Articulograph AG500 - data format

contents

data flow

Each sensor picks up a signal that is induced by the magnetic field from the 6 transmitters arranged at the EMA-Cube.

The resulting voltage from each sensor is sampled 100 000 times per second. It is the input data for the “senrec” program in the PSR-Computer. The signals arrives as 16 Bit digital words. All 12 channels are read simultaneously. This data stream (2,4Mbyte/s) is called raw data.

- raw data = 100kHz sensor signal from 12 channels.

The “senrec” program demodulates this data stream and yields amplitudes from the 6 transmitter signals for each sensor with 200Hz time resolution.

The field strength is calculated as complex amplitudes (real and imag. component) .

- sample = 6 complex amplitudes for each sensor (200Hz signal).

The samples are sent to the “AGMaster” program over TCP/IP connection. Only for display, the program calculates the value for the actual field strength and shows it.

During sweep recording, this data stream is stored to disk without any modification into one file. The speech movement data appears on the disk as a sequence of samples. Other parameter are stored in an additional header file.

- sweep = data file with a sequence of sample + header file

The data for one sweep are stored in two files (*.KOF and *.hdr). A sweep contains the data to calculate the position and orientation for all sensors 200 times per second.

For calibration, it is necessary to know the sensor positions and orientations at fix points. Therefore a sweep is recorded at each position. For each sensor, the mean values are calculated out of all the complex amplitudes and stored in text files inside the folder “ausgewertet”.

- ausgewertet = folder with calibration data

The result of the calibration procedure ist stored in a copy from the ag500.ini file inside the folder for each sensor (channel = Kanal in German). This file also holds the actual recording parameter.

- ag500.ini = data recording parameters and calibration results.

Speech movement data – sweep

A sweep is the continuous recording of speech movement data measured by the sensors. Stored to disk, the Information for each sweep are stored in two files: A binary data file with the extension *.KOF and a header file in ASCII format.

binary data file

Inside the resulting data file a sweep appears as a sequence of samples.

A sample consists of 144 words (each 8 byte) in the following order:

word no.	complex amplitudes	transmitter	channel
0	real	Ku_1	1
1	real	Nl_2	1
2	real	Sl_3	1
3	real	Ko_4	1
4	real	Sr_5	1
5	real	Nr_6	1
6	real	Ku_1	2
7	real	Nl_2	2
8	real	Sl_3	2
9	real	Ko_4	2
10	real	Sr_5	2
11	real	Nr_6	2
12	real	Ku_1	3
13	real	Nl_2	3
..	..	..	..
..	..	..	..
70	real	Sr_5	12
71	real	Nr_6	12
72	imaginary	Ku_1	1
73	imaginary	Nl_2	1
..	..	..	..
..	..	..	..
142	imaginary	Sr_5	12
143	imaginary	Nr_6	12

The second sample will start as the 144th word of the binary file. There is no separator between samples.

header file

The header file contains one parameter in each line, starting with the parameter name.

a) comment=

In the first line, the AGMaster program inserts one of the following comments here if it is a spezial sweep:

calibration sweep:

comment=golist entry <move number>

angle adjust sweep:

comment=angle adjust sweep

measuring complex offset:

comment=complex offset sweep 1

b) complex offset and phase shift = angle offset

The following lines are always one group of three lines for each sensor – transmitter combination.

The sensor – transmitter combination is coded in the name of the parameter. The format is:

<parameter type>_<sensor/channel no>_<transmitter no>

For example, the last three lines contain the parameter for sensor/channel number 12 in compination with transmitter 6 (Nr6):

Complex_Cos_12_6= -16.75

Complex_Sin_12_6= 4.84

AngleOfs_12_6= 3.12

c) Complex_Cos_12_6=

The offset value that has to be subtracted from the real amplitudes. In this example it has to be subtracted from the word no. 137 in each sample

d) Complex_Sin_12_6=

The offset value that has to be subtracted from the imaginary amplitudes. In this example it has to be subtracted from the word no. 143 in each sample

e) AngleOfs_12_6=

To define the sign of the sensor signals, the resulting phase angle is corrected to zero by adding the according offset angle.

calculating the sensor signal

For each sensor – transmitter combination the measured signal is calculated out of the stored values in the binary data file corrected by the complex offset.

With the above example (sensor 12 and transmitter 6), the following procedure is necessary to get the sensor's signal measured from transmitter 6:

CosValue:= Cos_coefficient - Complex_Cos_12_6

SinValue:= Sin_coefficient - Complex_Sin_12_6

Where the Cos_coefficient is the word no.137 and Sin_coefficient is word no. 143 in each sample.

Amplitude := sqrt(sqr(CosValue) + sqr(SinValue));

Phase angle := ArcTan2(SinWert, CosWert) + AngleOfs;

Raw data

analog to digital converter values

The AD converter data are sampled and processed with 100kHz all the time. For basic system tests, the raw data can also be stored to disk.

According to the selected “Number of blocks” and “Blocksize in kS/C” (kSample/Channel), the data will be stored in the “rawdata” folder in the file ”raw0000.bin” (the next in ”raw0001.bin”).

The raw data file is a sequence of 16 Bit Integer Values – starting with sensor 1.

integer word no.	channel/sensor
0	1
1	2
2	3
3	4
4	5
5	6
6	7
7	8
8	9
9	10
10	11
11	12
12	1
13	2
..	..
..	..

reference values

The AGMaster Rev. 6.07 (and later) also stores also the reference information for demodulating. This file has the same filename as the raw data file with the extension *.ref.

The reference information is valid for all channels. There are 12 words (each 8 byte) belonging to one 100kHz ADC measuring in the following order:

word no.	reference	transmitter	time [µs]
0	cosine	Ku_1	0
1	cosine	Nl_2	0
2	cosine	Sl_3	0
3	cosine	Ko_4	0
4	cosine	Sr_5	0
5	cosine	Nr_6	0
6	sine	Ku_1	0
7	sine	Nl_2	0
8	sine	Sl_3	0
9	sine	Ko_4	0
10	sine	Sr_5	0
11	sine	Nr_6	0
12	cosine	Ku_1	10
13	cosine	Nl_2	10
..	..	..	..
..	..	..	..

Calibration data – preliminary

The calibration is done with the AutoKal. At each position a sweep is recorded and the mean amplitudes (see: calculating the sensor signal) is calculated. The result for each position is stored together with the position information in an ASCII file named “result.txt”.

A line in the “result.txt” has the following format:

Ku_1	Nl_2	Sl_3	Ko_4	Sr_5	Nr_6	x	y	z
3067.96	3217.20	1413.32	1762.78	1381.37	603.62	20.7800	50.9100	-29.3800

For detailed information of the ag500.ini file please refer to:

http://www.phonetik.uni-muenchen.de/~andi/EMAPage/