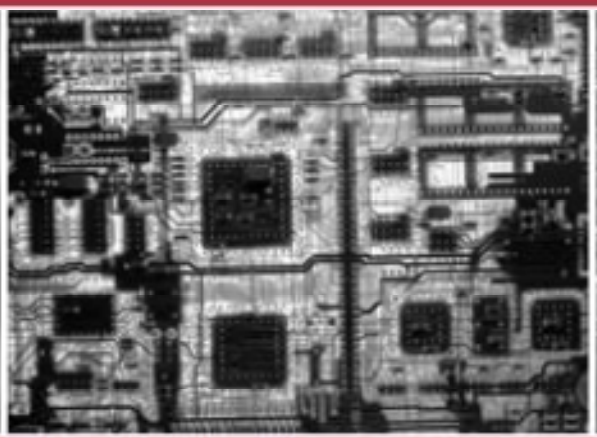
Automatic Classification of Audio Data

*Carlos H. C. Lopes, Jaime D. Valle Jr. & **Alessandro L. Koerich***

IEEE International Conference on Systems, Man and Cybernetics
The Hague, The Netherlands
October 2004

Pontifical Catholic University of Paraná
Department of Computer Science

Alessandro L. Koerich, *Ph.D. ing.*
alekoe@ppgia.pucpr.br

Motivation

- The amount of multimedia now available on–line has created a surge for efficient tools to organize and manage such a huge amount of data.

- Digital music is one of the most important data types distributed in the web.

- How to effectively organize and process such large variety and quantity of musical data to allow efficient indexing, searching and retrieval is a real challenge.

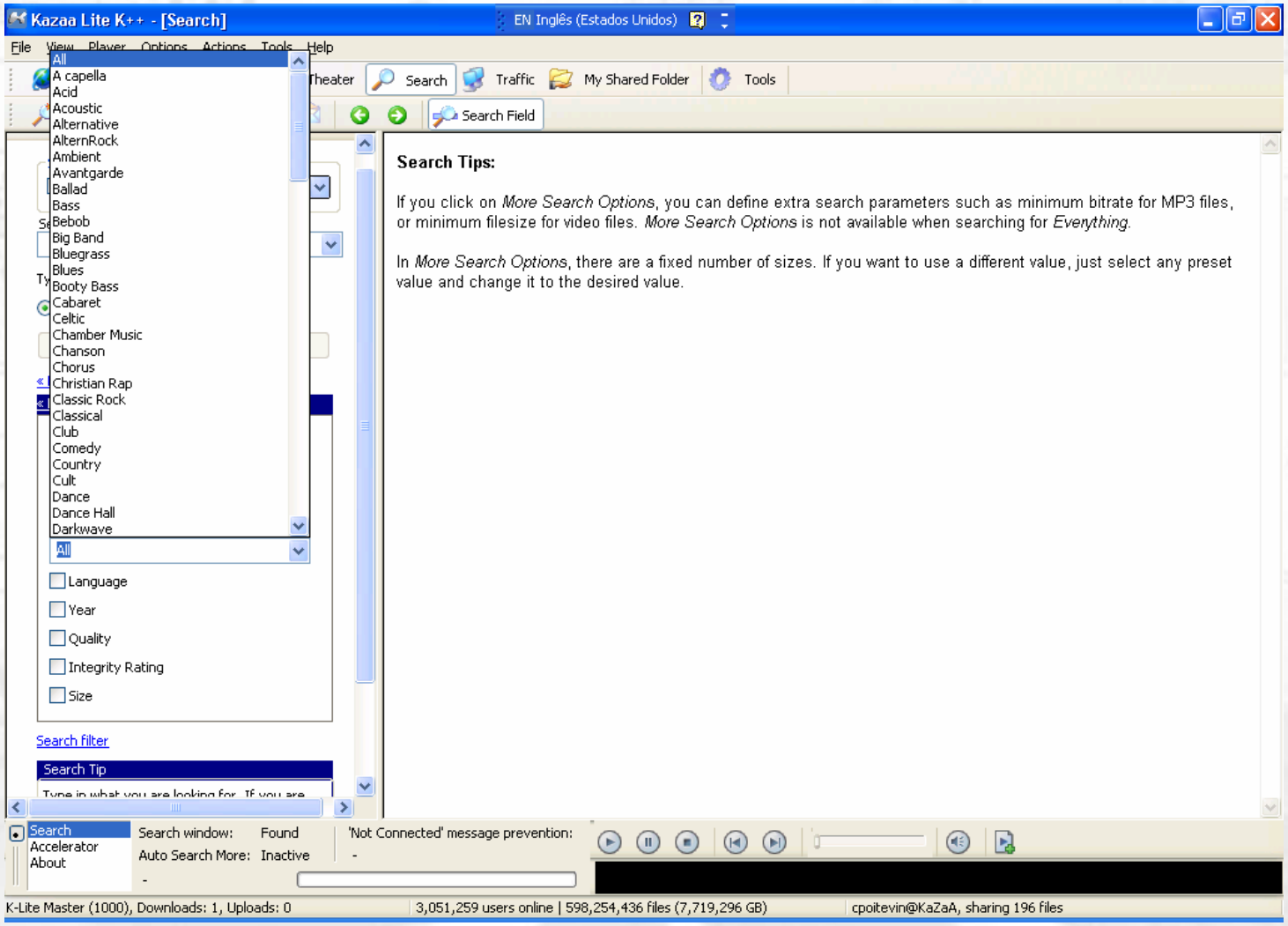
Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Motivation

- At present, multimedia data is usually classified based on textual meta–information.
- While such information is very useful for indexing, sorting, comparing and retrieval, it is manually generated.
- Extracting the information through an automatic and systematic process might overcome such problems.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Motivation


Kazaa Lite K++ - [Search]
EN Inglês (Estados Unidos)
File View Player Options Actions Tools Help
Theater
Search
Traffic
My Shared Folder
Tools
Search Field
All
A capella
Acid
Acoustic
Alternative
AlternRock
Ambient
Avantgarde
Ballad
Bass
Bebob
Big Band
Bluegrass
Blues
Booty Bass
Cabaret
Celtic
Chamber Music
Chanson
Chorus
Christian Rap
Classic Rock
Classical
Club
Comedy
Country
Cult
Dance
Dance Hall
Darkwave
All
Language
Year
Quality
Integrity Rating
Size
Search filter
Search Tip
Search Tips:
If you click on More Search Options, you can define extra search parameters such as minimum bitrate for MP3 files, or minimum filesize for video files. More Search Options is not available when searching for Everything.
In More Search Options, there are a fixed number of sizes. If you want to use a different value, just select any preset value and change it to the desired value.
Search
Accelerator
About
Search window: Found
Auto Search More: Inactive
'Not Connected' message prevention:
K-Lite Master (1000), Downloads: 1, Uploads: 0
3,051,259 users online | 598,254,436 files (7,719,296 GB)
cpoitevin@KaZaA, sharing 196 files

Proposal

- Musical genre is an important description that has been used to classify and characterize digital music and to organize the large collections available on the web

- Musical genres are categorical labels created by humans to characterize music clips.

- These characteristics are related to the instrumentation, rhythmic structure, and harmonic content of the music.

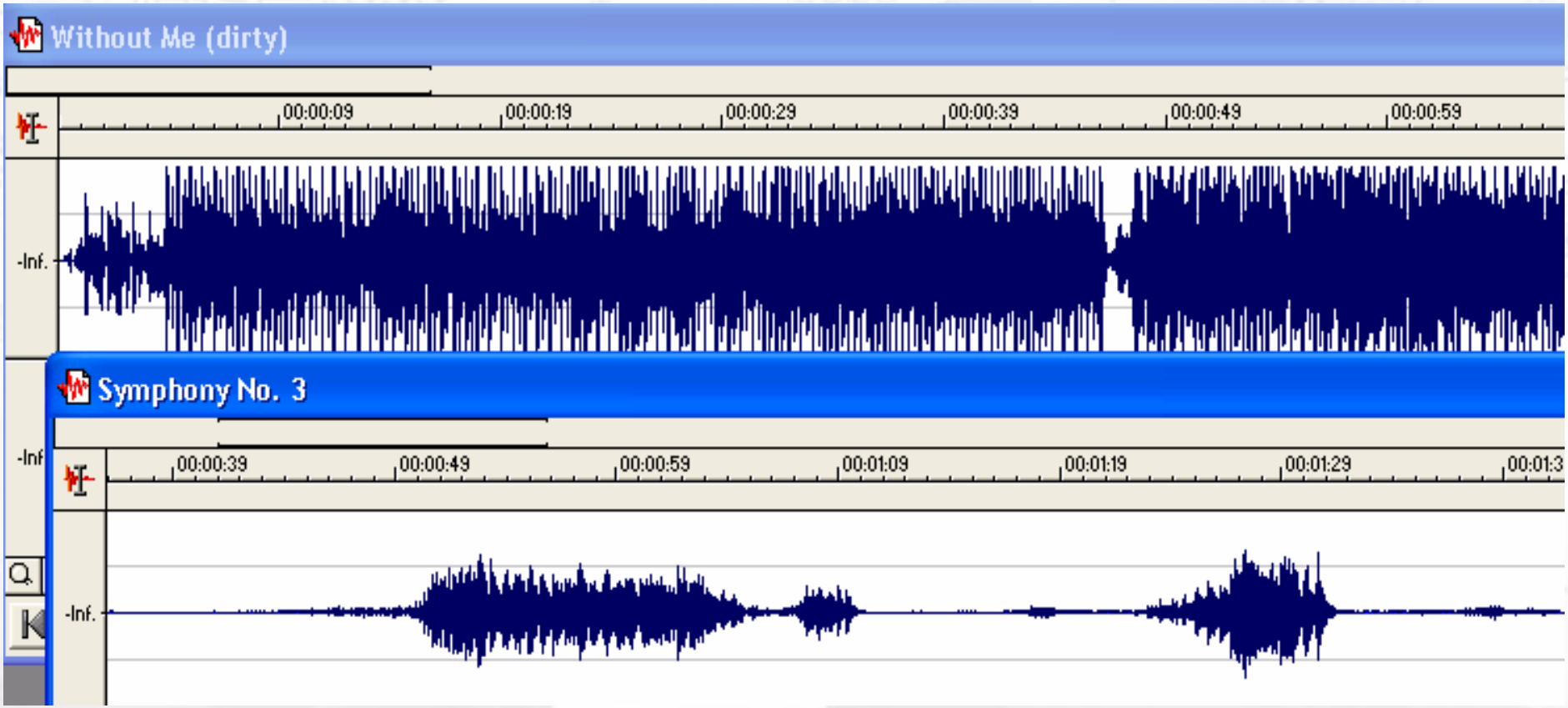
Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Novelty

- A novel approach for content–based musical genre classification based on the *combination of classifiers*.

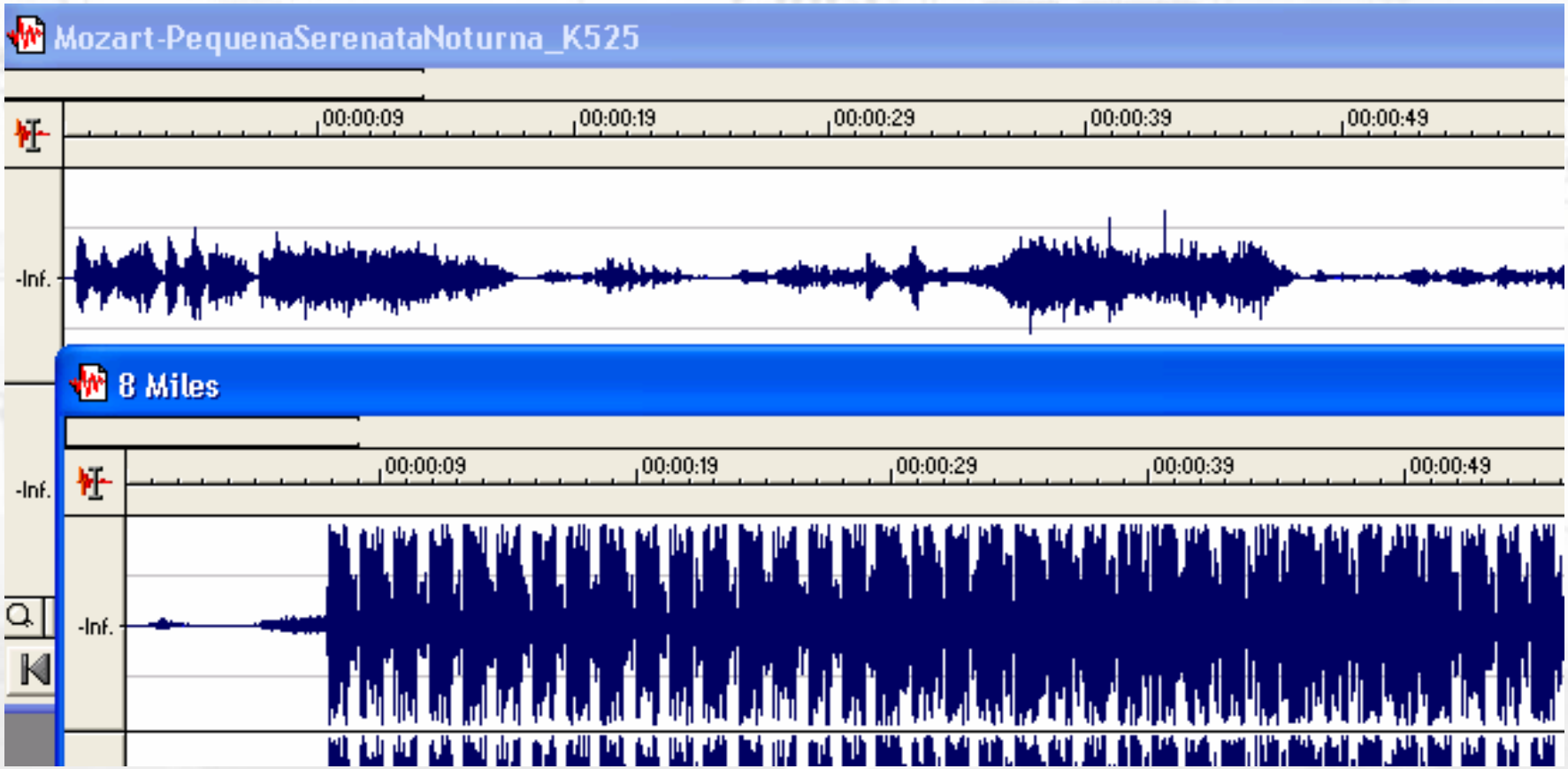
Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Content-Based Approach



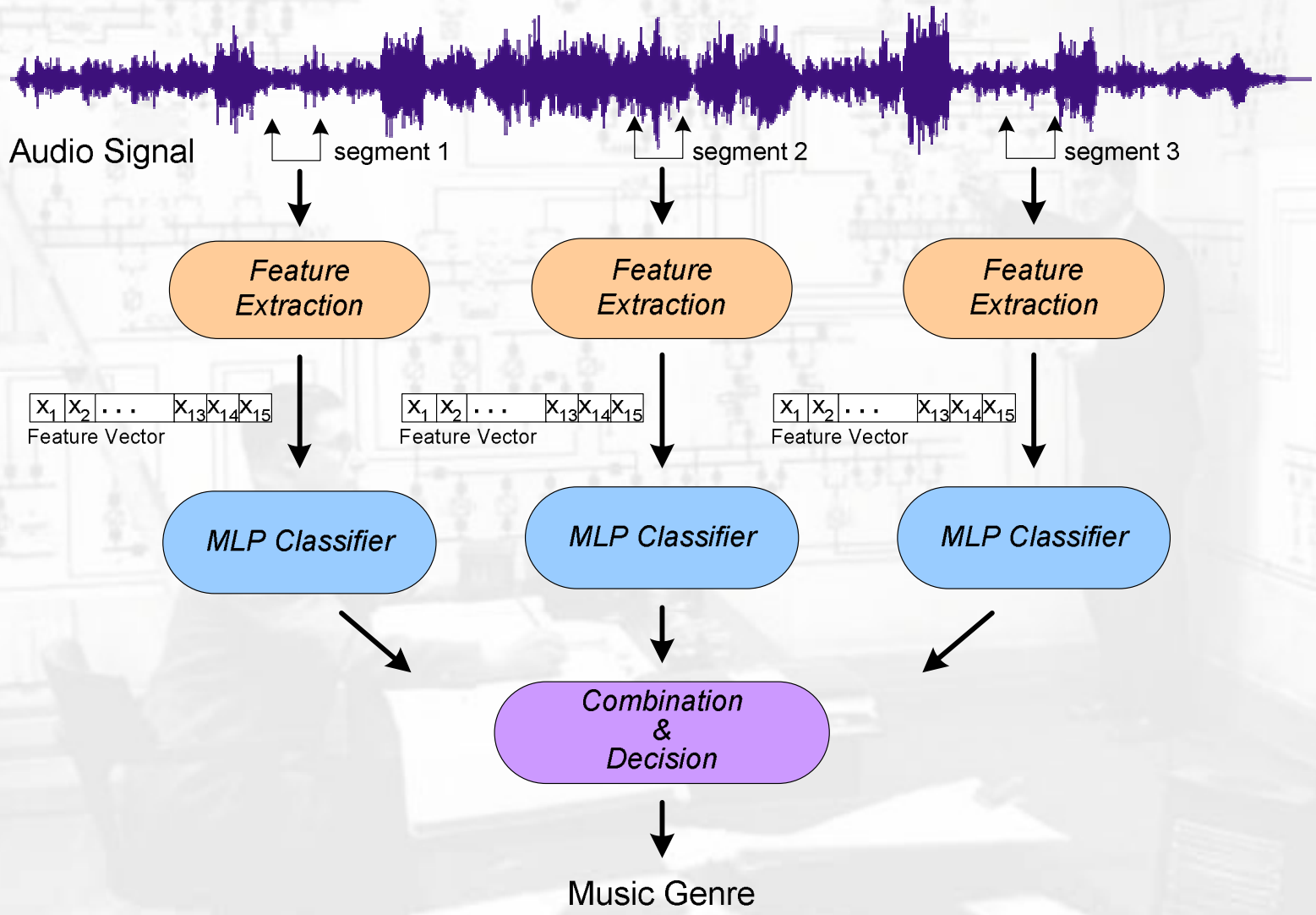
Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Content-Based Approach



Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Overview



Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Feature Extraction

- We have considered the problem of content–based musical genre classification as a pattern classification problem.

- In such a way $\rightarrow$ Extract relevant features from music clips

- Feature extraction is the process of representing a segment of audio by a compact but descriptive vector.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Feature Extraction

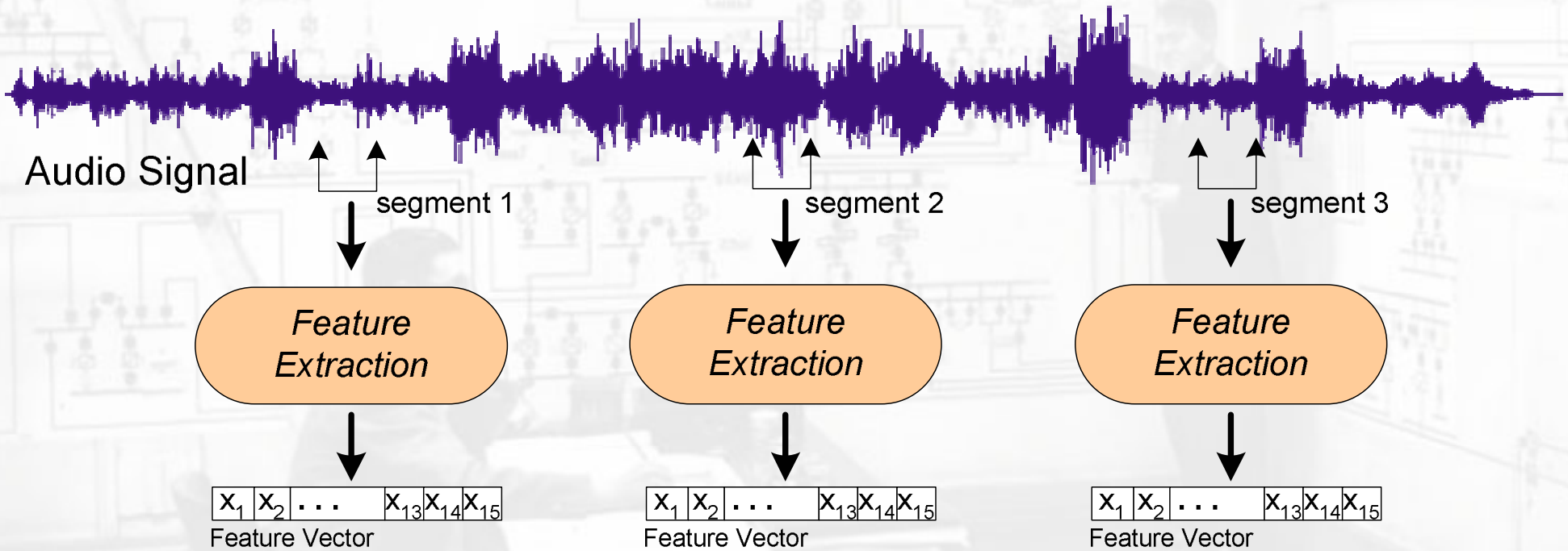
- Since digitized music in good sound quality has an 1MB/minute rate, it would be very time consuming to extract the feature vector from the whole music.

- In such a way feature extraction is carried out only on segments of the music clip.

- Three segments are chosen according to the duration and bit rate of the music.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Feature Extraction



Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Feature Extraction

- The feature set used in this paper was originally proposed by Tzanetakis et al. 2002.

- Two different types of features:
  - musical surface features: mean and average of the spectral centroid, flux, zero–crossing rate, and low energy.
  - beat–related features: relative amplitudes and beats per minute.

- These features form 15–dimensional feature vectors.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Classification Problem

- The basic problem in musical genre classification is:

- Given a music clip represented by a feature vector $X = (x_1\, x_2 \ldots x_D)$ where D is the dimension of the vector, assign a class, i.e. a musical genre $g \in G$ that best matches to the input vector.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Classification

- Instance–based method: k–nearest neighbor (*kNN*) algorithm.

- A multilayer perceptron (MLP) classifier with one hidden layer trained with the backpropagation algorithm.

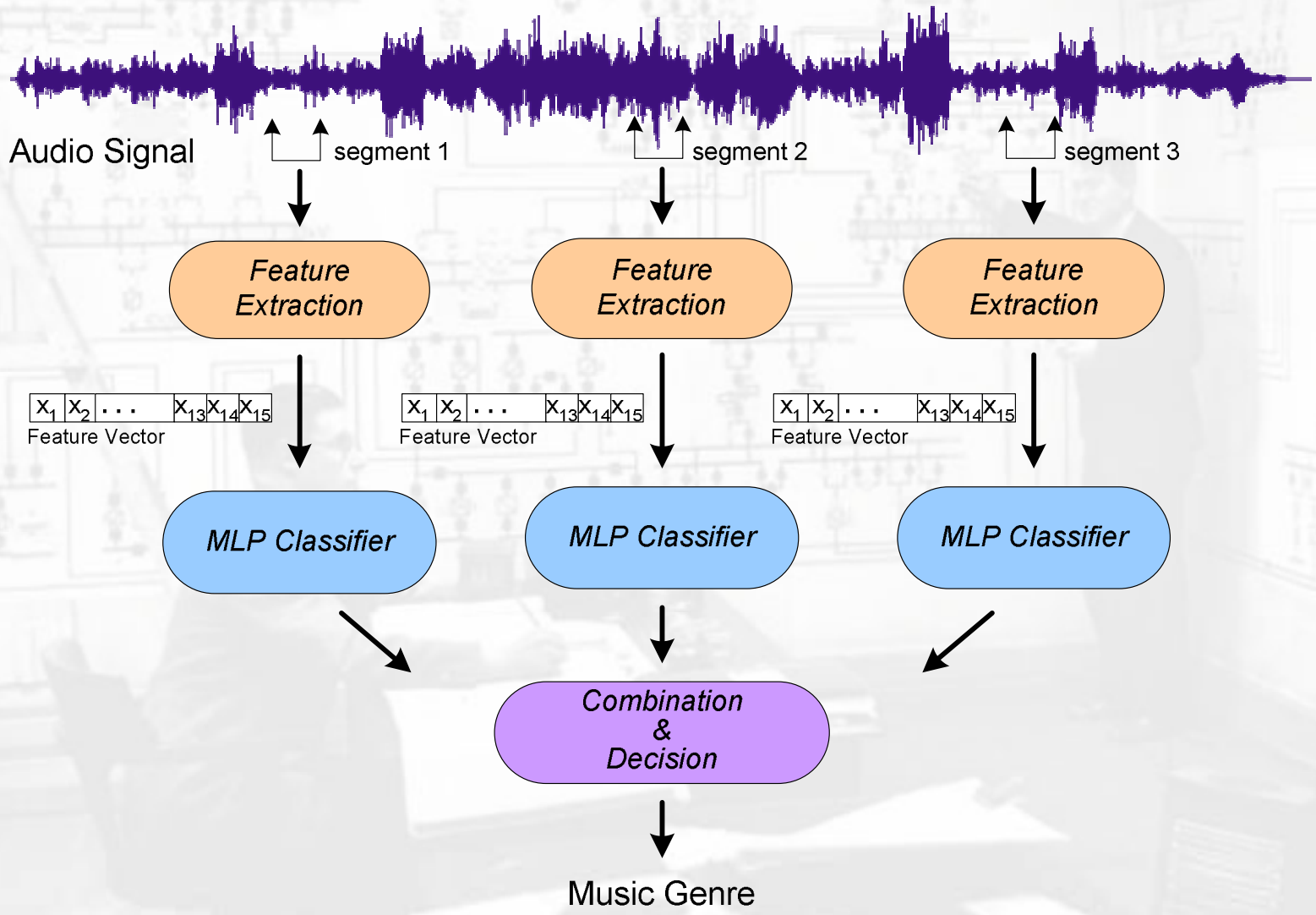
Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Combination

- The three feature vector are extracted from the same music clip.
- The output of the classifiers that take at the input each feature vector can be combined to optimize the classification performance.
- We have considered only the majority voting scheme.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Overview



Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Experimental Results

- Dataset: 414 music clips (207 rock, 207 classic)
  - Training set: 208 samples
  - Validation set: 82 samples
  - Test set: 122 samples

- Three feature vectors were extracted from each music clip $\rightarrow$ 1,242 feature vectors.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Experimental Results

Table 1: Correct musical genre classification rates for single feature vectors and single classifiers

Dataset	Correct Classification Rate (%)				
	1–NN	3–NN	5–NN	7–NN	MLP
Training	—	—	—	—	92.1
Validation	84.0	87.5	90.0	90.0	90.7
Test	83.0	85.5	84.0	83.0	89.5

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Experimental Results

Table 2: Correct musical genre classification rates for the MLP classifier considering three feature vectors for each music clip

	Correct Classification Rate (%)		
Dataset	Segment 1	Segment 2	Segment 3
Training	90.1	92.1	92.2
Validation	86.3	90.7	88.9
Test	85.5	89.5	88.7

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Experimental Results

Table 3: Correct musical genre classification rates for the k–NN classifier considering three feature vectors for each music clip

Dataset	Correct Classification Rate (%)											
	Segment 1				Segment 2				Segment 3			
	1	3	5	7	1	3	5	7	1	3	5	7
Validation	85.5	91.5	92.5	90.5	84.0	87.5	90.0	90.0	80.5	85.0	88.0	90.5
Test	78.0	79.0	76.0	74.0	83.0	85.5	84.0	83.0	79.0	84.5	82.5	80.5

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Experimental Results

Table 4: Correct musical genre classification rates for the combination of classifiers output using the majority voting rule

Dataset	Correct Classification Rate (%)				
	1-NN	3-NN	5-NN	7-NN	MLP
Validation	83.3	89.1	84.2	79.9	91.3
Test	82.3	86.3	83.1	81.5	90.3

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Conclusion

- Automatic musical genre classification is a difficult pattern recognition task.

- We have presented a novel approach to musical genre classification that combines three feature vectors extracted from different regions of music clips.

- The feature vectors are combined at classification level through the combination of the outputs of single classifiers.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br

Conclusion

- A slight improvement in the correct musical genre classification was achieved.

- The combination rule used is very simple.

- Future work will include other combination strategies that take into account the confidence scores provided by the classifiers as well as a rejection mechanism to further improve the reliability of the system.

Alessandro L. Koerich
alekoe@ppgia.pucpr.br