

Programme Overview

Start	End	22-Oct	23-Oct	24-Oct
8:00	8:30	Tutorials		
8:30	9:00			
9:00	9:30		Technical Sessions	Technical Sessions
9:30	10:00		Break	Break
10:00	10:30			
10:30	11:00		Technical Sessions	Technical Sessions
11:00	11:30			
11:30	12:00	Welcome		
12:00	12:30	Reception & Opening	Lunch	Lunch
12:30	13:00		WiAPSIPA	
13:00	13:30			
13:30	14:00	Keynote 1	Keynote 2	Keynote 3
14:00	14:30			
14:30	15:00	Technical Sessions	Technical Sessions	Technical Sessions
15:00	15:30			
15:30	16:00	Break	Break	Break
16:00	16:30			
16:30	17:00	Technical Sessions	Technical Sessions	Founders Forum & Closing
17:00	17:30			
17:30	18:00			
18:00	18:30			
18:30	19:00			
19:00	19:30			
19:30	20:00			
20:00	20:30		Banquet	
20:30	21:00			
21:00	21:30			

Onsite registration opens at 10am on 22 Oct.

Best Paper Finalists are given 12-min presentation 3-min Q&A at Conference Awards Sessions I & II

All contributed full papers are given 5-min presentation and 30-min poster interaction (max A0 portrait) in the same room.

All late-breaking reports are given 5-min presentation only.

All authors are required to report to their respective sessions at least 15 mins ahead of the session to upload and test their slides.

Paper Allocations in Technical Sessions

Session	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
Active Noise Control I	56	57	72	124	285	289	301	345	511	553	610	626
Active Noise Control II	63	87	134	265	305	310	330	340	458	490	497	513
Advanced Multimedia Applications	113	158	297	364	367	377	382	423	473	516		
Advanced Signal Processing and Machine Learning for Acoustic Scene Analysis and Signal Enhancement	51	184	186	201	234	293	418	439	442	474	501	507
Advanced Signal Processing and Resource Management for Sensing and Communication	144	167	169	219	220	341	400	416				
Advanced Topics in Audio Understanding of Sound Events, Scenes, and Beyond	58	103	117	157	211	258	275	328	389	398	459	461
Advanced Topics on Music Processing	177	283	298	319	337	349	356	424	441	453	496	517
Advances in Multimodal AI for Multimedia Applications	59	68	96	245	249	252	268	317	354	359	436	
Audio Processing	129	140	174	408	409	434	492	583				
Best Paper Finalists	105	402	69	284	36	355	530	586				
Biomedical Signal Processing and Systems I	26	148	173	214	240	241	336	407	413	415	447	227
Biomedical Signal Processing and Systems II	454	466	531	535	591	630	631	632				
Deep Learning: Algorithm, Implementations, and Applications	49	55	84	108	185	372	385	386	387	422	470	
Emerging Technologies and Applications of Image Processing and Computer Vision	54	143	171	196	269	332	445	518	548	568	572	
Generative AI Models for Vision-Based Applications	91	92	147	242	251	533	536	561	615			
Grand Challenge & Audio Processing	110	138	139	228	401	588	634	642	648	651	652	653
Interactive, Natural, Expressive and Robust Conversation System	42	60	73	78	102	112	152	199	217	321	483	
Machine Learning: Algorithms and Application I	22	47	216	229	274	294	313	361	576	578	609	
Machine Learning: Algorithms and Application II	83	207	232	309	343	352	370	412	460	557	644	625
Machine Learning: Algorithms and Application III	35	239	291	311	312	314	411	457	493	520	559	
Machine Learning: Information and Medical Applications	24	260	278	300	303	324	431	558	580	594	595	
Multimedia Security & Forensics	104	135	215	348	379	410	464					
Multimodal AI	23	155	193	299	510	541	550	562				
Music and Singing	71	94	141	307	362	391	421	451	482	498	569	
Privacy and Security in Speech AI	86	133	194	200	417	495	515	538	547	608	623	
Recent Advances in Multimedia Enrichment, Security and Privacy	89	109	136	192	210	218	225	292	320	325	469	
Recent Advances in Multimodal Learning for Computer Vision	52	53	160	205	302	334	363	375	388	399	604	
Research Frontiers in Learned Visual Data Coding and Processing	64	106	161	204	335	383	465	468	532	539	621	
Scalable and Efficient Signal Processing for Multimodal AI Systems (ESPRESSO)	115	153	206	267	481	505	523	534	563	599	602	
Signal & Information Processing I	46	61	66	76	159	202	244	259	271	318	627	
Signal & Information Processing II	137	322	347	381	393	394	455	480				
Speaker Modeling Beyond Speaker Recognition	41	81	101	131	132	182	189	197	333	590		
Speech and Language Processing I	95	127	142	150	179	213	236	253	262	446	571	
Speech and Language Processing II	97	118	120	154	264	308	331	338	440	587	603	
Speech and Language Processing III	27	130	163	168	191	208	235	237	248	261	414	
Speech and Language Processing IV	29	116	449	478	502	503	514	551	560	577	641	
Speech Communication and Media Generation	256	272	344	384	427	428	450	452	472	487	616	
Speech Recognition and Understanding I	195	282	390	406	426	443	475	524	592	593		
Speech Recognition and Understanding II	151	226	238	280	296	342	392	405	542	556	575	
Speech Synthesis, Enhancement, and Recognition	17	48	98	149	243	304	491	500	554	600		
Tracing the Fake: Deepfake Detection, Attribution & Spoof-aware Speaker Verification Across Languages & Accents	126	187	188	203	212	254	276	327	508	509		
Wireless Communications & Networking	39	50	74	128	209	270	287	290	351	456		

22 October

Start	End	Island Ballroom	Lotus I	Lotus II	Hibiscus III	Peony I	Peony II	
8:00	8:30	Exhibition Setup	Tutorial 1: Recent Advances in End-to-End Learned Image and Video Coding	Tutorial 2: Deep Speaker Modeling: Theories, Applications and Practice	Tutorial 3: From Detection to Direction: An Overview of Sound Event Localization and Detection	Tutorial 4: Adaptive Sensor Networks in Digital Health	Conference Subcomm Meeting	
8:30	9:00							
9:00	9:30							
9:30	10:00							
10:00	10:30	Registration						
10:30	11:00							
11:00	11:30							
11:30	12:00	Welcome Reception & Opening Ceremony						
12:00	12:30							
12:30	13:00							
13:00	13:30							
13:30	14:00	Keynote 1 by Emanuel Habets						
14:00	14:30							
14:30	15:00	Perspective 1: Voice Privacy & Security	Advanced Signal Processing and Machine Learning for Acoustic Scene Analysis and Signal Enhancement	Interactive, Natural, Expressive and Robust Conversation System	Generative AI Models for Vision-Based Applications	Multimedia Security & Forensics	Wireless Communications & Networking	
15:00	15:30							
15:30	16:00							
16:00	16:30	Break						
16:30	17:00	Perspective 2: Utilization of the Foundation Models and the Future	Tracing the Fake: Deepfake Detection, Attribution & Spoof-aware Speaker Verification Across Languages & Accents	Speaker Modeling Beyond Speaker Recognition	Three-Minute Thesis (3MT) Competition	OC Meeting	Advanced Multimedia Applications	
17:00	17:30							
17:30	18:00							
18:00	18:30	APSIPA BoG Meeting						
18:30	19:00							
19:00	19:30							
19:30	20:00							
20:00	20:30							
20:30	21:00							
21:00	21:30							

23 October

Start	End	Island Ballroom	Lotus I	Lotus II	Hibiscus III	Peony I	Peony II
8:00	8:30	Registration					
8:30	9:00	Active Noise Cancellation Workshop Keynote	Advanced Topics in Audio Understanding of Sound Events, Scenes, and Beyond	Speech and Language Processing I	Research Frontiers in Learned Visual Data Coding and Processing	Multimodal AI	Biomedical Signal Processing and Systems I
9:00	9:30						
9:30	10:00						
10:00	10:30	Break					
10:30	11:00	Active Noise Cancellation Panel Discussion	Advanced Topics on Music Processing	Speech and Language Processing II	Emerging Technologies and Applications of Image Processing and Computer Vision	Biomedical Signal Processing and Systems II	Machine Learning: Algorithms and Application I
11:00	11:30						
11:30	12:00						
12:00	12:30	Lunch					
12:30	13:00	Women in APSIPA Forum	TC Meeting (BioSIPS - WCN)	TC Meeting (SPS)	TC Meeting (AAM - SLP)	TC Meeting (IVM - MSF)	TC Meeting (SIPTM - MLDA)
13:00	13:30						
13:30	14:00						
13:30	14:00	Keynote 2 by Jane Wang					
14:00	14:30						
14:30	15:00	Perspective 3: Neural Speech Assessment and Its Application	Active Noise Control I	Speech and Language Processing III	Machine Learning: Information and Medical Applications	Audio Processing	Signal & Information Processing I
15:00	15:30						
15:30	16:00						
16:00	16:30	Break					
16:30	17:00	Education Forum	Active Noise Control II	Speech and Language Processing IV	Recent Advances in Multimedia Enrichment, Security and Privacy	OC Meeting	Advances in Multimodal AI for Multimedia Applications
17:00	17:30						
17:30	18:00						
18:00	18:30	Banquet					
18:30	19:00						
19:00	19:30						
19:30	20:00						
20:00	20:30						
20:30	21:00						
21:00	21:30						

24 October

Start	End	Island Ballroom	Lotus I	Lotus II	Hibiscus III	Peony I	Peony II
8:00	8:30	Registration					
8:30	9:00	Sadaoki Furui Prize Papers and Conference Award Session I	Music and Singing	Speech Communication and Media Generation	Deep Learning: Algorithm, Implementations, and Applications	Advanced Signal Processing and Resource Management for Sensing and Communication	Machine Learning: Algorithms and Application II
9:00	9:30						
9:30	10:00						
10:00	10:30	Break					
10:30	11:00	Conference Award Session II	Grand Challenge & Audio Processing	Speech Recognition and Understanding I	Recent Advances in Multimodal Learning for Computer Vision	Signal & Information Processing II	Machine Learning: Algorithms and Application III
11:00	11:30						
11:30	12:00						
12:00	12:30						
12:30	13:00	Lunch	TC Meeting	TC Meeting	TC Meeting	TC Meeting	TC Meeting
13:00	13:30						
13:30	14:00	Keyntoe 3 by Xiao-Ping (Steven) Zhang					
14:00	14:30						
14:30	15:00	Perspective 4: Leveraging LLMs in Interdisciplinary Study of Speech and Language	Speech Synthesis, Enhancement, and Recognition	Speech Recognition and Understanding II	Privacy and Security in Speech AI	Late Breaking Reports	Scalable and Efficient Signal Processing for Multimodal AI Systems (ESPRESSO)
15:00	15:30						
15:30	16:00						
16:00	16:30	Break					
16:30	17:00						
17:00	17:30	APSIPA General Assembly, Founders' Forum & Closing Ceremony					
17:30	18:00						

Date	Start	End	Island Ballroom	Chair
	8:00	10:00		
	10:00	10:30	Set Up	
	10:30	11:30		
22	11:30	13:30	Welcome Reception & Opening Ceremony	Woon Seng Gan
Oct	13:30	14:30	Keynote 1 by Emanuël Habets	Shoji Makino
2025	14:30	16:00	Perspective 1: Voice Privacy & Security	Liping Chen
Wed	16:00	16:30	Break	
	16:30	18:00	Perspective 2: Utilization of the Foundation Models and the Future	Isao Echizen
	18:00	19:00		
	19:00	21:30		
	8:30	10:00	Active Noise Cancellation Workshop Keynote	Woon Seng Gan
	10:00	10:30	Break	
	10:30	12:00	Active Noise Cancellation Panel Discussion	Woon Seng Gan
23	12:00	13:30	Women in APSIPA Forum (12.20-13.30)	Bonnie Law
Oct	13:30	14:30	Keynote 2 by Jane Wang	Jay Kuo
2025	14:30	16:00	Perspective 3: Neural Speech Assessment and Its Application	Hsin-Min Wang
Thu	16:00	16:30	Break	
	16:30	18:00	Education Forum	Anthony Kuh
	18:00	19:00	Turnover for Banquet	
	19:00	21:30	Banquet	
	8:30	10:00	Sadaoki Furui Prize Papers and Conference Award Session I	Sang Hoon Lee
	10:00	10:30	Break	
24	10:30	12:00	Conference Award Session II	Yan Wu
Oct	12:00	13:30	Lunch	
2025	13:30	14:30	Keynote 3 by Xiao-Ping (Steven) Zhang	Haizhou Li
Fri	14:30	16:00	Perspective 4: Leveraging LLMs in Interdisciplinary Study of Speech and Language	Tan Lee
	16:00	16:30	Break	
	16:30	18:00	APSIPA General Assembly, Founders' Forum & Closing Ceremony	Woon Seng Gan

Date	Start	End	Lotus I	Chair	Co-Chair
	8:00	10:00			
	10:00	10:30	Tutorial 1: Recent Advances in End-to-End Learned Image and Video Coding		
	10:30	11:30			
22	11:30	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Advanced Signal Processing and Machine Learning for Acoustic Scene Analysis and Signal Enhancement	Shoji Makino	Kouei Yamaoka
Wed	16:00	16:30	Break		
	16:30	18:00	Tracing the Fake: Deepfake Detection, Attribution & Spoof-aware Speaker Verification Across Languages & Accents	Haizhou Li	Madhusudan Singh
	18:00	19:00	Turnover for BoG Meeting		
	19:00	21:30	BOG Meeting	Woon Seng Gan	
	8:30	10:00	Advanced Topics in Audio Understanding of Sound Events, Scenes, and Beyond	Nobutaka Ono	Keisuke Imoto
	10:00	10:30	Break		
	10:30	12:00	Advanced Topics on Music Processing	Tetsuro Kitahara	Eita Nakamura
23	12:00	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Active Noise Control I	Bai Mingsian	Shi Dong Yuan
Thu	16:00	16:30	Break		
	16:30	18:00	Active Noise Control II	Thushara Abhayapala	Jiancheng Tao
	18:00	19:00			
	19:00	21:30			
	8:30	10:00	Music and Singing	Soham Gangopadhyay	Tomohiko Nakamura
	10:00	10:30	Break		
24	10:30	12:00	Grand Challenge & Audio Processing	Ramesh R	Yang Xiao
Oct	12:00	13:30			
2025	13:30	14:30			
Fri	14:30	16:00	Speech Synthesis, Enhancement, and Recognition	Dyah A. M. G. Wisnu	Saisha Suresh Bore
	16:00	16:30	Break		
	16:30	18:00			

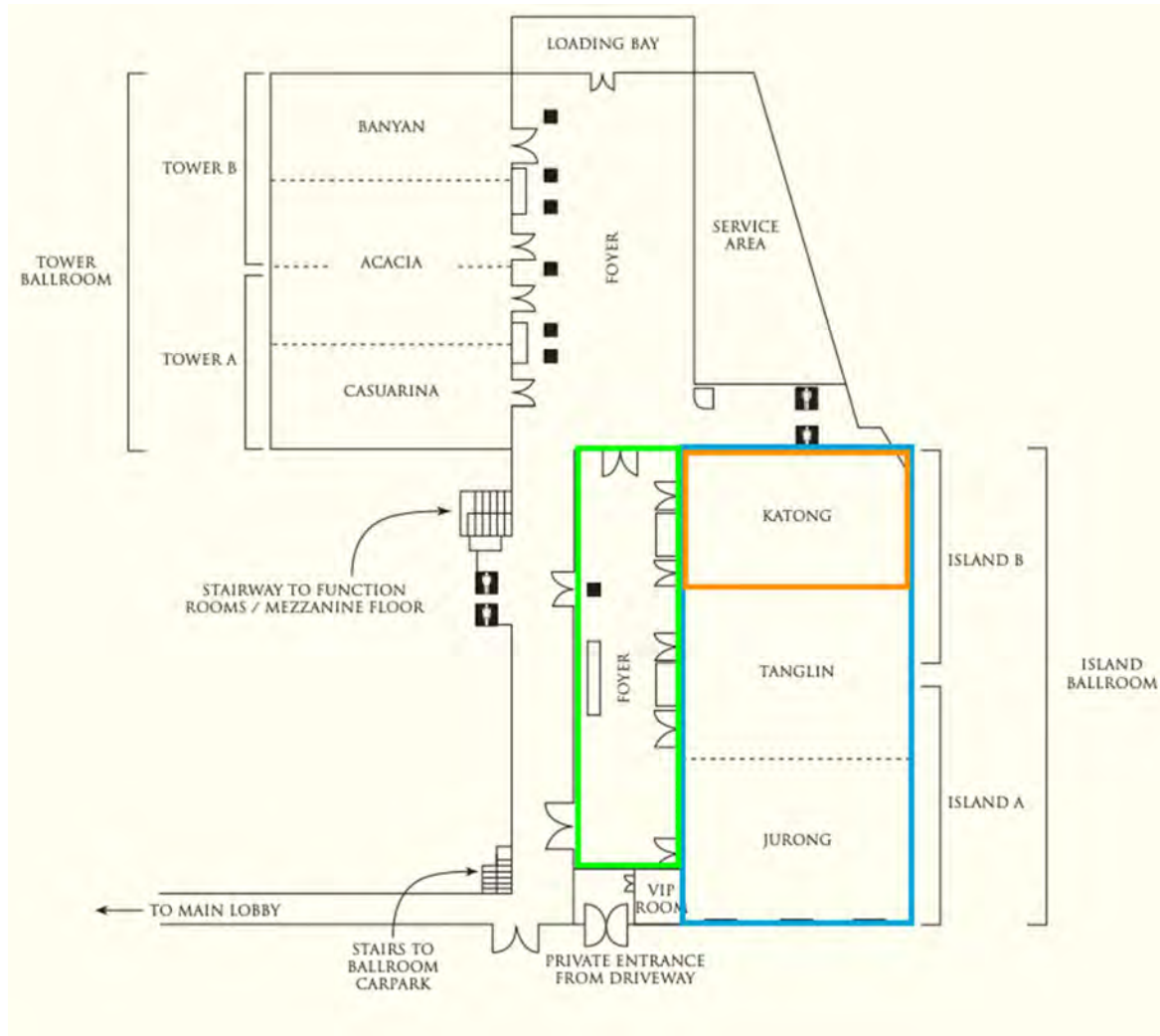
Date	Start	End	Lotus II	Chair	Co-Chair
	8:00	10:00			
	10:00	10:30	Tutorial 2: Deep Speaker Modeling: Theories, Applications and Practice		
	10:30	11:30			
22	11:30	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Interactive, Natural, Expressive and Robust Conversation System	Tetsuji Ogawa	Tzu-Quan Lin
Wed	16:00	16:30	Break		
	16:30	18:00	Speaker Modeling Beyond Speaker Recognition	Shuai Wang	Yanmin Qian
	18:00	19:00	Turnover for BoG Meeting		
	19:00	21:30	BOG Meeting	Woon Seng Gan	
	8:30	10:00	Speech and Language Processing I	Yang Ai	Eliathamby Ambikairajah
	10:00	10:30	Break		
	10:30	12:00	Speech and Language Processing II	Ashish Panda	Bagus Tris Atmaja
23	12:00	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Speech and Language Processing III	Liping Chen	Ivan Kukanov
Thu	16:00	16:30	Break		
	16:30	18:00	Speech and Language Processing IV	Xin Wang	Madhusudan Singh
	18:00	19:00			
	19:00	21:30			
	8:30	10:00	Speech Communication and Media Generation	Wen-Chin Huang	Kengo Ohta
	10:00	10:30	Break		
24	10:30	12:00	Speech Recognition and Understanding I	Sayaka Shiota	Virender kadyan
Oct	12:00	13:30			
2025	13:30	14:30			
Fri	14:30	16:00	Speech Recognition and Understanding II	Yu Tsao	Huijing Zhan
	16:00	16:30	Break		
	16:30	18:00			

Date	Start	End	Hibiscus III	Chair	Co-Chair
	8:00	10:00			
	10:00	10:30	Tutorial 3: From Detection to Direction: An Overview of Sound Event Localization and Detection		
	10:30	11:30			
22	11:30	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Generative AI Models for Vision-Based Applications	Po-Chyi Su	Hsiao-Chi Li
Wed	16:00	16:30	Break		
	16:30	18:00	Three-Minute Thesis (3MT) Competition	Woon Seng Gan	
	18:00	19:00			
	19:00	21:30			
	8:30	10:00	Research Frontiers in Learned Visual Data Coding and Processing	Wen-Hsiao Peng	Xiem Hoang
	10:00	10:30	Break		
	10:30	12:00	Emerging Technologies and Applications of Image Processing and Computer Vision	Chang-Su Kim	Keunsoo Ko
23	12:00	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Machine Learning: Information and Medical Applications	Kazushi Ikeda	Ting Dang
Thu	16:00	16:30	Break		
	16:30	18:00	Recent Advances in Multimedia Enrichment, Security and Privacy	Masaki Kawamura	Shoko Imaizumi
	18:00	19:00			
	19:00	21:30			
	8:30	10:00	Deep Learning: Algorithm, Implementations, and Applications	Shoji Makino	Stephen Karungaru
	10:00	10:30	Break		
24	10:30	12:00	Recent Advances in Multimodal Learning for Computer Vision	Akshita Abrol	Shogo MURAMATSU
Oct	12:00	13:30			
2025	13:30	14:30			
Fri	14:30	16:00	Privacy and Security in Speech AI	Tomoki Toda	Yiming Wang
	16:00	16:30	Break		
	16:30	18:00			

Date	Start	End	Peony I	Chair	Co-Chair
	8:00	10:00			
	10:00	10:30	Tutorial 4: Adaptive Sensor Networks in Digital Health		
	10:30	11:30			
22	11:30	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Multimedia Security & Forensics	Hong Huy Nguyen	Mehmet Sinan YILDIRIM
Wed	16:00	16:30	Break		
	16:30	18:00	OC Meeting		
	18:00	19:00			
	19:00	21:30			
	8:30	10:00	Multimodal AI	Chi-Chun (Jeremy) Lee	Keunsoo Ko
	10:00	10:30	Break		
	10:30	12:00	Biomedical Signal Processing and Systems II	Ingon Chanpornpakdi	Kazushi Ikeda
23	12:00	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Audio Processing	Natsuki Ueno	Huawei Zhang
Thu	16:00	16:30	Break		
	16:30	18:00	OC Meeting		
	18:00	19:00			
	19:00	21:30			
	8:30	10:00	Advanced Signal Processing and Resource Management for Sensing and Communication	Koichi Adachi	Kazunori Hayashi
	10:00	10:30	Break		
24	10:30	12:00	Signal & Information Processing II	Candy Olivia Mawalim	Candy Olivia Mawalim
Oct	12:00	13:30			
2025	13:30	14:30			
Fri	14:30	16:00	Late Breaking Reports (638 & 654-667)	Ying-Ren Chien	Jihui Aimee Zhang
	16:00	16:30	Break		
	16:30	18:00			

Date	Start	End	Peony II	Chair	Co-Chair
	8:00	10:00			
	10:00	10:30	Conference Subcomm Meeting		
	10:30	11:30			
22	11:30	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Wireless Communications & Networking	Kouji Hirata	Shu-Yu Lin
Wed	16:00	16:30	Break		
	16:30	18:00	Advanced Multimedia Applications	Zhi Hu	Dayan Perera
	18:00	19:00			
	19:00	21:30			
	8:30	10:00	Biomedical Signal Processing and Systems I	Jetsada Arnin	Wen Zhang
	10:00	10:30	Break		
	10:30	12:00	Machine Learning: Algorithms and Application I	Jia-Ching Wang	Guo-Shiang Lin
23	12:00	13:30			
Oct	13:30	14:30			
2025	14:30	16:00	Signal & Information Processing I	Dipanita Chakraborty	S. Nimishan
Thu	16:00	16:30	Break		
	16:30	18:00	Advances in Multimodal AI for Multimedia Applications	Phat Nguyen	Chen-Kuo Chiang
	18:00	19:00			
	19:00	21:30			
	8:30	10:00	Machine Learning: Algorithms and Application II	Junya Hara	Yuichi Tanaka
	10:00	10:30	Break		
24	10:30	12:00	Machine Learning: Algorithms and Application III	Koichi Shinoda	Chih-Yi Chiu
Oct	12:00	13:30			
2025	13:30	14:30			
Fri	14:30	16:00	Scalable and Efficient Signal Processing for Multimodal AI Systems (ESPRESSO)	Yang Xiao	Jisheng Bai
	16:00	16:30	Break		
	16:30	18:00			

Floorplan for Island Ballroom



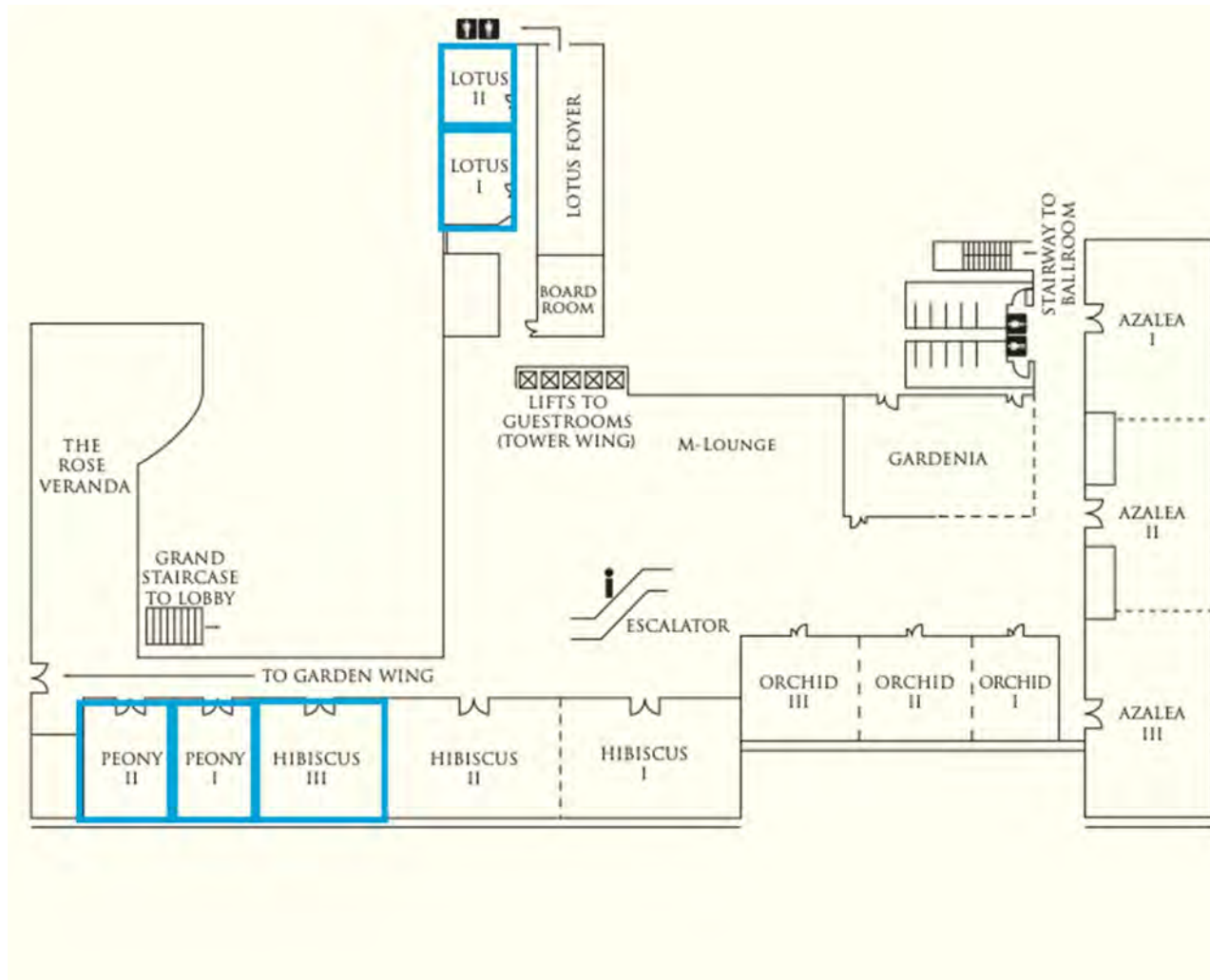
Registration: Foyer of Island Ballroom

Plenary Sessions: Island Ballroom

Banquet: Island Ballroom

Exhibition: Katong

Floorplan for Mezzanine Floor



Parallel Sessions:

- Lotus I
- Lotus II
- Hibiscus III
- Peony I
- Peony II