2009

BATVOX 3.0
PRODUCT OVERVIEW

1.   INTRODUCTION

The use of biometrics like fingerprint, DNA or voice has become since few years a best practice in forensic, law enforcement and intelligence organizations for bringing new evidences in court or identifying suspects.

The great advantage of the voice biometrics is that it can be remotely obtained, using the current communication infrastructures (land phones, GSM, VoIP, etc.) without any need to purchase expensive recording hardware like finger print scanners, cameras, etc. Sometime, the voice biometric recordings are the only evidences available, with no physical contacts with the criminals.

Most of the law enforcement and defense organizations in the World are already using voice interception systems capable of storing thousands of hours of audios per day but few only have deployed automatic systems for filtering out targeted speakers in a timely manner.

BATVOX is a world class expert tool for analyzing intercepted voices against recorded voices of known suspects and building expert reports to be used in court.

In its last version, BATVOX 3.0 integrates AGNTIO's latest generation of Speaker Identification (SID) technology that is text, language and channel independent.

BATVOX is a software application dedicated to forensic analysis of voice samples based on the state-of-the-art automatic speaker identification technology developed by AGNITIO.

Since 2006, AGNITIO has successfully participated in all NIST (National Institute of Standards and Technologies) evaluations of speaker recognition systems and is consistently ranked among the top 3.

AGNITIO's technology has been proven in the most demanding and challenging environments. Identification of suspects or criminals even with short voice samples obtained in the "field" makes it the most reliable voice biometrics technology available for Homeland Security applications.

BATVOX is currently used by qualified experts and scientific police in more than 20 countries worldwide that include North America, South America, Europe and Asia.

2. GENERAL DESCRIPTION

BATVOX 3.0 is an automatic system for speaker recognition based on biometric technology.

The system is able to store cases in which, one or more calculations can be run to identify unknown voice samples against a set of known speakers. The identification results consist of a list of potential suspects, ordered from highest to lowest matching scores with the suspect.

	Model Order 1		
	Name	**Score**	**Fit Level**
Impostor Microfono con silencio	model 1	0.518	Ok
Impostor 1 GSM	model 1	0.4683	Ok
Impostor 2 GSM	model 1	1.921	Ok
Dubitado Sospechoso GSM	model 1	5.1638	Ok
Impostor 3 GSM	model 1	0.9941	Ok

○ Audio File vs Model ○ Model vs Audio File **Threshold** 2.5

💾 Save Log   💾 Save Data

Progress
100%

❌ Identification 1 ❌ LR 1

Figure 1: Identification Results window

Voice audios are stored in the database, along with their main characteristics for facilitating the voice classification and search (i.e. audio storing date, recording channel, spoken language, speaker gender and speaking style).

Two types of calculation can be performed within a case:

- **Identification**: comparison between an unknown voice samples against a known target voice are performed by the system in order to determine which unknown voice has a higher probability to belong to the known speaker.
- **Likelihood Ratio (LR)**: as used for DNA verification, this calculus is a probability ratio, equal to the probability that the unknown voice sample belongs to the target divided by the probability that it belongs to the "rest of the world".

Two different architectures are available for BATVOX 3.0: BATVOX Basic and BATVOX Pro. The characteristics of each of these architectures are detailed in section 3 of this document.

3. BATVOX 3.0 products

There are two types of architectures available for BATVOX 3.0:

- **BATVOX Basic:** BATVOX Basic is a solution based on a single machine and designed for a single user. Therefore, the concepts of User Groups or User Administration do not apply for BATVOX Basic. Some other privileges such as the administration of the Quality Module are not available for this user.

- **BATVOX Pro:** BATVOX Pro is a distributed system Client-Server. Several users with a Client version of BATVOX can connect to the central Server and share information.

    Two types of profiles are defined for BATVOX Pro: Administrators and Users.

    An Administrator has the right to create users within the system, create groups of users and administrate the Quality Module of BATVOX 3.0, where several thresholds can be redefined in order to adapt the system to the needs of the users. Configuring the acceptance of the system the administrator can specify whether the system follows a more exigent approach regarding the quality of the audios accepted or not.

    Users within the same group of users share the cases they have created with the other users in the group, but not the sessions and the specific calculations performed. Users from different Groups cannot see the cases from other groups, and finally the Administrators can always have access to all the cases created within every group defined in BATVOX 3.0.

4. MAIN FUNCTIONALITIES

In this section, the main functionalities of BATVOX 3.0 is highlighted.

- **Creation of Cases and Sessions:** BATVOX 3.0 allows creating *Cases* in order to facilitate the organization of audios. Within a case, audios are introduced, to be part of future calculations. A case is structured in *Sessions*, where the calculations are created.

Figure 2: Case window

- **Storage of audio:** audios added in the system can be recorded from Identified Speakers or Non-Identified Speakers.

  In order to use audios in identification, they have to be registered in the system. After that, for each audio we want to compare against the rest of the

audios in the system (typically the known speaker´s audios) a biometrical model has to be generated. The characteristics of this biometrical model are described later in this section.

When several audio samples correspond to the same speaker but have been recorded in different sessions all of them can be associated to the speaker in order to improve the quality and length of the audio associated.

BATVOX 3.0 accepts audio files in the following format: .wav files with Linear PCM coding, sampling frequency 8 KHz, 16-bit resolution and mono.

- *Registration, modification and deletion* of audio files. As mentioned before, the first step for running calculations is to register the audios in the system. Once the audios needed for the calculations are registered, it is possibile to modify some of the data associated to them or delete them.

Figure 3: New audios window

- *Training of voice models:* when one or several audios have been added for a speaker and before performing an identification which includes that speaker, another process has to take place: the training of the voice model.

Each individual can have only one voice model associated. This voice model represents the specific characteristics of the individual´s voice, allowing the comparison of voice samples against this model. A net speech length of 40 seconds is required to generate a model, considering all the audios included for a speaker.

The process of creating a voice model has to be launched by the user, who can follow up the process through a progress bar in the interface and can check

when it is finished, meaning this that the speaker is ready to be included in a new identification.

- **Identifications:** an identification calculation implies the comparison of one or more biometric models to one or more audio files. The result of the comparison shows a list of Audio files versus Models or Models versus Audio files. In the first way, the models with best scores against each test audio are listed and, in the second way, test audios with best scores against each model are listed.

  The scores shown in the results indicate which of the audio files compared to a model has a higher probability to belong to the Individual whose biometric characteristics are represented in the model.

  When the audios used in identification are recorded in heterogeneous conditions, there are a series of characteristics that could influence the results of these comparisons such as the recording channel, the spoken language, etc. For example, two audios with the same recording characteristics (for example, both recorded via telephone line) could get a higher score (that is to say, look alike) in the comparison between themselves even though they might be voices from different speakers instead of two audios from the same speaker with different characteristics (one recorded via microphone, the other via telephone, for example). For this reason, when working under circumstances in which the homogenous quality of the audios cannot be guaranteed, it is essential to minimize the effects of these characteristics.

  BATVOX 3.0 includes the *Normalization Process* to minimize the effects of these characteristics.

  Normalization is the process of correcting the effects that the lack of alignment has on statistical scoring. It enables effective scores comparisons.

  A case can include as many as needed identifications.

- **Likelihood ratios:** a Likelihood Ratio (LR) is a special case of verification.

  Verification is a 1 to 1 comparison where the result can conclude if the identity of the speaker matches the sample voice. The verification result in BATVOX 3.0 is expressed as a relationship of probabilities.

  The LR is calculated as the quotient of the likelihood that the test voice belongs to the suspect divided by the likelihood that the test voice belongs to someone else (in other words that it does not belong to the suspect).
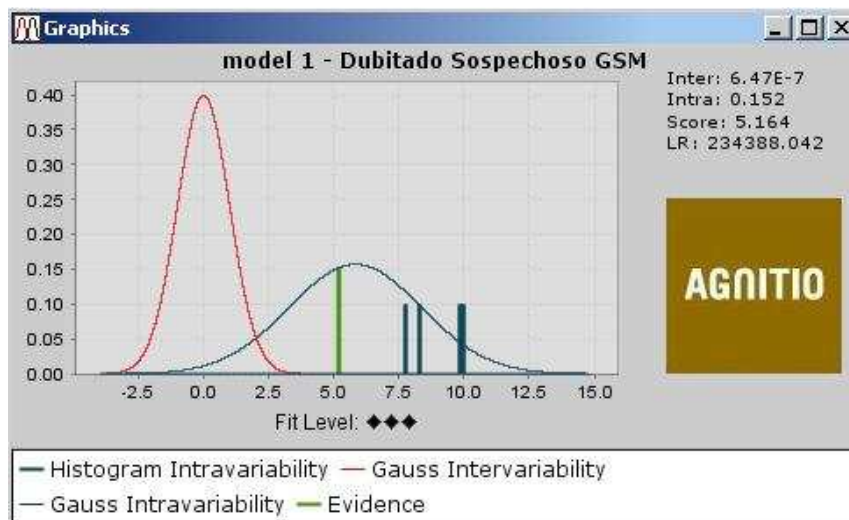
Figure 4: LR graphic

This value gives a very accurate idea of the verification results, in a probabilistic way, which can be easily presented and justified in legal processes, always with the support of an expert.

Figure 5: LR Results window

- *Export the results of an identification or likelihood ratio to an html file:* The results of an identification/LR can always be exported to an html file.

- *Checking of the quality of the results:* the user of BATVOX 3.0 is always going to be able to check the reliability of the identifications and LRs performed. According to the quality of the audios used in the identification, the models generated and some other characteristics, a *Fit Level* is going to be defined indicating the user several degrees of reliability in the result obtained.

  In the case of the Administrator user we have defined for the architecture of BATVOX Pro, the thresholds determining the quality of the audios and therefore the quality of the calculation can be configured.

- *Administration of users:* also in the case of Administrators there is a possibility to manage the creation, modification and suspension of users. User groups can also be created in order to manage the access rights to the cases created in a group only for the users belonging to that group.

- *Wizard or user manual:* at any step of the processes in BATVOX 3.0, a Wizard and a User manual are available for the user. The theoretical principles of the tool as well as the process steps which need to be followed are described in both documents.

5.   ARQUITECTURE

As mentioned in section 3, two architectures are offered for BATVOX 3.0:

- **BATVOX Basic** is local application which does not need an application server in order to manage its database or the rest of the modules within the application. It can be run in a single machine including all the modules needed for running all the biometric tasks involved in the identification processes.
- **BATVOX Pro** includes a centralized server and a set of clients connected to that server. The common modules such as the database are stored in the central server and shared by all the clients connected to the server.

ASIS 2.0 is a tool provided by AGNITIO that allows to build centralized databases of voices of known or unknown suspects and to identify new voice samples against the stored voices. This tool can be connected to BATVOX 3.0 in order to share information from the cases and identifications in ASIS 2.0 and contrast the results with the results in BATVOX 3.0.

6. AGNITIO: PRODUCTS FOR HOMELAND SECURITY

AGNITIO's proprietary technology is the most sophisticated, robust and reliable technology in the voice biometrics market.

AGNITIO has developed the most advanced solutions for text and language independent voice recognition.

AGNITIO's technology has been deployed over 20 countries worldwide and has been proven in the most demanding environments: forensic and law enforcement agencies globally.

BATVOX
Forensic laboratories

ASIS
Law enforcement

c Speaker Identification
eployed for building
databases of voice
known criminals and
e investigations on
ices samples.

SIFT
Field investigation

BS3
Intelligence agencies

Voice models

it
of
al
s.

tric Speaker Spotting
matically identifying
;" in large databases or
pted voices.