

# Operations manual of GritTec's Speaker-ID: The Mobile Client

## Version 1.90

2011

<b>Title</b>	Operations manual of GritTec's Speaker-ID: The Mobile Client
<b>Short name of product</b>	GritTec's Speaker-ID: The Mobile Client
<b>Version</b>	1.90
<b>Document Number</b>	Manual_GritTec's Speaker-ID_The mobile Client_ENG.doc
<b>Abstract</b>	The software solution of GritTec Speaker-ID: The mobile client is intended for automatic voice identification of a speech signal of unknown voice by comparing with target voice samples of "speaker cards", existing in the database of system.

## Copyright Information

© 2011 GritTec Ltd. All rights reserved.

This document is protected by copyright. No part of this document may be reproduced in any form by any means without prior written authorization of GritTec Ltd.

All brand names and product names mentioned in this document are trademarks or registered trademarks of their respective owners.

### Contact Address:

**Address:** 91 Bld 2, Svobody str., Moscow, 125481, Russia Federation

**Phone:** +7 (495) 796 24 18

**E-mail:** [info@grittec.com](mailto:info@grittec.com)

**Home Page:** <http://www.grittec.com>

## Revision history

Date	Version	Author	Amendment
2011.01.25	1.60		First edition
	1.80		Fixed problem with resetting of Database by default.
2011.03.24	1.90		Updating have been made at the expense of upgrading in a new version of GritTec's Speaker-ID (Ver.3,00) engine in which integrated of noise cancellation engine and was made additional tuning of the noise compensation module. Added possibility of rebuilding of all target voice models for all speaker cards.

## CONTENTS

1. DEFINITIONS AND ABBREVIATIONS .....	5
2. OVERVIEW.....	5
3. SIGNAL REQUIREMENT .....	5
4. SUPPORTED SIGNAL FORMATS.....	6
5. AVAILABILITY.....	6
6. GENERAL FEATURES.....	6
7. ARCHITECTURE OF GRITTEC'S SPEAKER-ID.....	6
8. OPERATIONS MANUAL .....	7
8.1. Database of target speaker cards .....	8
8.2. Rebuilding of all target samples for all speakers cards .....	9
8.3. Create/Change common Database .....	9
8.4. Target speaker cards.....	10
8.5. List of identified Groups.....	12
8.6. The identified parameters.....	13
8.7. The Identification Results .....	14
8.8. Target speaker search (the identification request) .....	15
8.9. Enter the license key .....	16
9. TESTING.....	16
10. REFERENCES.....	16
APPENDIX 1 .....	17

## TABLES

Table 1. Definition and abbreviations.....	5
Table 2. Signal Requirement.....	5

## FIGURES

Fig 1. Scheme of architecture of GritTec's Speaker-ID .....	7
Fig 2. The main panel of GritTec's Speaker-ID: The mobile client .....	8
Fig 3. The panel of target speaker cards Database.....	8
Fig 4. The panel of System Configuration with «Database» section .....	9
Fig 5. Set new directory for Database.....	9
Fig 6. «Database» section with new Database files .....	10
Fig 7. reset directory for Database files .....	10
Fig 8. «Database» section with current Database files.....	10
Fig 9. The panel of preview speaker card.....	11
Fig 10. Add audio file with voice sample into «Voice samples» database .....	11
Fig 11. Structure of Speaker Card .....	12
Fig 12. The panel of list identified groups .....	13
Fig 13. The panel of the identification parameters.....	14
Fig 14. The panel of the identification results .....	15
Fig 15. The panel of speaker search. ....	16

## 1. Definitions and abbreviations

Name	Description
GritTec's Speaker-ID	GritTec's Automatic Text Independent Speaker Identification engine.
Speaker Card	Speaker card besides information about current speaker (first, last name, birthday, gender, and so on) is characterized by examples of audio files with the speaker voice. Each example of audio file is described by the acoustic voice model, error model (FAR, FRR, EER) and noise model, describing surrounding noises and channel distortion, existing in audio file. For the full description of each speaker card it is sufficiently 1 - 3 audio files with the speaker voice, recorded for different telephone lines and duration of each one not less than 60 sec.
FRR	False Rejection Rate
FAR	False Acceptance Rate
ERR	Equal Error Rate (EER = FRR = FAR).

**Table 1. Definition and abbreviations**

## 2. Overview

This document describes the operations manual of GritTec's Speaker-ID: The Mobile Client software. GritTec Speaker-ID: The mobile client (Version 1.60/ Engine 2.990) is intended for automatic voice identification of a speech signal of unknown voice by comparing with voice samples of "speakers cards", existing in the Database of system.

This software product supports the high-quality GUI interface and may be interest for centers of criminalistics, polices, call centers and the banks which purpose is language/text independent voice identification of unknown audio phonogram of telephone negotiations.

Designed algorithm of voice identifications is based on duel comparison spectra features of unknown voice with the spectra features of target voice. Spectra features are calculated with provision of dynamic determinations of channel distortion level and external hindrances and noises.

GritTec's Speaker-ID can be effectively used:

- For automatic voice identification of unknown voice by phonogram of telephone negotiations;
- In systems with high safety level, for instance, when access to digital information is limited by circle of given persons;
- Applications where it's necessary to identify a person using peculiarities of his voice.

## 3. Signal requirement

For application GritTec's Speaker-ID the input signal must be:

Parameters	Value
Signal format	16 bits linear
Sampling rate	8000 Hz
Channels	1 (mono)
Frequency range:	300-3400 Hz or better
Signal to Noise Ratio (SNR), see <a href="#">Exhibit 1</a>	at least 10 db

**Table 2. Signal Requirement**

## 4. Supported signal formats

- |   |
|---|
|   |
| <ul style="list-style-type: none"><li>• Windows PCM,</li><li>• A/mu-Law Wave,</li><li>• Microsoft ADPCM (MS ADPCM),</li><li>• Intel ADPCM (IMA ADPCM),</li><li>• Microsoft ACM GSM 6.10 (ACM Waveform).</li></ul> |

## 5. Availability

MS Windows XP, Vista, Windows 7 under Intel, AMD x86 platform.

## 6. General Features

- |  |
|--|
|  |
| <ul style="list-style-type: none"><li>• Operation with low SNR;</li><li>• Fast adaptation to changing of channel distortion and external noises;</li><li>• Using of noise compensation model and Noise Cancellation engine;</li><li>• Minimum duration of a speech signal with a voice example used for correct reception of voice parameters for the target speaker - not less 15 seconds;</li><li>• Minimum duration of a speech signal with a voice example used for voice identification or voice verification - not less 7 seconds;</li><li>• Speaker identification reliability not less than 90% if both of speech signals were recorded in the same channel;</li><li>• Speaker identification reliability not less then 85% if both of speech signals were recorded in different channels (cross channels);</li><li>• The system supports Databases of Paradox format, with a maximum quantity of records in volume no more than 50 000;</li><li>• Possibility to create and establish a new Databases;</li><li>• Automatic voice identification doesn't require special skills.</li></ul> |

## 7. Architecture of GritTec's Speaker-ID

Architecture of GritTec's Speaker-ID consists of 4 modules:

- The module of registration of audio (acoustic) signals;
- The module of receiving of voice biometric parameters from an audio signal;
- The module of creation of speaker card by received voice biometric parameters;
- The module of automatic speaker identification.

Scheme of architecture of GritTec's Speaker-ID is shown in the figure below.

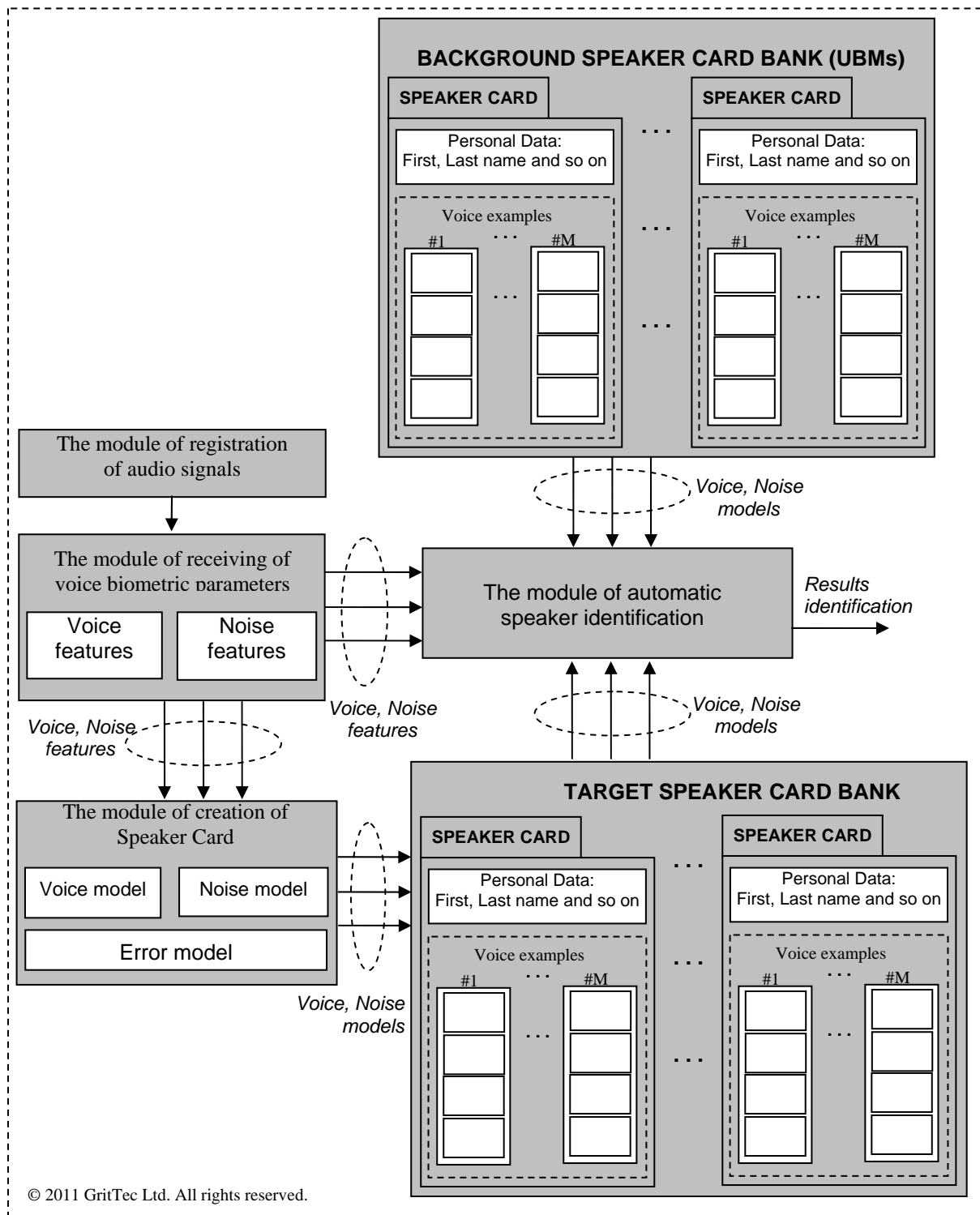
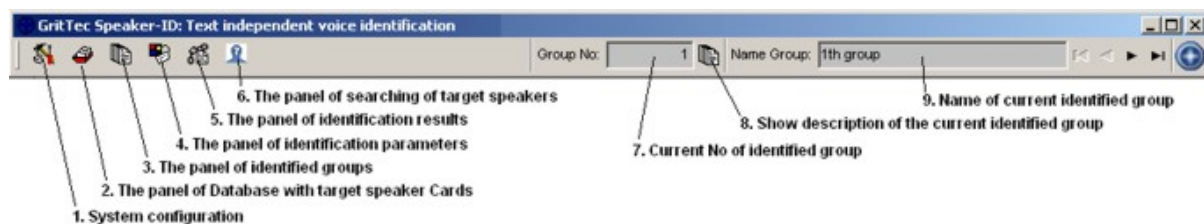


Fig 1. Scheme of architecture of GritTec's Speaker-ID

## 8. Operations manual

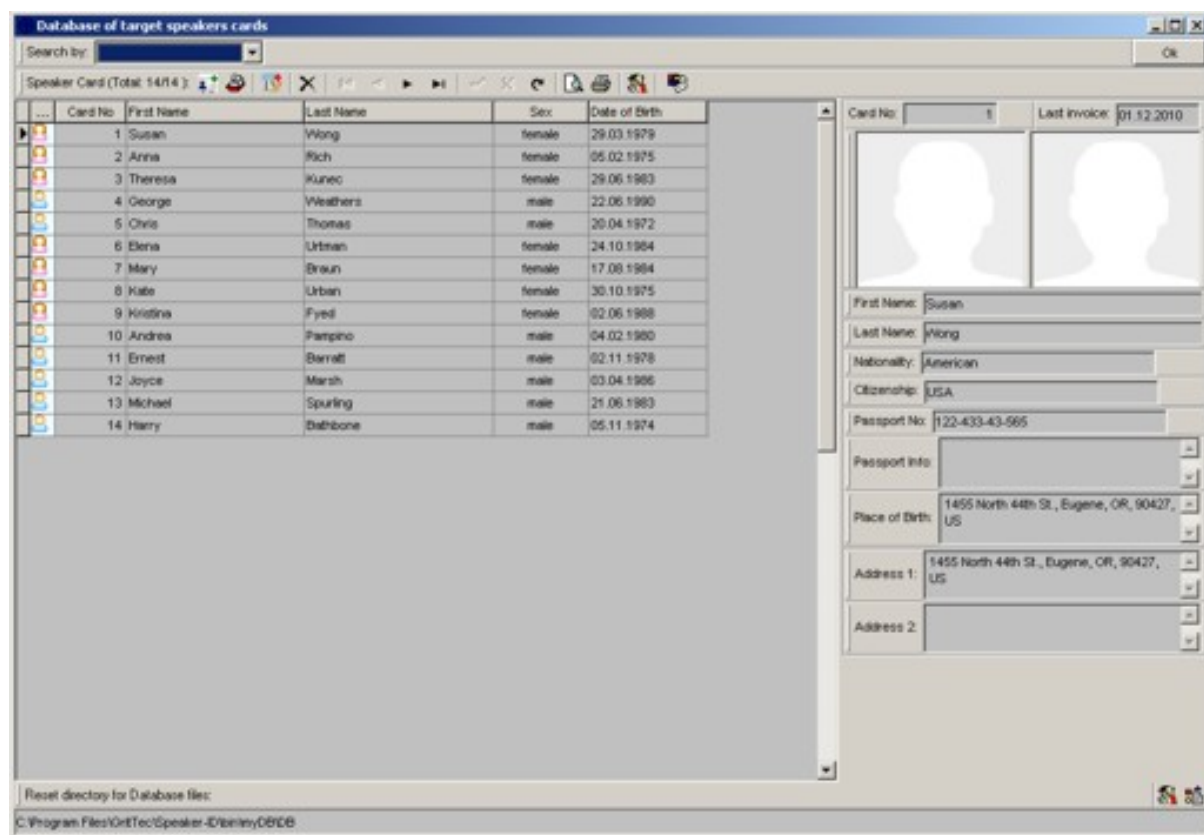
The software consists from the main panel (see Fig 2) which exercises administration of other panels.



**Fig 2. The main panel of GritTec's Speaker-ID: The mobile client**

In current state in the program are installed Database by default. Click "2" button in the main panel of program for preview and edit the current Database of target speakers cards.

### 8.1. Database of target speaker cards



**Fig 3. The panel of target speaker cards Database**

In opened panels of Database of current speakers cards you can see records of target speaker with personal information.

To edit of current record you should press the button «Edit the current speaker card» or you should twice to click a mouse on the selected record. Then the window «Speaker card» will access for changes.

To add of new record of the target speaker you should press the button «Create the new speaker card». Then the window «Speaker card» will open, in which you can write data for the new target speaker and to add audio samples with voice of the current target speaker.

Notice to add new record with target speaker it's possible also at previewing of the current target speaker in the opened window «Speaker card».

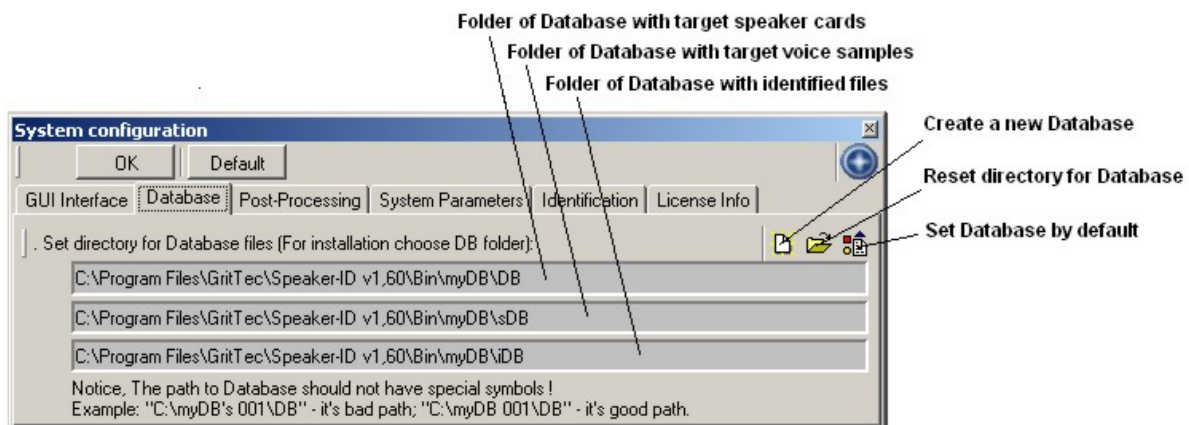
## 8.2. Rebuilding of all target samples for all speakers cards

Rebuilding of all target samples for all speakers cards means updating of voice, error, noise models for each audio sample of all speakers cards at occurrence of release of the new version of GritTec's Speaker-ID engine of voice identification.

The given operation can be made by pressing of button "Rebuild all voice models for all target speakers" in a window "Database of target speakers cards". Also the given operation can be executed in a window "View of speaker card".

## 8.3. Create/Change common Database

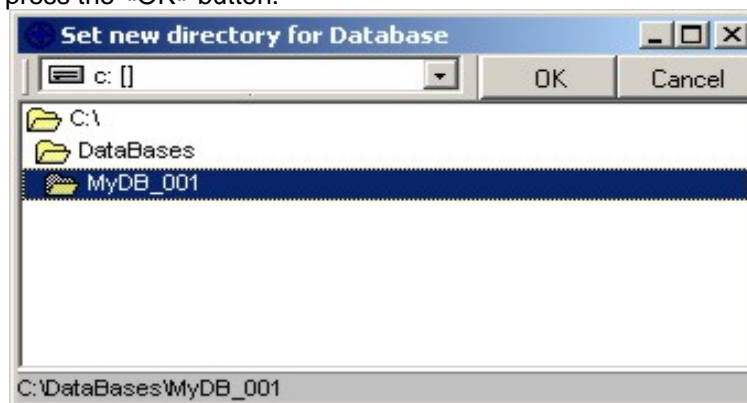
To create a new Database or choose other Database (earlier created) you should open the panel of «System configuration» and «Database» section.



**Fig 4. The panel of System Configuration with «Database» section**

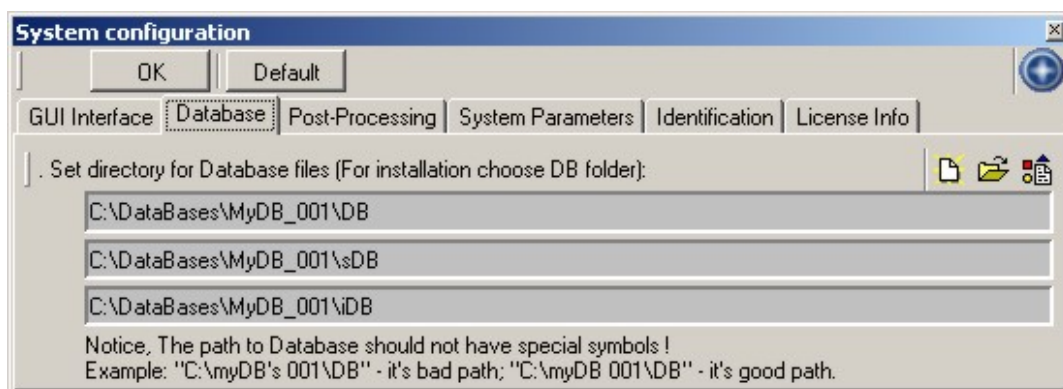
To set Database by default, you should press the button «Set Database by default». In result will be established Database which is stored in the folder «myDB».

To create a new Database you should press the button «Create a new Database», choose or create a folder on the hard disk in which will be stored Database of speakers cards , target voice samples and identified files and press the «OK» button.



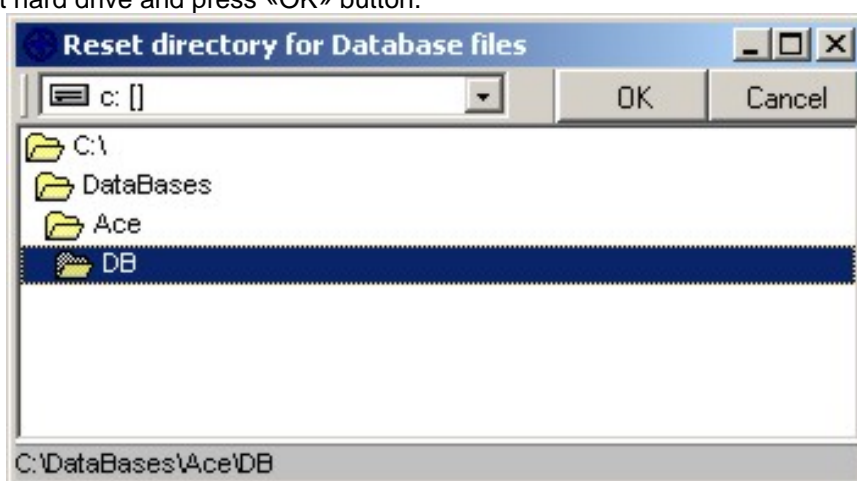
**Fig 5. Set new directory for Database**

Created new Database to be displayed in section «Database» as follows:



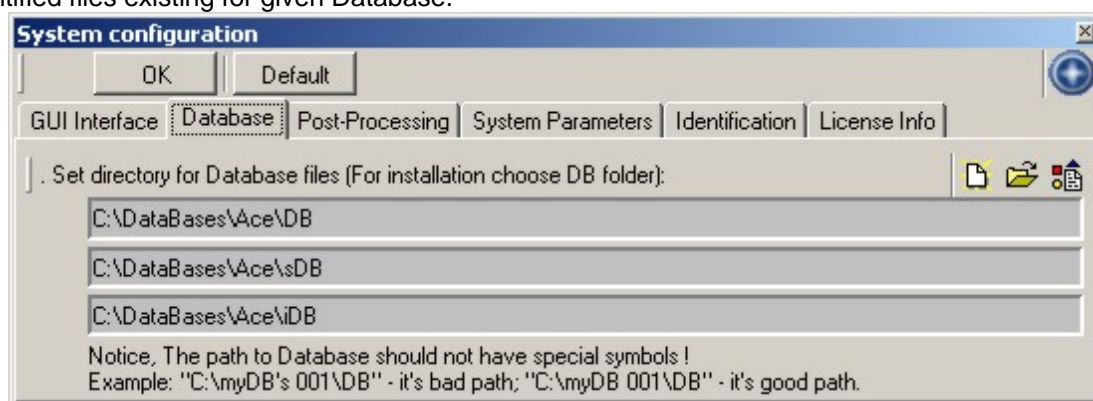
**Fig 6. «Database» section with new Database files**

To open other Database you should press the button «Reset directory for Database files», and select «DB» folder at hard drive and press «OK» button.



**Fig 7. reset directory for Database files**

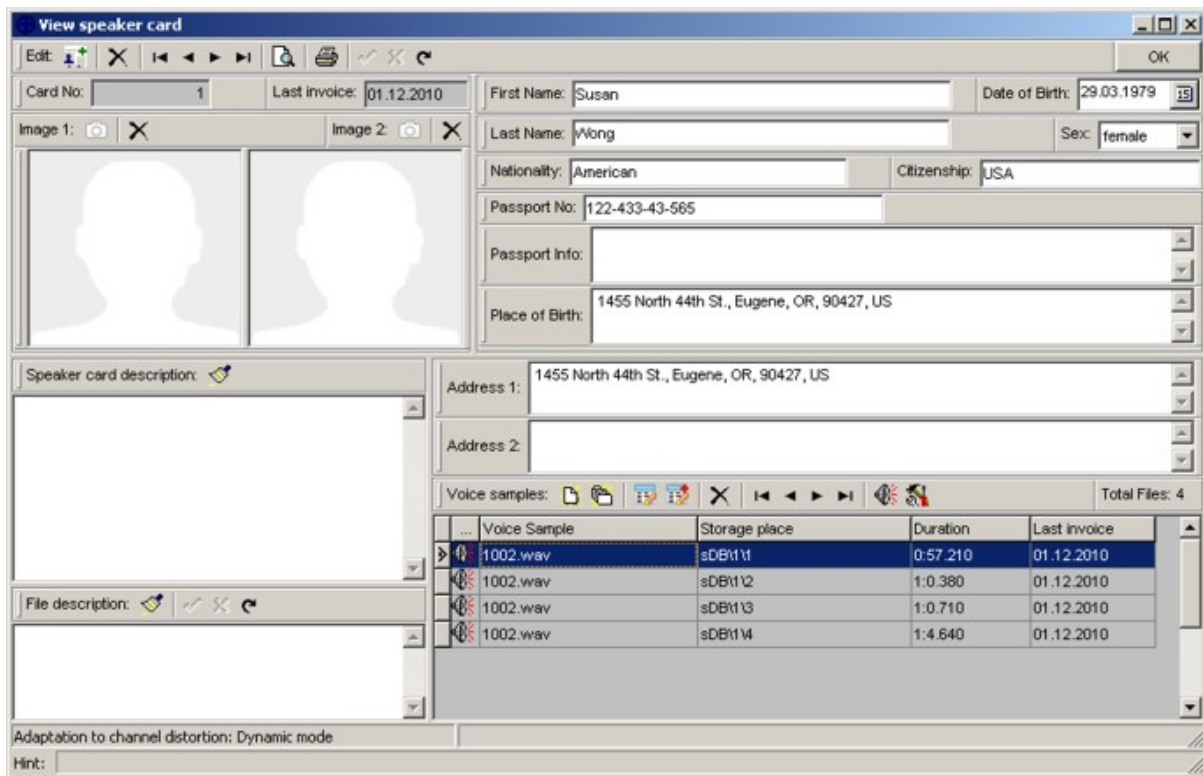
Further the system automatically to reset all Databases for speakers cards, voice samples and the identified files existing for given Database.



**Fig 8. «Database» section with current Database files**

#### 8.4. Target speaker cards

Speaker card besides information about current speaker (first, last name, birthday, gender, and so on) is characterized by examples of audio files with the speaker voice.

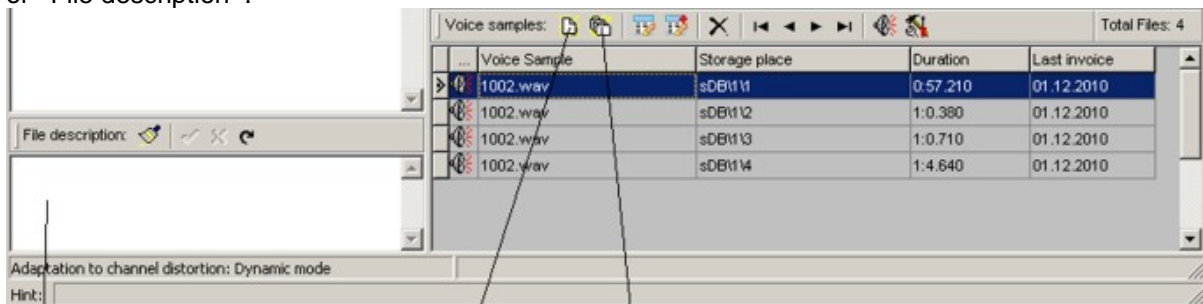


**Fig 9. The panel of preview speaker card**

Each example of audio file is described by the acoustic voice model, error model (FAR, FRR, EER) and noise model, describing surrounding noises and channel distortion, existing in audio file. For the full description of each speaker card it is sufficiently 1 - 3 audio files with the speaker voice, recorded for different telephone lines (cell, voip, spot) and duration of each one not less than 30 sec.

Enrollment (construction) of voice model, error model and noise model for each audio file with voice sample of the target speaker makes automatically at its addition in Database of «Voice samples». To add a new audio file with voice of the target speaker in Database of «Voice samples» you should press the button «Add a new voice sample file». After a choice of audio file, the system will automatically copy the chosen audio file in Database «Voice samples» and will build voice model, error model and noise model.

For each audio file with a voice of the target speaker it's possible to leave comments where and in what conditions the current audio file has been recorded. You can leave such comments in the panel of «File description».



Set your comments for current audio file with voice sample of current target speaker

Add a new audio file with voice sample

Add new audio files with voice sample from the folder

**Fig 10. Add audio file with voice sample into «Voice samples» database**

Also exist possibility for current speaker card add not one audio file, and group of audio files with speaker voice at once. To add group of audio files with voice of the current target speaker you should press the button «Add new voice samples files from the folder and all sub folders». After a choice of folder with audio files, the system automatically will begin copying of audio files into Database «Voice samples» and will begin to build the voice model, error model and noise model for each audio file.

Add a few audio files with voice of the target speaker it's recommended in case audio files with voice of the target speaker are recorded in different communication channels (example, cell and voip). Further at identification it will allow to lower influence of changing cross channel on identification accuracy.

Full structure of speaker card is shown in below.

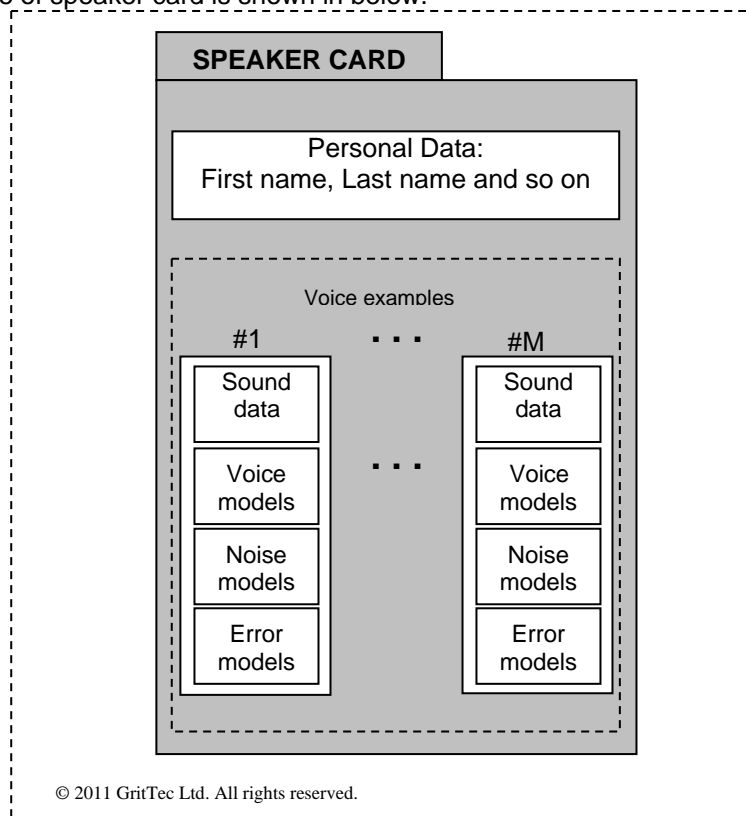
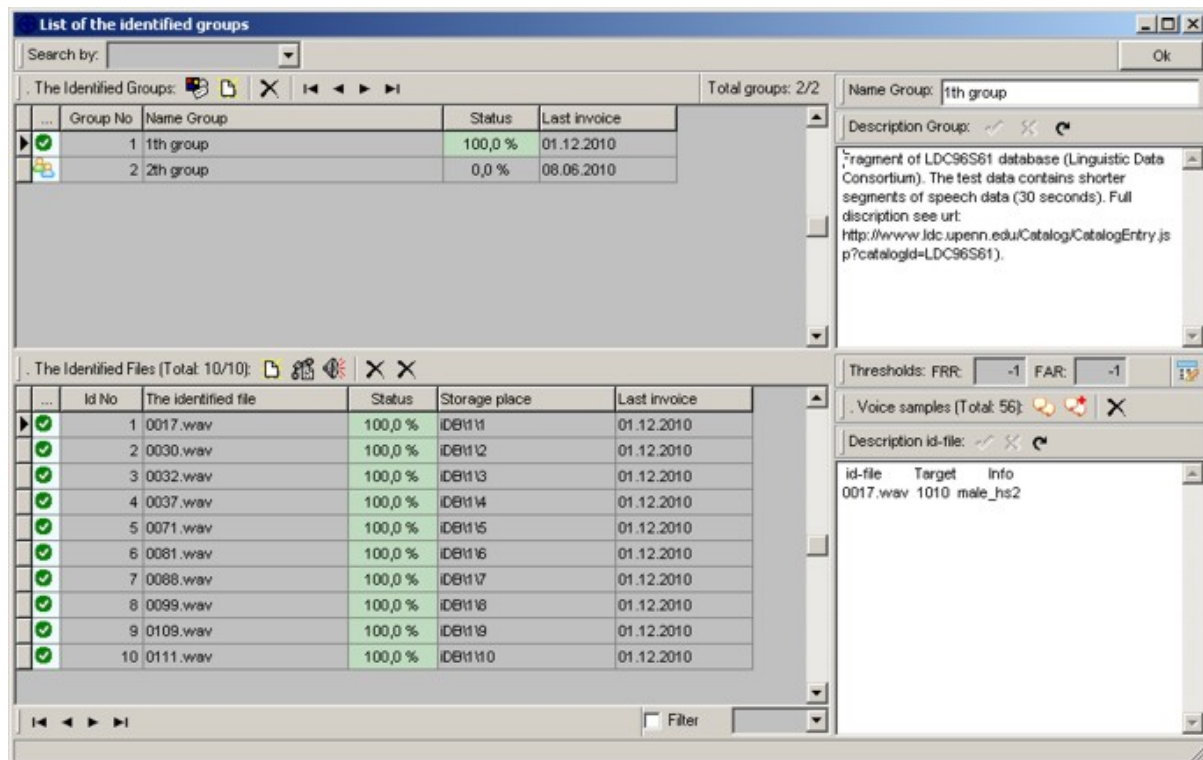


Fig 11. Structure of Speaker Card

## 8.5. List of identified Groups

To open window with list of identified group you should press the button «Show list of identified groups» in the main panel of program.

In opened panels of the identified groups you can see the records of existing groups and lists corresponding to them identified files.



**Fig 12. The panel of list identified groups**

To add new group you should press the button «Create the new identified group». For each identified group it's possible to make comments in a window «Description Group».

To add audio files which in a consequence will be identified, you will choose group for which wish to add audio files and press the button «Add identified file (s)».

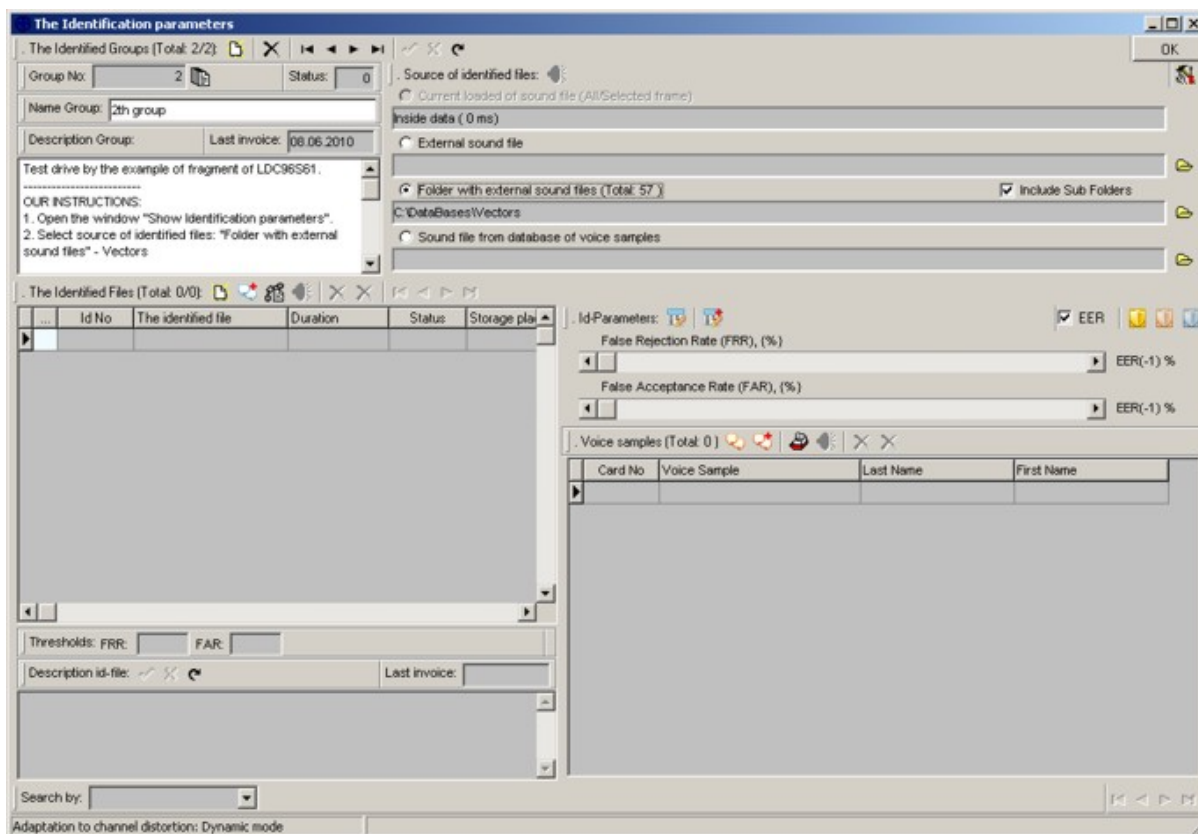
## 8.6. The identified parameters

To add audio file for identification you should open window «The identified parameters» from the main panel of program or open window «The identified parameters» from the window «List of identified groups».

Will choose a source of addition of audio files from «Source of identified files» section and you should press the button «Add the identified file (s) from current Source». Then, in Database of identified files for current group will be copied audio file which in a consequence will be identified.

Further you should set lists of target samples with which will execute identification for each identified audio file. There are a few variants of setting:

1. Addition of all voice samples from all target speakers cards for all identified files. To execute it you should press the button «Add all voice samples for all identified files».
2. Addition of the limited quantity of voice samples from the current target speakers cards for the chosen identified file. To execute it you should press the button «Add all voice samples using for identification» in section «Voice Samples (Total: N)». To change quantity of voice samples you should press button «Add/Change voice samples using for identification». In the opened window you can choose the necessary voice samples from speakers cards which will be used for identification.



**Fig 13. The panel of the identification parameters**

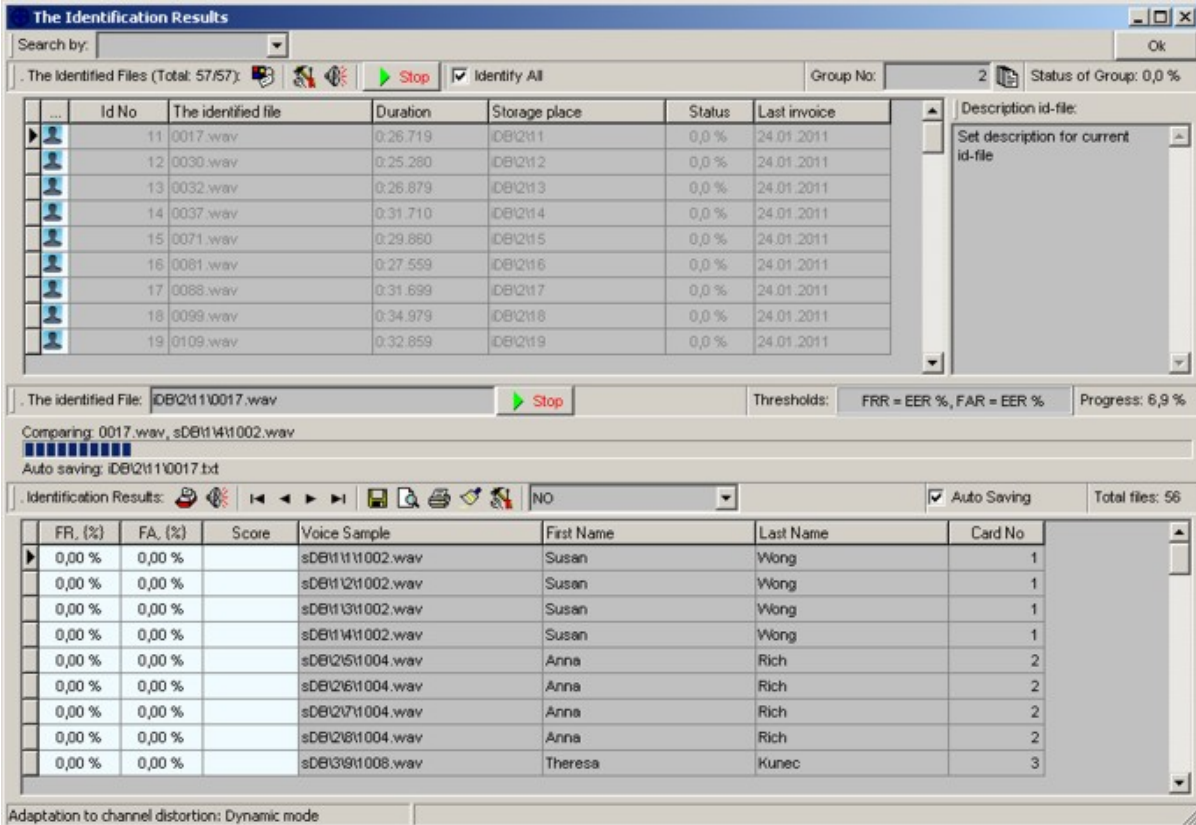
In the process of addition of target samples lists for each identified file, thresholds of identification (FRR, FAR) for each identified file also establishes. By default they are established by EER, i.e.  $FRR = FAR = EER$ . We recommend to use the these thresholds, but they may be changed.

After addition of target samples for identified files to start identification process, you should press the button «Start identification for the current identified file». In result will opened the window «The Identification Results».

## 8.7. The Identification Results

To begin identification process you should press the button «Run identification group» for identification of all audio files or press the button «Run identification for current identified file» for identification of the selected the identified file.

Process of voice identification is executed in an automatic mode, without post processing of identified files. For this purpose it is enough to choose the identified group which will be identified, and to start identification process of group or the current identified file.



The screenshot shows the 'The Identification Results' window. At the top, there is a search bar and a 'Search by:' dropdown. Below that, a status bar indicates 'The Identified Files (Total: 57/57)', a 'Stop' button, and an 'Identify All' checkbox. The main area contains a table of identified files:

...	Id No	The identified file	Duration	Storage place	Status	Last invoice
	11	0017.wav	0:26.719	IDB\Q\11	0,0 %	24.01.2011
	12	0030.wav	0:25.280	IDB\Q\12	0,0 %	24.01.2011
	13	0032.wav	0:26.879	IDB\Q\13	0,0 %	24.01.2011
	14	0037.wav	0:31.710	IDB\Q\14	0,0 %	24.01.2011
	15	0071.wav	0:29.860	IDB\Q\15	0,0 %	24.01.2011
	16	0081.wav	0:27.559	IDB\Q\16	0,0 %	24.01.2011
	17	0088.wav	0:31.699	IDB\Q\17	0,0 %	24.01.2011
	18	0099.wav	0:34.979	IDB\Q\18	0,0 %	24.01.2011
	19	0109.wav	0:32.859	IDB\Q\19	0,0 %	24.01.2011

Below the table, there is a section for 'The identified File: IDB\Q\11\0017.wav' with a 'Stop' button and 'Thresholds: FRR = EER %, FAR = EER %' and 'Progress: 6,9 %'. It also shows 'Comparing: 0017.wav, sDB\11\1\002.wav' and 'Auto saving: IDB\Q\11\0017.txt'. At the bottom, there is a detailed table of identification results:

FR, (%)	FA, (%)	Score	Voice Sample	First Name	Last Name	Card No
0,00 %	0,00 %		sDB\11\1\002.wav	Susan	Wong	1
0,00 %	0,00 %		sDB\11\2\1\002.wav	Susan	Wong	1
0,00 %	0,00 %		sDB\11\3\1\002.wav	Susan	Wong	1
0,00 %	0,00 %		sDB\11\4\1\002.wav	Susan	Wong	1
0,00 %	0,00 %		sDB\2\5\1\004.wav	Anna	Rich	2
0,00 %	0,00 %		sDB\2\6\1\004.wav	Anna	Rich	2
0,00 %	0,00 %		sDB\2\7\1\004.wav	Anna	Rich	2
0,00 %	0,00 %		sDB\2\8\1\004.wav	Anna	Rich	2
0,00 %	0,00 %		sDB\3\9\1\008.wav	Theresa	Kunec	3

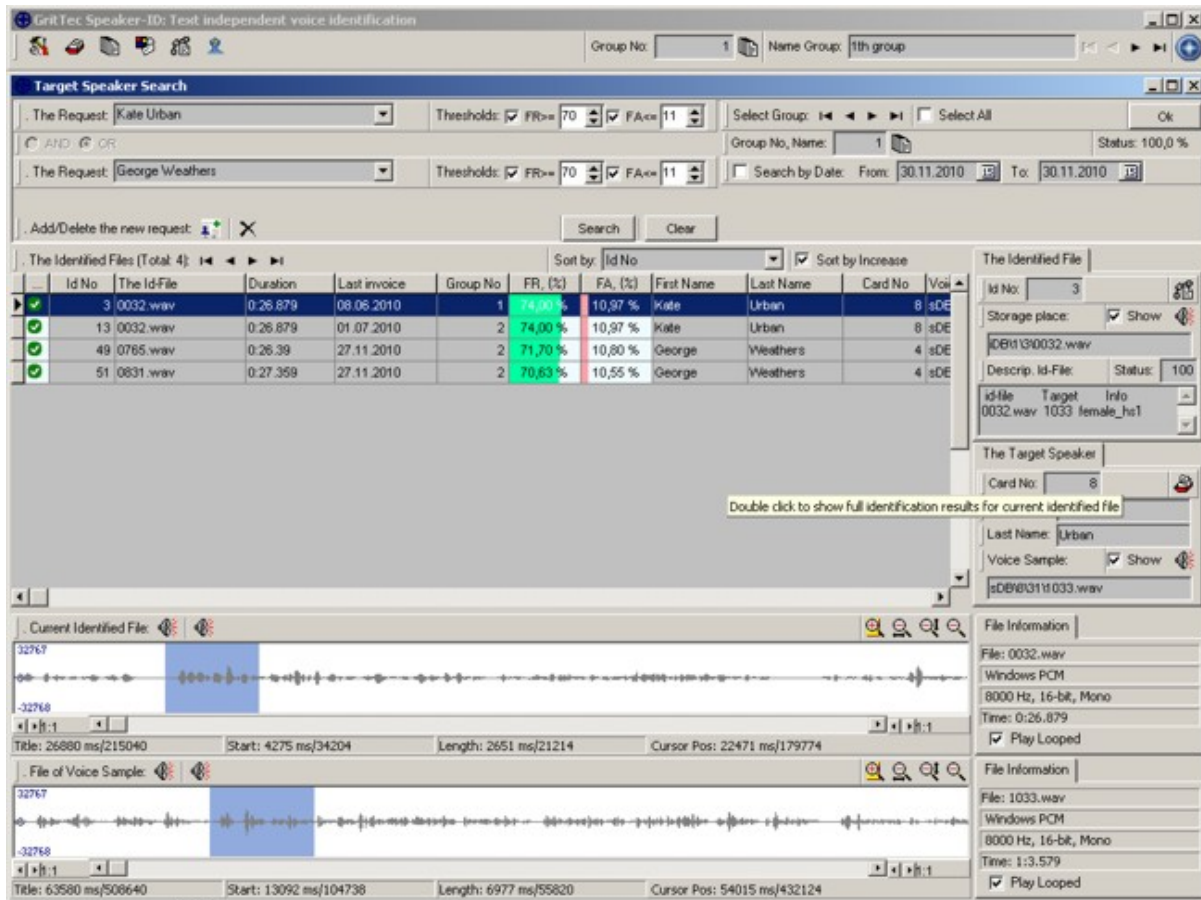
The bottom status bar shows 'Adaptation to channel distortion: Dynamic mode'.

Fig 14. The panel of the identification results

After finishing of identification process you can look identification results for each identified file.

## 8.8. Target speaker search (the identification request)

The search form of the target speakers allows making the automated search of target speakers in the Database of identified files. Search can be made by three requests of the required speakers, existing in the speaker cards database. The flexible system of request settings, display forms of search results and display signal forms of the current identified file and target sample file allow to make effectively search and studying of the received request results.



The screenshot shows the 'GritTec Speaker-ID: Text independent voice identification' application. The 'Target Speaker Search' panel is active, displaying search results for two target speakers: Kate Urban and George Weathers. The search criteria include 'The Request' (Kate Urban and George Weathers), 'Thresholds' (FR=70, FA=11), and 'Search by Date' (From: 30.11.2010, To: 30.11.2010). The 'The Identified Files' table lists four files with their respective FR and FA percentages and speaker names.

Id No	The Id-File	Duration	Last invoice	Group No	FR (%)	FA (%)	First Name	Last Name	Card No	Voic
3	0032.wav	0:26.879	08.06.2010	1	74,00 %	10,97 %	Kate	Urban	8	sDE
13	0032.wav	0:26.879	01.07.2010	2	74,00 %	10,97 %	Kate	Urban	8	sDE
49	0765.wav	0:26.39	27.11.2010	2	71,70 %	10,80 %	George	Weathers	4	sDE
51	0831.wav	0:27.359	27.11.2010	2	70,83 %	10,55 %	George	Weathers	4	sDE

The interface also shows a 'Current Identified File' section with two audio waveforms and their corresponding file information:

- File 1:** 0032.wav, Windows PCM, 8000 Hz, 16-bit, Mono, Time: 0:26.879, Play Looped.
- File 2:** 1033.wav, Windows PCM, 8000 Hz, 16-bit, Mono, Time: 1:3.579, Play Looped.

Fig 15. The panel of speaker search.

## 8.9. Enter the license key

To update or to enter the license key you should open the panel «System configuration» section «License Info» and press the button «Enter license key». You should see the window «Enter Key». To generate your license key you should send us your hardware fingerprint key for PC computer where you plan to use our software.

## 9. Testing

Testing of GritTec's Speaker-ID engine was conducted on the real telephone records and on specialized sound base LDC96S61 of English telephone records given by LDC consortium (Linguistic Data Consortium). [2].

## 10. References

1. GritTec's Speaker-ID: Automatic text independent speaker identification. GritTec Ltd. (<http://www.gritttec.com>).
2. LDC96S61 - 1996 Speaker Recognition Benchmark, Mark Przybocki, Alvin Martin. Linguistic Data Consortium, Philadelphia. (<http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC96S61>.)

## APPENDIX 1

Under the term of SNR the following equation is assumed:

$$SNR = 10 \cdot \log_{10} \left( \frac{E_S - E_N}{E_N} \right),$$

where  $E_S$  – is an average energy of speech signal (without pauses) for at least 20 second in length;  $E_N$  – is an average energy of pause signal for at least 20 second in length; the pause signal must be taken nearby of the speech signal used for calculation of  $E_S$ ; both speech and pause signals must not be preprocessed or filtered by any way; only original signal (without any preprocessing) must be used for calculation of SNR. The average energy of signal is calculated as a sum of all signal samples divided by the number of samples. If  $E_S < E_N$  then it is assumed that the equation  $E_S - E_N$  is equal to zero, i.e.  $\forall E_S < E_N, E_S - E_N = 0$ .