of generative AI on design fixation and divergent thinking
- Interaction design principles for AI-assisted cross-disciplinary teams

### 4. SESSION PAPERS

The session comprises seven papers spanning the organizer's own submissions and invited contributions from leading researchers in the field. All papers address the session's core theme from complementary perspectives.

#	Paper Title / Authors	Institution	MIPR Track	Status
---	-----------------------	-------------	------------	--------

#	Paper Title / Authors	Institution	MIPR Track	Status
1	<b>Generative Visual Mediation in Cross-Disciplinary Collaboration: Trajectory Analysis of Text-to-Image Alignment Across Iterative Design Cycles</b> Teng-Wen Chang, Jia-Rong Li, Wei-Lung Mao, Song Xu, Ching-Chih Chang ★ Session Organizer's Paper (Anchor)	National Yunlin University of Science and Technology, Taiwan	Generative & Foundation Models — Alignments in T2I Generation	Submitting
2	<b>Beyond Semantic Distance: Fusing Physiological and Linguistic Signals for Multimodal Coordination Assessment in Cross-Disciplinary Design</b> Teng-Wen Chang, Jia-Rong Li, Ya-Chen Chang, Hsin-Yi Huang, Chen-Syuan Lin ★ Session Organizer's Paper	National Yunlin University of Science and Technology, Taiwan	Affective and Perceptual Multimedia — Mobile and Wearable Multimedia	Submitting
3	<b>IdeationWeb: Tracking the Evolution of Design Ideas in Human-AI Co-Creation</b> Hanshu Shen, Lyukesheng Shen, Wenqi Wu, Kejun Zhang	CHI 2025 paper — corresponding institution TBC	Generative & Foundation Models — Multimodal Agents	Invited
4	<b>From Fake Perfects to Conversational Imperfects: Exploring Image-Generative AI as a Boundary Object for Participatory Design of Public Spaces</b> Jose A. Guridi, Angel Hsing-Chi Hwang, Duarte Santo, Maria Goula, Cristobal Cheyre, Lee Humphreys, Marco Rangel	Cornell University, University of Southern California, USA	Content Understanding — Affective and Perceptual Multimedia	Invited
5	<b>Aldeation: Designing a Human-AI Collaborative Ideation System for Concept Designers</b> Wen-Fan Wang, Chien-Ting Lu, Nil Ponsa Campanyà, Bing-Yu Chen, Mike Y. Chen	National Taiwan University, Taiwan	Generative & Foundation Models — Alignments in T2I Generation	Invited
6	<b>The Effects of Generative AI on Design Fixation and Divergent Thinking</b> Wadinambiarachchi et al.	CHI 2024 — institution TBC	Content Understanding — Multimodal/Multisensory Interfaces	Invited
7	<b>An Ensemble Deep Learning Framework for Emotion Recognition Through Wearable Devices Multi-Modal Physiological Signals</b> Author TBC — Scientific Reports 2025	Institution TBC	Machine/Deep Learning — Mobile and Wearable Multimedia	Invited

Note: Papers marked ★ are confirmed submissions by the session organizer. Papers marked 'Invited' are under active invitation; confirmation pending. Green rows = organizer submissions.

## 5. SESSION COHERENCE AND INTELLECTUAL STRUCTURE

The seven papers form a coherent intellectual arc across three levels of analysis:

### Level 1 — Individual Human-AI Co-Creation (Papers 3, 5, 6)

IdeationWeb, Aldeation, and the design fixation study examine how individual designers interact with AI-generated content in ideation workflows. Together they establish the state of the art on human-AI co-ideation and identify both the benefits (enhanced creativity, faster exploration) and the risks (fixation,

reduced divergent thinking) of generative AI in design.

### Level 2 — Multi-Party Visual Mediation and Alignment (Papers 1, 4)

The session organizer's P2 paper and the Guridi et al. boundary object paper both examine how AI-generated images function as mediators in multi-stakeholder collaborative settings. Both papers converge on the same counter-intuitive finding — that visually ambiguous, imperfect images produce better collaborative outcomes than highly specified, polished images — but approach it from complementary methodological positions: computational trajectory analysis (P2) vs. qualitative participatory design (Guridi et al.).

### Level 3 — Multimodal Coordination Monitoring (Papers 2, 7)

The session organizer's P1 paper and the wearable deep learning paper both address the challenge of monitoring collaboration states through physiological and semantic signals. P1 introduces the bimodal ASCI framework; Paper 7 provides the deep learning architecture for robust HRV-based emotion classification. Together they establish the technical and theoretical foundations for real-time affective-aware collaboration monitoring systems.

Across all three levels, the session is unified by the trajectory-based assessment paradigm: each paper treats collaboration as a temporally evolving process rather than a static agreement state. This represents a distinctive and coherent methodological position that differentiates the session from existing multimedia collaboration tracks.

## 6. SESSION ORGANIZER

<b>Name</b>	Prof. Teng-Wen Chang, PhD
<b>Position</b>	Professor, Department of Digital Media Design, National Yunlin University of Science and Technology (YunTech), Taiwan
<b>Email</b>	tengwen@yuntech.edu.tw
<b>Research Areas</b>	Interaction Design, Human-Computer Interaction, Scenario-Based Interaction, Computational Design, AI-Mediated Collaboration
<b>Citations</b>	874 total citations   H-index: 13   i10-index: 21 (Google Scholar, March 2026)
<b>Relevant Publications</b>	Chang et al. (2025). Empathy Game Toolkit. CoDesign. Chang et al. (2024). Discovering semantic and visual hints with ML. Advanced Engineering Informatics, 59, 102244. Li, Chang et al. (2023). Constructing a Cross-Disciplinary Idea Convergence System Using AIGC. IEEE IV 2023, 352–357. Chang et al. (2026). Tracking semantic coordination dynamics (under review). Advanced Engineering Informatics.

## 7. EXPECTED CONTRIBUTIONS TO MIPR 2026

This special session makes four distinct contributions to the MIPR 2026 programme:

- It introduces trajectory-based assessment as a new analytical paradigm for multimedia collaboration research — shifting from endpoint outcomes to temporal dynamics of alignment, ideation, and affective engagement.
- It provides the first integrated computational framework combining generative AI (text-to-image), semantic NLP (SBERT), and physiological sensing (HRV) into a unified multimodal collaboration monitoring system — a genuinely novel multimedia systems contribution.
- It establishes a distinctive Asia-Pacific research perspective by centering contributions from YunTech and National Taiwan University alongside leading North American and European researchers, directly serving MIPR 2026's Bangkok location and regional orientation.
- It creates a productive intellectual tension within the session — papers showing AI enhances creativity (Aldeation) alongside papers showing AI causes fixation (Wadinambiarachchi et al.) alongside papers showing AI functions as a productive boundary object (Guridi et al.) — that will generate substantive discussion and future research directions.

## 8. PROPOSED SESSION FORMAT

<b>Session Duration</b>	90 minutes (7 papers × 10 minutes + 20 minutes panel discussion)
<b>Presentation Format</b>	10-minute presentation per paper (8 min talk + 2 min Q&A) followed by 20-minute structured panel discussion on the session's unifying theme
<b>Demo Opportunity</b>	The session organizer will bring the Emotion Wings wearable device for a live demonstration of real-time HRV-based affective feedback during the panel discussion — a highly memorable and concrete illustration of the session's multimodal theme.
<b>Panel Discussion Theme</b>	"When AI mediates collaboration: should we optimize for convergence, or for productive divergence?" — a discussion question that directly emerges from the contrasting findings across the session's papers.
<b>Estimated Attendance</b>	30–50 attendees, drawing from MIPR's multimedia systems, affective computing, and HCI research communities.

---

Thank you for considering this proposal. I am committed to organizing a high-quality, well-attended session that advances the MIPR community's understanding of multimodal AI-mediated collaboration. I welcome any questions or suggestions regarding the session scope, paper list, or format.

**Prof. Teng-Wen Chang**

National Yunlin University of Science and Technology, Taiwan

tengwen@yuntech.edu.tw | softlab.tw

March 27, 2026