MIREX 2011 Evaluation Results

J. Stephen Downie & IMIRSEL
University of Illinois at Urbana-Champaign
email: jdownie@illinois.edu

Audio Music Similarity

SubID	Participants	Avg. Fine Score
SSPK2	Seyerlehner, Schedl, Knees, Pohle	58.64
CTCP2	Charbuillet, Peeters, Cornu, Tardieu	58.59
SSKS3	Seyerlehner, Schedl, Knees, Sonnleitner	58.13
PS1	Schnitzer, Pohle	57.70
CTCP1	Charbuillet, Peeters, Cornu, Tardieu	57.27
CTCP3	Charbuillet, Peeters, Cornu, Tardieu	56.21
DM2	Leon, Martinez	50.49
DM3	Leon, Martinez	50.35
ZYC2	Zhang, Yang, Chen	50.04
ML1	McFee, Lanckrie	47.78
ML3	McFee, Lanckrie	47.78
ML2	McFee, Lanckrie	47.30
YL1	Lin, Cheng, Wu	42.37
HKHLL1	Han, Lee, Hyung, Kim, Lee	42.19
STBD1	Sammartino, Bandera, Tardón, Barbancho	33.91
GKC1	Gkiokas, Katsouros, Carayannis	31.84
STBD2	Sammartino, Bandera, Tardón, Barbancho	30.56
STBD3	Sammartino, Bandera, Tardón, Barbancho	30.39

Audio Onset Detection

SubID	Participants	Avg. F-measure
SB2	Böck	0.83
AR3	Roebel	0.81
AR4	Roebel	0.80
SB1	Böck	0.80
AR5	Roebel	0.80
AR1	Roebel	0.80
AR2	Roebel	0.80
BT1	Thoshkahna	0.73

Audio Chord Estimation

SubID	Participants	Weighted overlap ratio
NMSD2	Ni, Mcvicar, Santos-Rodriguez, De Bie	0.97
NMSD3	Ni, Mcvicar, Santos-Rodriguez, De Bie	0.82
KO1	Khadkevich, Omologo	0.82
NM1	Ni, Mcvicar	0.81
CB2	Cho, Bello	0.80
CB3	Cho, Bello	0.80
NMSD1	Ni, Mcvicar, Santos-Rodriguez, De Bie	0.78
KO2	Khadkevich, Omologo	0.78
CB1	Cho, Bello	0.78
UUOS1	Ueda, Uchiyama, Ono, Sagayama	0.76
PVM1	Pauwels, Varewyck, Martens	0.73
RHRC1	Rocher, Hanna, Robine, Conklin	0.72
UUROS1	Balazs, Ueda, Uchiyama, Raczynski, Ono, Sagayama(5)	0.34
BUURO3	Balazs, Ueda, Uchiyama, Raczynski, Ono, Sagayama	0.34
BUURO1	Balazs, Ueda, Uchiyama, Raczynski, Ono, Sagayama	0.23
BUURO4	Balazs, Ueda, Uchiyama, Raczynski, Ono, Sagayama	0.19
BUURO2	Balazs, Ueda, Uchiyama, Raczynski, Ono, Sagayama	0.16
BUURO5	Balazs, Ueda, Uchiyama, Raczynski, Ono, Sagayama	0.12

Structural Segmentation - MIREX '09

SubID	Participants	FP clustering F-measure
MHRAF3	Martin, Hanna, Robine, Allali, Ferraro	0.55
MHRAF2	Martin, Hanna, Robine, Allali, Ferraro	0.55
MHRAF1	Martin, Hanna, Robine, Allali, Ferraro	0.54
CL1	Chen, Li	0.54
GP6	Peeters, Cornu	0.50
SBVRS1	Sargent, Bimbot, Vincent, Raczynski, Sagayama	0.48

Structural Segmentation - MIREX '10

SubID	Participants	SB Eval Measure @ 3sec
MHRAF2	Martin, Hanna, Robine, Allali, Ferraro	0.56
SBVRS1	Sargent, Bimbot, Vincent, Raczynski, Sagayama	0.55
GP6	Peeters, Cornu	0.53
MHRAF1	Martin, Hanna, Robine, Allali, Ferraro	0.53
MHRAF3	Martin, Hanna, Robine, Allali, Ferraro	0.49
CL1	Chen, Li	0.43

Symbolic Music Similarity

SubID	Participants	Avg. Fine Score (0-1)
UL1	Urbano, Lloréns, Morato	0.59
UL2	Urbano, Lloréns, Morato	0.57
UL3	Urbano, Lloréns, Morato	0.55
WK1	Wolkowicz, Keselj	0.52
LJY2	Lee, Jo, Yoo	0.49
WK2	Wolkowicz, Keselj	0.49
LJY1	Lee, Jo, Yoo	0.48
WK5	Wolkowicz, Keselj	0.47
WK4	Wolkowicz, Keselj	0.46
WK6	Wolkowicz, Keselj	0.46
WK3	Wolkowicz, Keselj	0.43

Real-time Audio to Score Alignment

SubID	Participants	Total Precision
SUROS1	Suzuki, Ueda, Raczynski, Ono, Sagayama	67.11%
JC1	Chen, Jang	64.90%

Multi-F0 Estimation

SubID	Participants	Accuracy
YR2	Yeh, Roebel	0.68
YR4	Yeh, Roebel	0.68
YR1	Yeh, Roebel	0.66
YR3	Yeh, Roebel	0.65
KD1	Dressler	0.63
BD1	Benetos, Dixon	0.57
RFF1	Reis, Fernandéz, Ferreira	0.49
RFF2	Reis, Fernandéz, Ferreira	0.48
LYC1	Lee, Yang, Chen	0.47

Audio Tempo Estimation

SubID	Participants	P-Score
GKC3	Gkiokas, Katsouros, Carayannis	0.83
FW2	Wu	0.74
ZG1	Zapata, Gómez	0.73
SP1	Pauws	0.71
GKC6	Gkiokas, Katsouros, Carayannis	0.68
SB5	Böck	0.66

Audio Key detection

SubID	Participants	Weighted Key Score
GP1	Peeters	0.82
DTBS2	Bandera, Tardón, Barbancho, Sammartino	0.82
PVM2	Pauwels, Varewyck, Martens	0.82
DTBS1	Bandera, Tardón, Barbancho, Sammartino	0.81
UUOS2	Ueda, Uchiyama, Ono, Sagayama	0.76
KO3	Khadkevich, Omologo	0.65
KO4	Khadkevich, Omologo	0.60
RHR1	Rocher, Hanna, Robine	0.39

Audio Cover Song Identification

SubID	Participants	Total Precision: Mixed	Total Precision: Mazurkas
ALL1	Ahonen, Lemström, Linkola	0.13	0.70
CWWJ1	Chuan-Yau Chan, Ju-Chiang Wang, Hsin-Min Wang,Shyh-Kang Jeng	0.21	0.88

Audio Melody Extraction

SubID	Participants	Overall Accuracy: MIREX'05	ADC'04	INDIAN'08
SG2	Salamon, Gómez	0.68	0.74	0.84
SG1	Salamon, Gómez	0.66	0.74	0.83
PJY1	Park, Jo, Yoo	0.65	0.81	0.71
YSLP1	Yoon, Song, Lee, Park	0.65	0.85	0.73
LYRS1	Liao, YEH, Roebel, Su	0.59	0.73	0.72
CWJ1	Chien, Wang, Jeng	0.57	0.73	0.69
TOS1	Tachibana, Ono, Ono, Sagayama	0.57	0.59	0.72
TY4	Yeh	0.51	0.47	0.70
TY3	Yeh	0.51	0.47	0.70
HCCPH1	Hsueh, Coover, Chen, Popp, Han, Pardo	0.45	0.44	0.64

Audio Melody Extraction

SubID	Participants	Overall Accuracy: MIREX'09 0db	MIREX'09 +5db	MIREX'09 -5db
SG1	Salamon, Gómez	0.78	0.85	0.61
SG2	Salamon, Gómez	0.78	0.85	0.61
TOS1	Tachibana, Ono, Ono, Sagayama	0.74	0.82	0.62
PJY1	Park, Jo, Yoo	0.74	0.83	0.54
CWJ1	Chien, Wang, Jeng	0.53	0.62	0.40
TY3	Yeh	0.52	0.56	0.41
TY4	Yeh	0.52	0.56	0.41
YSLP1	Yoon, Song, Lee, Park	0.52	0.66	0.39
HCCPH1	Hsueh, Coover, Chen, Popp, Han, Pardo	0.50	0.59	0.39
LYRS1	Liao, YEH, Roebel, Su	0.47	0.54	0.36

Query By Tapping

SubID	Participants	Task 1A MRR	Task1B MRR	Task 2 MRR
CJ1	Chen, Jang	0.60	0.51	0.59

ILLINOIS UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
mirex
imirsel

The Andrew W. Mellon Foundation
nema

MIREX 2011 Evaluation Results

J. Stephen Downie & IMIRSEL
University of Illinois at Urbana-Champaign
email: jdownie@illinois.edu

Query By Singing/Humming

SubID	Participants	Task1A MRR	Task1B MRR	Task2 MRR
TY1	Yeh	0.93	0.44	8.74
JSSLP1	Jang, PARK, Song, Shin, JANG, Lee, Lee, Seo	0.90	0.91	9.28
TY2	Yeh	0.88	0.85	8.74

Audio Beat Tracking

		F-Measure	
SubID	Participants	MCK	MAZ
SB3	Böck	52.69	40.29
SB4	Böck	50.86	51.17
KFRO1	Khadkevich, Fillon, Richard, Omologo	50.67	29.27
KFRO2	Khadkevich, Fillon, Richard, Omologo	50.45	35.04
GP5	Peeters	50.32	47.02
GKC2	Gkiokas, Katsouros, Carayannis	50.10	42.18
GP4	Peeters	50.09	49.12
GP3	Peeters	49.56	40.16
GP2	Peeters	49.29	41.80
GKC5	Gkiokas, Katsouros, Carayannis	48.54	37.31
FW1	Wu	47.84	67.56
SP2	Pauws	43.53	31.03

Multi-F0 Note Tracking - Mixed Dataset

SubID	Participants	Avg. F-measure Onset Only	Avg. Overlap
YR1	Yeh, Roebel	0.56	0.89
YR3	Yeh, Roebel	0.55	0.89
BD3	Benetos, Dixon	0.45	0.85
BD2	Benetos, Dixon	0.45	0.86
RFF1	Reis, Fernandéz, Ferreira	0.41	0.86
LYC1	Lee, Yang, Chen	0.39	0.83
RFF2	Reis, Fernandéz, Ferreira	0.36	0.86

Multi-F0 Note Tracking - Piano Only

SubID	Participants	Avg. F-measure Onset Only	Avg. Overlap
YR1	Yeh, Roebel	0.61	0.82
BD3	Benetos, Dixon	0.59	0.82
YR3	Yeh, Roebel	0.59	0.82
BD2	Benetos, Dixon	0.53	0.82
LYC1	Lee, Yang, Chen	0.53	0.79
RFF1	Reis, Fernandéz, Ferreira	0.52	0.80
RFF2	Reis, Fernandéz, Ferreira	0.44	0.79

Audio Tag Classification - Major Minor

SubID	Participants	Class. F-Measure	Affinity AUC-ROC
PH2	Hamel	0.56	0.91
TCCP2	Tardieu, Charbuillet, Cornu, Peeters	0.50	0.79
SSKS1	Seyerlehner, Schedl, Knees, Sonnleitner	0.49	0.89
SBC1	Sordo, Bogdanov, Celma	0.48	0.87
TCCP1	Tardieu, Charbuillet, Cornu, Peeters	0.48	0.79
SC1	Sordo, Celma	0.48	0.87
BA2	Bourguigne, Agüero	0.47	0.78
BA1	Bourguigne, Agüero	0.47	0.78
BA3	Bourguigne, Agüero	0.47	0.78
JR6	Ren, Chang	0.46	0.83
CLCB1	Coviello, Lanckriet, Chan, Barrington	0.37	0.81
JR4	Ren	0.36	0.85
JR5	Ren	0.35	0.85
ECL1	Ellis, Coviello, Lanckriet	0.34	0.80
CCL1	Lanckriet, Coviello, Ellis	0.34	0.79

Audio Tag Classification - Mood

SubID	Participants	Class. F-Measure	Affinity AUC-ROC
PH2	Hamel	0.51	0.87
SSKS1	Seyerlehner, Schedl, Knees, Sonnleitner	0.49	0.87
TCCP2	Tardieu, Charbuillet, Cornu, Peeters	0.48	0.87
SBC1	Sordo, Bogdanov, Celma	0.47	0.84
SC1	Sordo, Celma	0.47	0.85
TCCP1	Tardieu, Charbuillet, Cornu, Peeters	0.47	0.86
BA2	Bourguigne, Agüero	0.40	0.78
BA1	Bourguigne, Agüero	0.39	0.79
BA3	Bourguigne, Agüero	0.39	0.79
JR6	Ren, Chang	0.37	0.80
JR5	Ren	0.34	0.84
JR4	Ren	0.28	0.84
CLCB1	Coviello, Lanckriet, Chan, Barrington	0.27	0.74
ECL1	Ellis, Coviello, Lanckriet	0.26	0.69
CCL1	Lanckriet, Coviello, Ellis	0.26	0.69

Special Thanks to: The Andrew W. Mellon Foundation, the National Science Foundation (Grant No. NSF IIS-0327371 and No. NSF IIS-0328471), IMIRSEL members (M. Bay, K. Choi, J. Downie, A. Ehmann, & G. Zhu), X. Hu, K. West, J. Jones, the content providers, Evalutron graders, the MIR community, and the ISMIR 2011 organizing committee