

	<p>SemanticHiFi <i>IST-507913</i></p>
<p>D 8.3: Public Report of WP8 “HiFi System” Covering period: December 2003 – November 2005</p> <p>Report version: 1.0 Report preparation date: 30.11.2006 Writers: Taufan Zimmer, Dr. Wilhelm Hagg, Sony EuTEC Control: Hugues Vinet, IRCAM Classification: Public Contract start date: December, 1st 2003 Duration: 36 months Project coordinator: Hugues Vinet, IRCAM Involved Partners: Sony EuTEC (WP Coordinator), Ircam, FhG, Sony CSL, UPF, BGU, NI</p>	
	<p>Project funded by the European Community under the "Information Society Technology" Programme</p>

Table of contents

1	WP Overview	3
1.1	Objectives	3
1.2	Partners' roles	3
1.3	WP contribution to the project	3
1.4	Synthesis of main achievements	4
2	WP Results and Achievements	5
2.1	Market requirements study	5
2.2	Functional description	5
2.2.1	User interface	5
2.2.2	Basic HiFi system features	6
2.2.3	Browsing/Searching	6
2.2.4	Playback	7
2.2.5	Sharing	7
2.3	Hardware Implementation	8
2.4	Software Implementation	9
2.4.1	Operating System	9
2.4.2	OS Packages	9
2.4.3	OS Libraries	10
2.5	System Architecture	11
2.6	Position over the state-of-the art	13
3	User Interface	15
3.1	The Touchscreen Remote Controller	15
3.2	Usability aspects	16
3.3	Design	17
3.4	PDA GUI Functionality	18
3.4.1	Advanced Browsing	18
3.4.2	Similar songs	19
3.4.3	Synchronized Lyrics	19
3.4.4	Organize Track Lists by Attributes	20
3.4.5	Query by Humming	21
3.4.6	Personal Categories & Generalization	21
3.4.7	Time stretch	22
3.4.8	Play Audio summary	23
3.4.9	MixFiles Playback from Authoring System	23
3.4.10	Sharing	24
3.5	The Front Panel Touchscreen	25
3.5.1	Show Track Metadata and Artist Picture	25
3.5.2	Navigate on song segments	25

1 WP Overview

1.1 Objectives

The target of WP 8 of the SemanticHiFi project is to develop a HiFi system prototype by utilizing various technologies provided within the project. The HiFi System is targeting at a future HiFi component enabling new operation modes by utilizing features related to the description of musical contents. These features enable advanced methods for music selection in huge databases as well as new visualization methods of music content based on the features of a track. A sharing system finally allows sharing of musical metadata such as playlists with the community. Technically the key features of the HiFi System are the ability to store a big amount of music on a hard disc as well as the possibility to communicate via a network with other HiFi systems, with authoring systems and with internet based online data services.

1.2 Partners' roles

The role of Sony EuTEC in the development of the HiFi system is to provide a framework for the integration of different partner technologies and to do the integration of the corresponding software modules. Furthermore Sony EuTEC provides a graphical user interface that allows for a proper operation of the HiFi system.

The partners are responsible for providing their software modules for the HiFi target platform according to the specification.

1.3 WP contribution to the project

The contribution of WP 8 is to provide a “close to product” prototype of a HiFi system that integrates a selected set of partner technologies developed in the project. The HiFi system is part of a global sharing scenario that integrates different HiFi systems as well as authoring systems. While an authoring system is targeting at audio authoring with different effects and a high degree of interactivity, the HiFi system focus is on the living room scenario where people basically consume music.



1.4 Synthesis of main achievements

A fully operational HiFi system has been developed that offers a lot of new operation aspects related to advanced music selection and metadata display based on the semantics of music.

The main implemented features include :

- music titles management through user-defined categories and a generalization system based on automatic signal analysis;
- content-based inter-document browsing combining multiple heuristics, including musical descriptors automatically extracted from the audio recordings (tempo, tonality, timbre, etc.) and query by humming;
- new interfaces for intra-document browsing and assisted listening: slow and accelerated time-stretching, automatic analysis of the music temporal structure, automatic generation of music summaries, real time display of lyrics during playback, execution of small hypermedia applications;
- exchange of user-produced musical metadata through a dedicated peer-to-peer sharing middleware: user-defined categories, musical descriptors, etc.
- interoperability with the Authoring PC application through the peer-to-peer system: execution of mix files, playlists, metadata import, etc.

All of these features are at the level of the current state-of-the-art or beyond in digital audio signal analysis and processing and music information retrieval technology, and outperform those of existing products in the field of music access and listening.

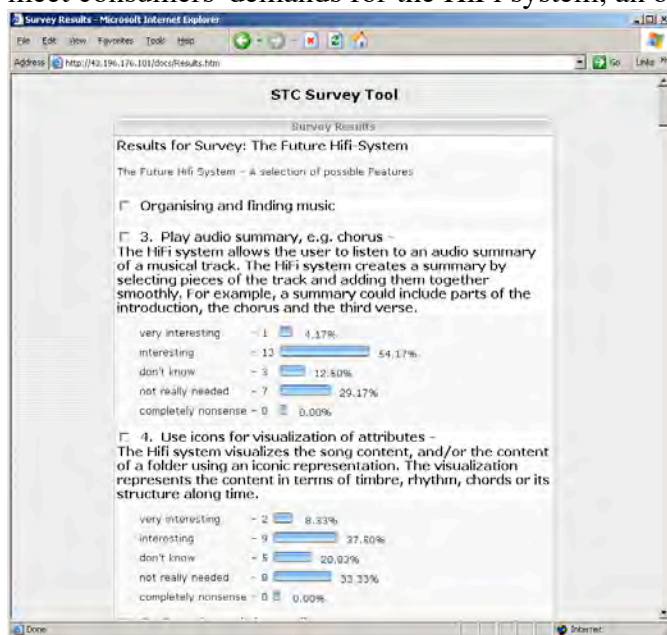
The developed system provides a consistent environment for the integration of all these features in terms of technical architecture and user interfaces. Several usability aspects have been evaluated in order to provide a system that is operable with limited I/O devices such as a PDA remote control. Targeting a living room scenario, the usability of new HiFi system features are regarded a key issue for product success.



2 WP Results and Achievements

2.1 Market requirements study

Today, much of the functionality used by consumers to organize and manipulate digital music is available as part of PC software applications. The Semantic HiFi project aims to bring this functionality, as well as new functionality in the areas of browsing, listening, interacting, performing and sharing, to the HiFi system. The different partners in the project have produced many new ideas and features. In order to determine which of these features best meet consumers' demands for the HiFi system, an online survey has been developed.



The result of the survey has been used for a selection of the most relevant features for integration into the HiFi system prototype. In addition to the user requirements from the survey, an evaluation of technical and market constraints has been done in order to select features that are relevant for integration into real products.

2.2 Functional description

The following list gives an overview of the system features that are realized in the HiFi System. Some basic features in the list below are not described in detail since they are trivial functionalities that don't need further explanation. However they had be implemented in the HiFi system.

2.2.1 User interface

While current advanced audio products still use traditional remote controllers, there is a strong demand from users for advanced user interfaces for HiFi Systems.

Touch screen remote control

The HiFi System uses an advanced remote control with integrated touch-screen.

Touch screen on HiFi unit

The HiFi System itself has an integrated touch-screen for basic operations and displaying song context such as the song structure.

2.2.2 Basic HiFi system features

This are the basic features people expect from a HiFi System. They are a mix of features people know from existing HiFi systems and PC based audio players.

Play, Stop, Pause, FF, FR, Next, Prev, Repeat, Shuffle

Change volume

Ripping of CDs

Transfer files from USB

Browsing by artist/album/title

Search by artist/album/title

Playlist creation, saving, execution, deleting

2.2.3 Browsing/Searching

These features describe the advanced methods for music selection. Music selection is regarded as a key issue for future HiFi Systems from a product/marketing point of view.

Searching with metadata

The HiFi System allows the user to search for tracks according to attribute values e.g. retrieve all songs from the 90'ies. Multiple selection on different metadata fields are allowed in order to shrink the number of matching tracks as desired by the user.

Playlist generation based on metadata properties

The HiFi System allows the user to generate a playlist according to attribute values e.g. include songs from year 1997-1999 belonging to genre blues with less than 100 beats per minute.

Organize playlists by metadata attributes

The HiFi System allows the user to organize playlists according to the available metadata attributes, e.g. in order of increasing tempo. The average of all songs in a playlist is computed in order to compute the ranking of the playlists.

Finding similar songs

The HiFi System allows the user to search for songs that are similar to a given song (based upon key acoustical attributes of the music).

Query by Humming

The HiFi System allows the user to search for a song by humming. The user hums a piece of music into a microphone. The HiFi System searches the database for the song.

Audio summary

The HiFi System allows the user to listen to an audio summary of a musical track. The HiFi System creates a summary by selecting pieces of the track and adding them together smoothly. For example, a summary could include parts of the introduction, the chorus and the third verse.

Building personalized categories

The HiFi System allows the user to create new individual genre types. The user may want to specify his own genre for identifying or grouping tracks e.g. the songs for relaxing. He does this by creating a new genre type “relaxing”.

2.2.4 Playback

Providing additional information on songs and artists is regarded as a key element especially for HiFi Systems that have any type of display. The following features were identified as the most attractive from a user perspective.

Show metadata of actual song

Show the song title, artist and album information.

Display lyrics during playback

The HiFi System allows the user to view the lyrics of a song on its display as the song is being played. The lyrics move in time with the song.

Show album cover/artist pictures

The HiFi System shows the album cover and/or artist pictures during playback. The pictures could be grabbed from the Internet if not available locally.

Show Flash animations (synchronized with song)

Show Flash animations that respond in some way on the music signal. This is basically an interface to develop new animations that could be loaded as plugins.

Navigate on song segments

The HiFi System allows the user to navigate within a track graphically. With this feature, the user can view the structure of the track on the HiFi System display, e.g. separate blocks are shown for the introduction, for the chorus and for each verse. The visualization of the blocks represents the content in terms of timbre, rhythm, chords and its structure along time. Using these blocks the user can skip to a particular section within the track.

Time stretch (ff, fr)

The HiFi System allows the user to move forward and backward within the song without the usual audio distortion. Instead song segments are skipped and the remaining segments are aligned with a smooth audio signal transition.

2.2.5 Sharing

The sharing component allows the user to share metadata with other HiFi systems and authoring systems on the network.

P2P user groups (create, join)

The HiFi System allows the user to create a user group on the P2P network. The user may wish

to create a group containing people with whom he wants to share his data, e.g. the members of a club, a class, a group of friends etc.

P2P sharing of musical metadata

The HiFi System allows the user to share and to obtain the attributes of a track from other users on the P2P network. For example, users may exchange attributes such as artist, title, year, genre, tempo, timbre, rhythm etc.

Sharing of playlists

The HiFi System allows the user to share playlists with other users on the P2P network.

Sharing of user generated music classifications

The HiFi System allows the user to share user generated music classifications with other users on the P2P network.

Interaction with Authoring Application

The HiFi System allows playback of mix files that are created on the authoring system.

2.3 Hardware Implementation

The HiFi System has two components: the HiFi unit itself and a remote controller. The HiFi System is using a HTPC case that is available in the market. Beside a DVD player slot as well as hidden slots for card readers and front USB connectors, the big LCD display is a highlight of this case and makes it an ideal candidate for displaying semantics related musical context such as the song structure.



Inside the HiFi component case the following components are used:

- MSI 915GM Speedster FA4 board

- Intel Pentium 4 Mobile 735 (1700MHz) - Dothan
- 100 GB HDD
- 512MB DDR-RAM
- DVD-Burner
- Noise compensation components

The **MSI 915GM Speedster FA4** board enables the use of a Pentium Mobile processor that has a substantial lower power consumption compared to state of the art desktop processors. This solves the cooling problem that is a potential source of noise which is not acceptable in a HiFi System.

The remote controller is based on a PDA. It offers a modern design as well as a high resolution touch screen. The simple design allows using the remote controller also in landscape mode. The PDA is connected by IEEE 802.11 wireless technology to the HiFi system.



2.4 Software Implementation

The following software components are used in the HiFi system:

2.4.1 Operating System

The Ubuntu Linux distribution was selected as operating system. Since linux and other linux-like operating systems are a choice already for a couple of current products, the choice of linux in the HiFi System facilitates a future migration to a real product.

2.4.2 OS Packages

Ubuntu provides the following packages that were used in the HiFi System:

- Java Virtual Machine (JDK 1.5)

Java is used as an integration layer and for implementing the core system. This enables fast and reliable implementation of the required features.

- Macromedia Flash Player 6/7

The user interface has been implemented in flash which enables efficient development and state of the art design for an attractive presentation on the HiFi System. Since Flash Player 7 was not available for PocketPC, version 6 had to be used on the remote controller. On the HiFi system Flash 7 is used.

- MySQL Server 5.0 (SQL Database)
- Jackd (an audio realtime engine)
- Wine (for running the mixfile playback engine)
- Eclipse, subversion and trac have been used for development

2.4.3 OS Libraries

Several specific linux system libraries are used for the purpose of feature extraction on top of the Ubuntu base distribution:

- libstdc++5
- atlas3-base
- libmad0
- libfaad2-0
- fftw3
- mp3info
- cd-discid

Background processing

This module is responsible for the initialisation and control of all tasks that are not triggered by the user interface such as ripping CD's and extracting features from the audio data.

Meta-Data Model

This module represents the HiFi System Metadata Model. It is responsible for initialising the metadata instances. It also provides information about the data model in place, especially about dynamic fields specified by the user.

Meta-Data Manager

This module allows for easy implementation of metadata models. This is the functionality provided by MCM.

SQL DB Access

This module provides access to an SQL database holding the meta-data.

Data Access

This module provides methods for moving, removing, deleting of data on the file system. All corresponding entries in the Meta-Data database need to be changed accordingly by the module. This allows to keep the data in sync with the metadata.

Extraction Process Control

This module is responsible for starting, stopping and controlling the data extraction processes. It provides information on system load, running extraction tasks and so on. It enables to control the system load by stopping extraction processing and allows to assign priorities to extraction tasks. It does a scheduling of processes based on priority and allows to pause and continue processing, e.g. after a system shutdown. The module is also responsible for storing the results in the Meta-Database.

TCP/IP

Provides basic networking functionality.

RT processing control

This module allows to load and unload realtime plugins within the realtime engine. Furthermore, it allows to connect processing and I/O modules by audio channels. It provides start, stop and monitoring methods for the processing chain.

Meta DB

This is an SQL database providing all metadata information.

Audio DB

This is the music database which is basically a file system. It also may hold certain context data such as pictures (CD cover). In general access to this database is read only in order to keep it consistent with the Meta-Data database. Modifications or deletion of material need to be done by the Data Access module.

Extraction Framework

Framework for extraction of metadata from audio files. This framework is in the most simple case based on standalone processes. Some methods for monitoring are provided.

Feature extraction plugin

The executable that is extracting a feature from the songs in the database. The plugin reports the extraction results in a determined way to the extraction control module.

Realtime engine

An audio processing engine with realtime capabilities. It provides basic modules such as decoders, encoders as well as mixing, rendering and recording modules for accessing the sound card hardware.

Signal processing plugins

This are plugins for realtime audio processing that are required for audio playback as well as different applications such as QbH. The plugins are based on the LADSPA plugin standard.

Audio driver

Provides access to audio hardware.

2.6 Position over the state-of-the art

Several categories of existing products are related to the developed features:

- traditional hi-fi systems,
- hard-disk home and public listening devices (juke-boxes),
- online music delivery servers and PC client and title management software,
- personal listening devices, including recent convergences with mobile phones

According to information currently available to project partners, the features of these products are outperformed by those available in the Hi-fi system, including:

- no automatic classification based on audio analysis from user-defined categories;
- no automatic title playlist generation based on signal descriptors;

- no inter-title browsing functions based on music descriptors automatically extracted from audio signals;
- no automatic music summary generation;
- no intra-title navigation interfaces beyond traditional navigation buttons (play, stop, rewind, next track) and linear cursors; no automatic analysis of the musical structure, including temporal; no variable speed, time-stretched playback;
- no peer-to-peer metadata sharing system based on user-produced metadata, and preventing the sharing of the music copyrighted audio files.

3 User Interface

3.1 The Touchscreen Remote Controller

The HiFi System uses an advanced remote controller with an integrated touchscreen. The main user interaction with the HiFi system is done via this advanced PDA style remote controller.



After starting the remote controller GUI a connection to the HiFi System is made. When the remote controller is connected with the HiFi unit, the following “Main Menu” appears on the PDA:



3.2 Usability aspects

Several usability tests have been applied in order to optimize the system access based on a PDA like device. The operation of a PDA like device is very different from operating a PC where mouse and keyboard can be used. The experiments have shown that the most important requirements for a PDA based GUI are:

1. The design must allow using only the thumb for interaction

People will use a stylus only if they are forced to do and thumb operation is much more attractive for most users.

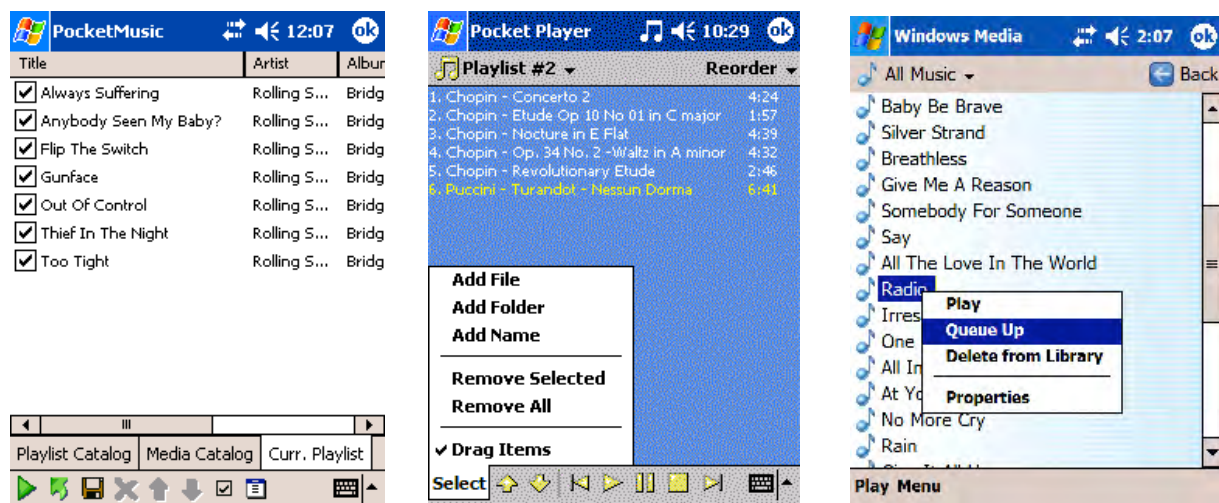
2. The external navigational keys should not be used

The limited space on a PDA device results in a poor usability of the hardware buttons found on the device.

3. The design should be in landscape mode

Experiments have shown that the accessibility of buttons on the screen is much better in landscape mode as long as both hands are used for operating the device.

The figures below show some prototypes that have been developed in order to apply user tests. First a typical PDA application is shown that is operated by stylus.



The next figure shows a prototype that is accessible by thumb operation. To enable thumb operation for complex tasks like scrolling of lists is quite challenging especially on small devices.



The usability issues addressed here are just a side issue within the context of SemanticHiFi and are not fully implemented in the current HiFi system.

3.3 Design

The following figure shows the initial design for the HiFi GUI that has been optimized based on the various user tests. All buttons can easily be accessed by thumb operation. A clear and simple design makes the operation of the system easier and enables the operation by non-expert users. The final prototype has been implemented in portrait mode due to performance problems of the Flash implementation on the PocktPC.



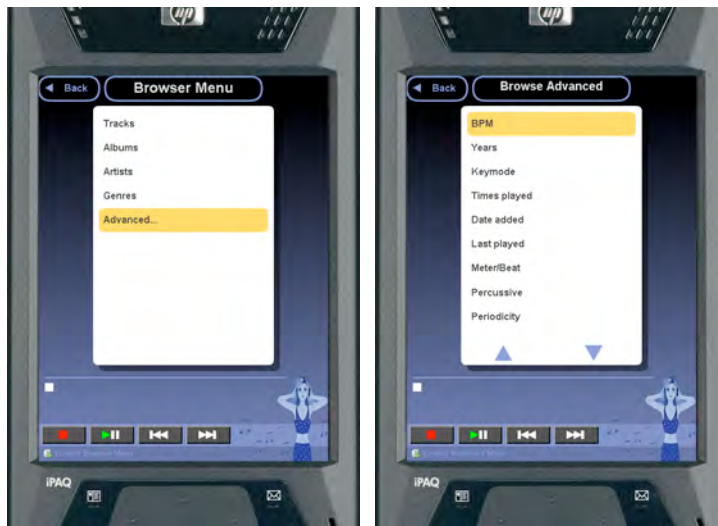
3.4 PDA GUI Functionality

This section highlights some functional aspects of the HiFi system GUI.

3.4.1 Advanced Browsing

The HiFi System allows the user to search for tracks using advanced metadata features, according to attribute values e.g. retrieve the fastest or slowest songs in the Database or songs which has a certain harmonic key or certain level of percussivity etc.

The following figure demonstrates the advanced browsing.



If the user selects for example “BPM” (Beats per Minute) then a menu appears to select songs which have the highest down to the songs which has the lowest BPM in the database. If the user selects “highest”, a list with songs which has the highest BPM will appear.



Consequently, if the user selects the “lowest” a list with songs which has the lowest BPM in the Database will appear.

3.4.2 Similar songs

The HiFi System allows the user to search for songs that are similar to a given song (based upon key acoustical attributes of the music / Timbre) and various other extracted metadata e.g. BPM, Percussivity, Harmonic Key, etc.

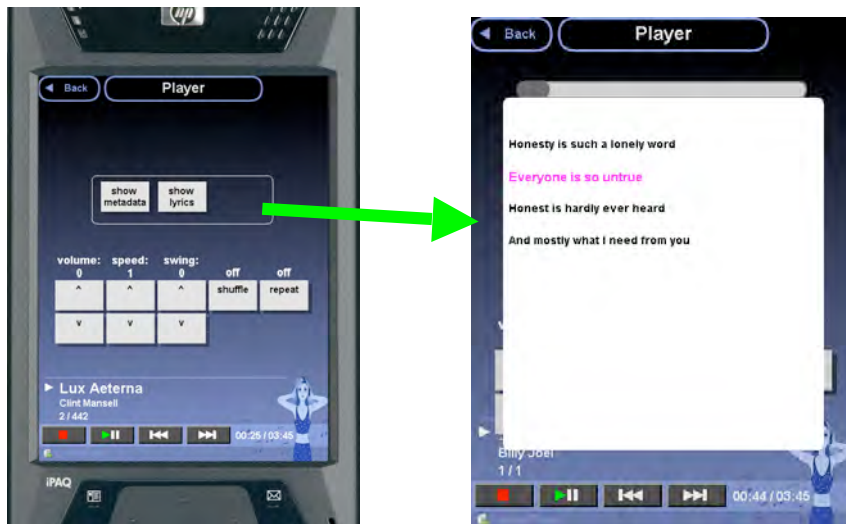


In the “Find Similar” screen, the user can drag and drop various *similarity selectors*. The center is showing the metadata value of the current selected song. The more the specific *similarity selector* is dragged to the center, the more weight it will gain in the similarity search.

After setting one or more desired selectors, the number of similar songs will be computed using the distance of the selector to the center as weight factor. The result is shown as a number in the lower left of the screen. If the user wants to see those similar songs, the “Show>” button can be pressed and a list with corresponding songs will appear.

3.4.3 Synchronized Lyrics

While a song is playing the user can opt to show the lyrics of the song. The system will then search for the corresponding lyrics file and if it is found it will be shown on the GUI. If the user clicks on a line of lyrics, the song will skip to the corresponding timeframe.



3.4.4 Organize Track Lists by Attributes

The HiFi System allows the user to organize track lists / playlists according to the available metadata attributes, e.g. in order of increasing tempo, etc.

First the user selects the desired attribute / metadata to show as column.



Then the list can be sorted by that column.



3.4.5 Query by Humming

The HiFi System allows the user to search for a song by humming. The user hums a piece of music into a microphone. The HiFi System searches the database for the song.



After that a list of tracks that are most similar to the hummed melody will appear on the screen.

3.4.6 Personal Categories & Generalization

The HiFi System allows the user to create new individual classification types using different categorized songs. The user may want to specify his own category for identifying or grouping tracks e.g. the songs for relaxing. He does this by creating a new category type "relaxing".



If the user has at least 2 categories, the system can automatically categorize all tracks in the database using these 2 categories. This automatic categorization will repeat every time the user adds or removes categories to any track in the database.



The automatically generalized classification based on the 2 categorized tracks specified by the user

3.4.7 Time stretch

The advanced time stretch feature of the HiFi System allows the user to play a song faster or slower without the usual audio distortion / skipping. Instead song segments are skipped and the remaining segments are aligned with a smooth audio signal transition.



3.4.8 Play Audio summary

The HiFi System allows the user to listen to an audio summary of a musical track. The HiFi System creates a summary by selecting pieces of the track and adding them together smoothly. For example, a summary could include parts of the introduction, the chorus and the third verse.



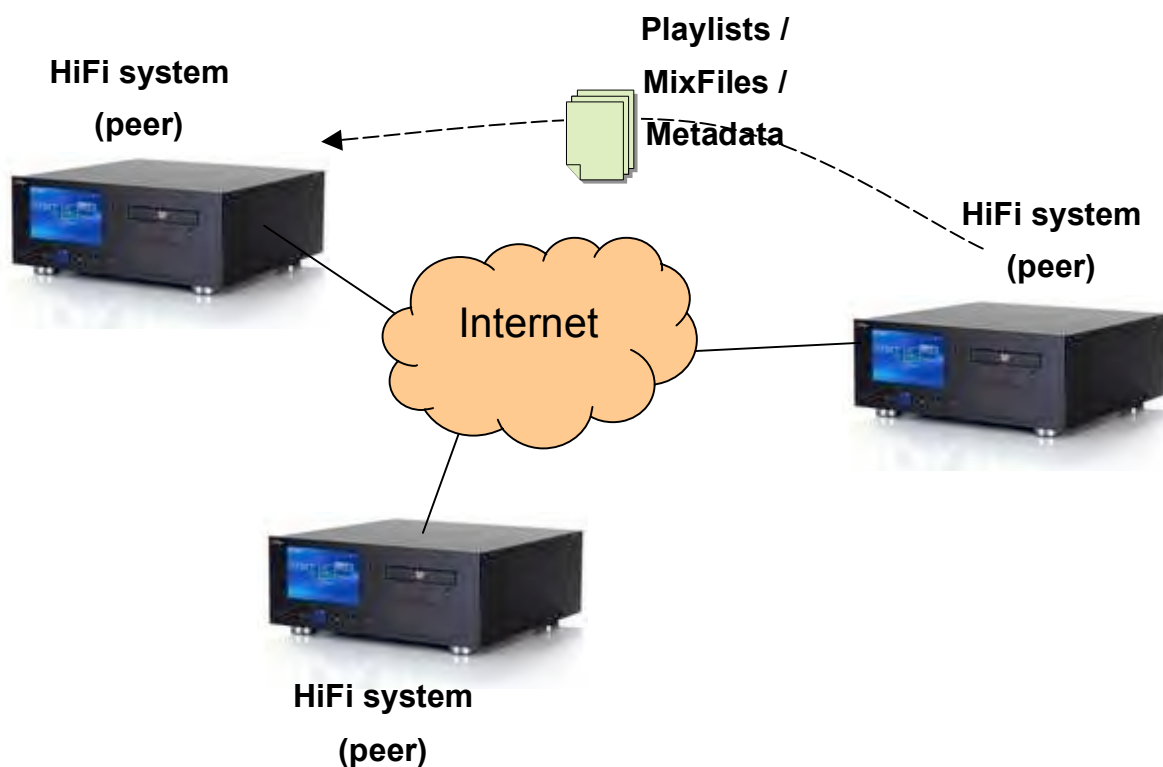
3.4.9 MixFiles Playback from Authoring System

The HiFi System allows playing MixFiles which were created using the Authoring Application.



3.4.10 Sharing

The sharing component allows the user to share song metadata, playlists, Mix-Files and classifications with other HiFi systems and authoring systems on the network.



The HiFi System allows the user to create a user group on the P2P network. The user may wish to create a group containing people with whom he wants to share his data, e.g. the members of a club, a class, a group of friends etc.

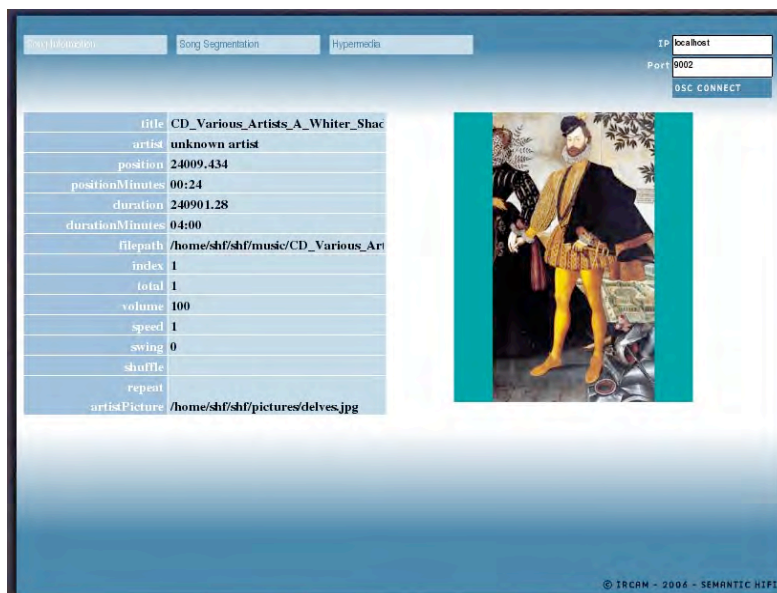
After a group has been joined, the user can select different data to search e.g. title, artist name, etc. In the following search results page, the user can select it and choose to download it. The selected item will then be added to the database and is available through the browsing / searching features of the HiFi System.

3.5 The Front Panel Touchscreen



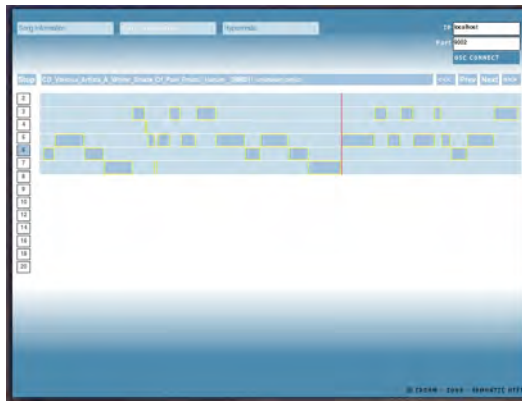
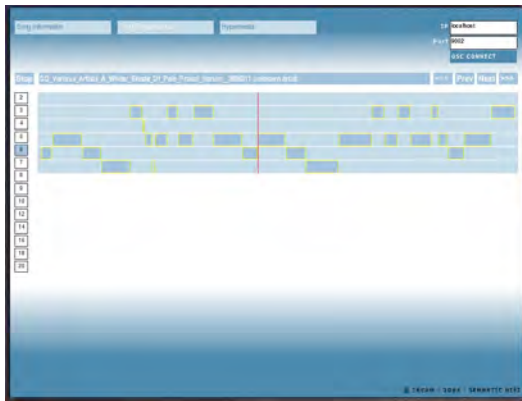
3.5.1 Show Track Metadata and Artist Picture

While a song is playing, the front panel touchscreen can show a picture of the artist of the current playing song alongside with some metadata of a song.



3.5.2 Navigate on song segments

The HiFi System allows the user to navigate within a track graphically. With this feature, the user can view the structure of the track on the HiFi System display, e.g. separate blocks are shown for the introduction, for the chorus and for each verse. The visualization of the blocks represents the content in terms of timbre, rhythm, chords and its structure along time. Using these blocks the user can skip to a particular section within the track.



The user can also select the depth / number of levels of song segments which should be shown by selecting a level (numbers) on the left.

