Datasets

Environmental sounds

Introduction

Audio data collection and manual data annotation both are tedious processes, and lack of proper development dataset limits fast development in the environmental audio research.

Currently, there are only handful of large datasets available and some of them might be hard to find (e.g. they are used in neighboring research fields). This page tries to maintain a list of datasets suitable for environmental audio research. In addition to the freely available dataset, also proprietary and commercial datasets are listed here for completeness.

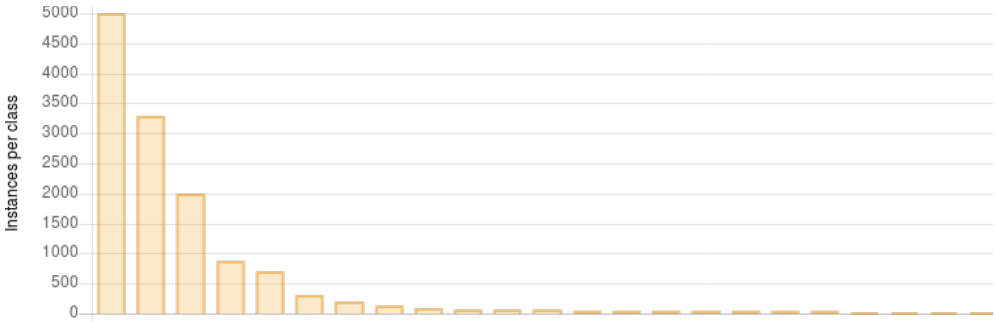
In addition to the datasets, also some of the on-line sound services are listed at the end of the page. These services can be used to form new datasets for special research needs.

If you know datasets which should be in this page or you have knowledge (statistics) about some of the commercial dataset, please let me know (mmailto:toni.heittola@tut.fi).

Environmental audio

The datasets are divided into two tables: *Sound events* table contains datasets suitable for research in the field of automatic sound event detection and automatic sound tagging. *Acoustic scenes* table contains datasets suitable for research involving the audio-based context recognition and acoustic scene classification.

Sound events



Bar Off

Provider	Name	Context selection	Annotation type	Event classes	Event count	Instances per class	Link
NYU	URBAN-SED	Mixed	Events	10	50000	5000.0	(http
Google	AudioSet	Mixed	Tags	632	2084320	3297.9	(http: /audios

Provider	Name	Context selection	Annotation type	Event classes	Event count	Instances per class	Link
MIVIA	Audio Events Data Set for Surveillance Applications	Single	Events	3	6000	2000.0	(http analysis
NYU	UrbanSound8K	Street	Tags	10	8732	873.2	(http /urbans /urbans
NII-SRC	RWCP Sound Scene Database	Mixed	Tags	14	9722	694.4	(http
NYU	UrbanSound	Single	Events	10	3075	307.5	(http /urbans
MIVIA	Audio Events Data Set for Road Surveillance Applications	Single	Events	2	400	200.0	(http analysis set/)
TUT	TUT Sound events 2017, Development dataset	Mixed	Events	6	729	121.5	(https
Dares	Amstel	Single	Events	13	1002	77.1	(http
ELRA	UPC-TALP database of isolated meeting-room acoustic events	Single	Events	14	1026	73.3	(http /produc
CICESE	Sound Events	Single	Tags	20	1367	68.3	(http /allSour
Freiburg	Freiburg-106, Audio Data Set for Human Activity Recognition	Single	Tags	24	1524	63.5	(http /pages/
TUT	TUT Sound events 2016, Development dataset	Mixed	Events	18	954	53.0	(https
TUT	CASA 2009	Mixed	Events	208	10326	49.6	

Provider	Name	Context selection	Annotation type	Event classes	Event count	Instances per class	Link
ESC	ESC-50	Mixed	Tags	50	2000	40.0	(https
ESC	ESC-10	Mixed	Tags	10	400	40.0	(https
ESC	Dataset for Environmental Sound Classification	Mixed	Tags	50	2000	40.0	(https /dataset /DVN/Y
IEEE AASP Challenge 2013	Event isolated	Single	Tags	16	639	39.9	(http /scenes /descrip
ELRA	FBK-Irst database of isolated meeting-room acoustic events	Single	Tags	16	576	36.0	(http /produc
TU Dortmund	Multi-channel acoustic event dataset	Single	Events	20	437	21.9	(http /dataset aed_do
IEEE AASP Challenge 2013	Event synthetic	Single	Events	15	310	20.7	(http /scenes /descrip
TU Dortmund	Acoustic event dataset	Single	Events	12	235	19.6	(http /icassp2
TUT	CASA 2010	Mixed	Events	289	4173	14.4	
IEEE AASP Challenge 2013	Event live	Single	Events	16	205	12.8	(http /scenes /descrip
Dares	G1	Mixed	Events	761	3214	4.2	(http
INRIA	NAR	Mixed	Tags	41	42	1.0	(https
Sound Ideas	BBC Sound Effects Library	Mixed	Description	0	1655		(http effects/l library.h
ELRA	CHIL 2007 Evaluation Package	Single	Events	0	0		(http /produc

Provider	Name	Context selection	Annotation type	Event classes	Event count	Instances per class	Link
QMUL	Freefield1010	Mixed	Tags	0	7690		(http /rdr/han
TUT	TUT-SED Synthetic 2016	Single	Events	16			(http /taslp20 syntheti
TUT	TUT Rare sound events 2017, Development dataset	Single	Events	3			(http /dcase2 2---dete

Instructions to use the table

Sort the table by clicking headers. Select more fields to be shown from -menu. The table can be filtered by selection condition from the select boxes. Filtering can be cleared with -button. Numerical data is also shown in bar-chart. The chart can be hiden and shown from Chart-button. Pagination can be enabled and disabled with -button.

Notation

Audio type is denoted as follows:

- **Isolated**, only one sound event is active per sample (no overlapping sounds).
- **Live**, real recordings where overlapping sounds may be presents, start and end times annotated.

Annotation type is denoted as follows:

- **Sound Events**, timestamps (onset and offset times) are indicated
- **Tags**, sample-wide (usually one word or short) textual label
- **Description**, sample-wide textual description of sounds / sound scene

Acoustic scenes

Bar Off

Provider	Name	Context count	Audio files	Instances per class	Mins	Link

Online services

Isolated sounds

- Freesound (http://www.freesound.org/), isolated sounds, tagged, creative commons
- Findsounds (http://www.findsounds.com/), isolated sounds, tagged, mixed licensing
- British Library Sound Archive (http://sounds.bl.uk/), isolated sounds and Live recordings, only available for UK universities, restricted licensing

Geotagged recordings

- Radio Aporee (http://aporee.org/maps/)
- Sound Around you (http://www.soundaroundyou.com/)
- Urban remix (http://urbanremix.gatech.edu/)

Related datasets

- Speech Datasets (https://wiki.inria.fr/rosp/Datasets), ISCA Special Interest Group on Robust Speech Recognition

Libraries

Free sound effect libraries by commercial provider:

- The Free Firearm Sound Library (https://www.stillnorthmedia.com/libraries/)
- Free Sound Effects by Airborne sound (http://www.airbornesound.com/sound-effects-library/free-sound-effects/), see end-user license (http://www.airbornesound.com/user-licence/).

Tools

Annotation tools

Software

- Audacity (http://www.audacityteam.org/), audio software with basic annotation capabilities. Use label tracks for the annotations, see more info here (http://manual.audacityteam.org/man/label_tracks.html).
- Audio Annotator (https://github.com/CrowdCurio/audio-annotator), Javascript web interface for annotating audio data.
- ELAN (https://tla.mpi.nl/tools/tla-tools/elan/), a linguistic annotation tool to create the textual annotations for audio and video files

Prototypes

- Soundscape (http://www.ai.rug.nl/~vdlinden/annotationtool/index.html), a tool for soundscape annotation
- I-SED (http://www.bongjunkim.com/ised/), an interactive sound event detector, see [Kim2017 (http://music.cs.northwestern.edu/publications/Kim_Pardo_IUI2017.pdf)]
- BAT (https://github.com/BlaiMelendezCatalan/BAT), BMAT Annotation Tool, see [Melendez-Catalan2017 (http://eecs.qmul.ac.uk/~keno/17.pdf)]
- audio-annotator (https://github.com/CrowdCurio/audio-annotator), Audio-annotator, see [Cartwright2017 (http://faculty.poly.edu/~onov/Cartwright_et_al_SONYC_CSCW_2017.pdf)]

Audio management tools

- Pumilio (http://ljvillanueva.github.io/pumilio/), a Web-Based Management System for Ecological Recordings

Audio augmentation tools

- muda (https://github.com/bmcfee/muda), Annotation-aware musical data augmentation, partly applicable for environmental audio (pitch shifting, time stretching). Documentation (https://muda.readthedocs.io/en/latest/)
- librosa (https://github.com/librosa/librosa/), See time stretching (https://librosa.github.io/librosa/generated/librosa.effects.time_stretch.html) and pitch shifting (https://librosa.github.io/librosa/generated/librosa.effects.pitch_shift.html) effects.
- TSM toolbox (https://www.audiolabs-erlangen.de/resources/MIR/TSMtoolbox/), MATLAB implementations of various classical time-scale modification (TSM) algorithm.