

## Visualization of Wine Attributes

Yuanxun Mei

# **Master Thesis**

## **Visualization of Wine Attributes**

Yuanxun Mei

2009

Department of Computer Science  
School of Mathematics and Systems Engineering (MSI)  
Växjö University

Supervisor:  
Prof. Dr. Andreas Kerren

## Abstract

As the development of the Internet and the rapid increase of data, information visualization is becoming more and more popular. Since human eyes receive visual information very quick and easy, the visualization can make complex and large data more understandable.

Describing sensory perceptions, such as taste, is a challenging task. For a customer, the visualization of the taste of a specific wine together with the other wine attributes such as color and grape type would help him/her choose the right one.

In the thesis, two suitable representations of wine attributes are implemented. And, the final system contains two parts. One is a user interface generating his/her fingerprint based on the two representations. The other one is generating the fingerprints of all wines in a database, and save these fingerprints as images. If the user compares his/her wine fingerprint with the wine images of the database, he/she possibly find the right one much easier.

**Keywords:** Visualization, Fingerprint, Visual representations, Wine, Hierarchical structures, Treemaps

## Acknowledgments

I would like to show my deepest gratitude to my supervisors, Prof. Dr. Andreas Kerren, a respectable, responsible and resourceful scholar, who has provided me with valuable help and guidance in every stage of the thesis work. Without his help, I could not have completed my thesis. Furthermore, my greatest gratitude goes to PhD-student Ilir Jusufi, who helped us with the visualization representation idea and gave me good advice during implementation. Finally, I would thank the persons who helped us to finish the evaluation of the system.

**Note:** In the report, the word “we” means Andreas Kerren, Ilir Jusufi and I. The word “I” means only myself.

# Contents

1. Introduction.....	1
1.1 Background .....	1
1.2 Motivation .....	1
1.3 Goal.....	2
1.4 Method.....	2
1.5 Report Structure .....	3
2. Research Review.....	5
2.1 Gathering Wine Attributes .....	5
2.2 Related Visualization Work.....	10
2.3 Summary .....	15
3. Analysis and Design.....	16
3.1 Requirements.....	16
3.1.1 Functional Requirements .....	16
3.1.2 Non-functional Requirements .....	18
3.2 Use Cases .....	19
4. Build the Database .....	22
4.1 Data Attributes .....	22
4.2 Database Details .....	23
5. Two Visualization Approaches.....	25
5.1 Circular Approach.....	25
5.2 Treemap Approach.....	27
5.3 Interface Scenarios.....	28
6. Visualization and Interaction .....	30
6.1 Circular Approach: Generate Fingerprint Based on Balloon Tree Layout ....	30
6.2 Treemap Approach: Generate Fingerprint Based on Treemap Layout .....	33
6.3 User Interface .....	34
6.4 Database Fingerprints Generating .....	36
7. Implementation .....	38
7.1 Tools, Language, and Libraries .....	38
7.2 Class Description.....	39
7.2.1 Class Description of Circular Approach .....	39
7.2.2 Classes Description of Treemap Approach.....	41
7.2.3 Classes of the User Interface.....	42
7.2.4 Classes of Database Generating System .....	43
8. Evaluation .....	45
9. Conclusions.....	47
9.1 Achievement .....	47
9.2 Experience.....	48
9.3 Future work.....	49
References.....	50
Appendix.....	52
A.Invitation Email.....	52

B.Information Sheet .....	52
C. Questionnaires .....	53

# List of Figures and Tables

Figure 1.1: Sir Edward Playfaire's Circles [26] .....	1
Figure 1.2: A wine label, taken from [1] .....	2
Figure 1.3: Fingerprint generated by using length information based on books of Jack London and Mark Twain, taken from [3].....	3
Figure 2.1: A section of <i>the Aroma Wheel</i> invented by Ann C. Noble detailing the fruity aromas that are most commonly encountered in table wines, taken from [8] .....	7
Figure 2.2: Red wine <i>aroma wheel</i> , taken from [9].....	8
Figure 2.3: White wine <i>aroma wheel</i> , taken from [9] .....	8
Figure 2.4: Hierarchical tree of aromas .....	9
Figure 2.5: Wine production by country, taken from [11] .....	10
Figure 2.6: Chernoff Face, taken from [13].....	11
Figure 2.7: Dr. Eugene Turner's award-winning map .....	12
Figure 2.8: Three visualization type about map, taken from [15] .....	13
Figure 2.9: World Map visualization locating USA .....	13
Figure 2.10: Balloon tree layout visualization from JUNG [17] .....	14
Figure 2.11: Radial Space-Filliping, taken from [18].....	15
Figure 2.12: Tree-map example, taken from [18].....	15
Figure 3.1: Sequence of Use Case 1 .....	19
Figure 3.2: Sequence of Use Case 2 .....	20
Figure 4.1: Database structure .....	22
Figure 4.2: The final database, filled with ten items .....	24
Figure 4.3: Grape Type table in the database.....	24
Figure 5.1: Balloon Tree Layout with the same nodes and structure of Figure 2.4 .....	25
Figure 5.2: A visualization example, taken from [21].....	26
Figure 5.3: Final representation based on balloon layout.....	26
Figure 5.4: A world image, generating by using visualization of Many Eyes [21] .....	27
Figure 5.5: Final representation based on tree-map.....	27
Figure 5.6: User Interface .....	28
Figure 5.7: Aroma selection interface .....	29
Figure 6.1: Fingerprint of Circular Approach, based on balloon tree layout.....	30
Figure 6.2: Circle layout represent "country" variable.....	31
Figure 6.3: Balloon layout represent "aroma" variable .....	31
Figure 6.4: 6 circles represent 7 variables .....	32
Figure 6.5: Represent "name", "producer", "grape type" as text information inside the circle .....	32
Figure 6.6: Fingerprint generated by treemap approach.....	33
Figure 6.7: A number of rectangles representations .....	34
Figure 6.8: User Interface .....	35
Figure 6.9: Pop-out frame for selecting aroma after clicking aroma button .....	36

Figure 6.10: Codes used to connect the database .....	37
Figure 6.11: Database fingerprint generating system .....	37
Figure 7.1: Architecture of wine visualization system .....	39
Figure 7.2: BalloonLayoutDemo-class structure .....	40
Figure 7.3: Initial-class structure .....	41
Figure 7.4: The structure of class-TreeMap.....	41
Figure 7.5: The structure of class-SourthMap .....	42
Figure 7.6: The structure of class- StarburstDemo .....	43
Figure 7.7: The structure of class- Database .....	44
Figure 7.8: The structure of class- ScreenImage .....	44
Figure 9.1: The system simulation.....	44
Figure 9.2: The final system .....	44
Table 3.1: Functional requirements .....	17
Table 3.2: Non-functional requirements .....	17
Table 3.3: Use Case 1 .....	19
Table 3.4: Use Case 2 .....	20
Table 3.5: Use Case 3, 4, and 5.....	20
Table 4.1: Fourteen wine attributes .....	22
Table 7.1: Tools, language, libraries .....	40
Table 7.2: Class-BalloonLayoutDemo description.....	40
Table 7.3: Class-Initial description .....	40
Table 7.4: Class-TreeMap description .....	41
Table 7.5: Class- SourthMap description.....	42
Table 7.6: Class- Rectangle description.....	42
Table 7.7: Class- Basicframe description .....	42
Table 7.8: Class-StarburstDemo description .....	43
Table 7.9: Class- Database description .....	43
Table 7.10: Class- ScreenImage description .....	44
Table 8.1: The Analysis result .....	45



# 1. Introduction

In this chapter, I will introduce the background of information visualization-why it is popular and when it is useful for us. Then, I will explain the goal of my thesis topic and the objective of my work. After that, I will introduce the structure of this report.

## 1.1 Background

Nowadays, the desire to visualize complex and large data sets becomes stronger and stronger. We are not satisfied with the textual data anymore, because it is very hard to understand if it becomes large and complex. How can we find a good way to make this complex data more understandable? Finding a good answer for that question is the aim of information visualization. In general, information visualization is to develop a good visual representation of the original data, which is easy to understand since human eyes acquire information better and faster through visual pictures. One classic example called Sir Edward Playfaire's Circles supports the benefits of information visualization. Figure 1.1 is the circles that was developed by Sir Edward Playfaire in order to stress that, in his day, the Russian Empire was far heavily taxed. The left-hand line is proportional to the gross national product and the right-hand line is proportional to the tax income. This Figure gives us a very straightforward and clear view to know the feature that is the Russian Empire taxed heavily.

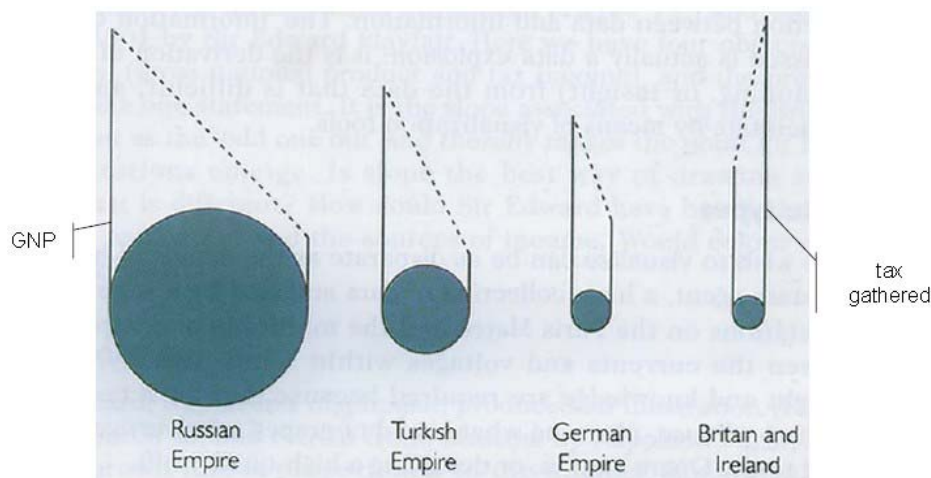


Figure 1.1 Sir Edward Playfaire's Circles [26]

## 1.2 Motivation

If a customer wants to buy wine in a wine store, he/she has to read a brief textual description on the wine labels. Unfortunately, the description does not contain all information of the wine, such as the taste, the color, the grape type and so on. Figure 1.2 shows a wine label, which is one of the most detailed described labels. Even though, the information that you can get from the label is just the wine name, vintage, producer, grape type, alcohol percentage, and bottle size. What can we do if the customer wants to know the description of the taste, the particular color of this wine or information on

the area? Our idea is making a suitable visual image to describe the wine information more concretely and clearly. Thus, the aim of this thesis is to find a way to visualize the textual wine information to help customers get a better understanding of wine information.

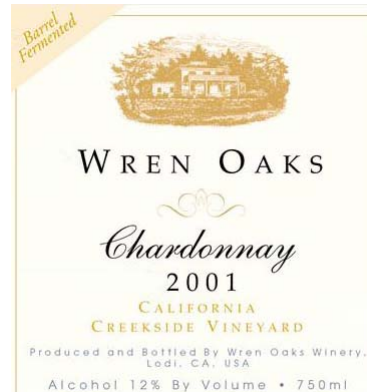


Figure 1.2: A wine label, taken from [1]

### 1.3 Goal

My thesis task is visualizing the textual wine information in such a way that is more understandable and clear for human eyes. The goal of this thesis is to explore and develop a suitable database schema of wine and visualize the attributes of the database properly.

The database should contain most common properties of wine, such as name, producer, taste, color, price and so on. For example, for the taste of a wine is not just sweet or dryness, we need a more concrete classification for it. By visualizing this database we may gain a number of advantages:

1. The possibility to visualize all wine items in our wine database. With the final tools, the text wine items would be visualized into fingerprints (pictures).
2. The ability to perceive all wine attributes easily and quickly. Which means the users of the tools could understand the meaning of the attributes in the fingerprint easily and quickly.
3. The possibility to visualize the user's specific wine item. The user could input his/her preference wine attributes into the tools. Then, the tools will generate the fingerprint according to these input data.

### 1.4 Method

The whole work is divided into four phases: Firstly, built a database, which contains significant wine attributes, such as wine name, vintage, producer, country, rating, and aroma. After that, fulfill some items in the database.

Secondly, found a suitable pattern to represent the attributes of the dataset into a picture, called fingerprint. In computer science, a fingerprinting algorithm is a procedure that maps an arbitrarily large data item to a much shorter bit string, its fingerprint, which uniquely identifies the original data for all practical purposes [2]. In this paper, a fingerprint is a mapping of wine features into a picture, which uniquely

identifies the original features. In order to support a better understanding of our idea, an example of the so-called “literature fingerprinting” [3] will explain it more clear. Figure 1.3 is a visual representation of the length information based on the books of Jack London and Mark Twain: each square represents a chapter of a book. The color of a square represents the length of this chapter. The length of a chapter is decreasing from blue to red. The way to find a suitable fingerprint is the challenge of this paper.

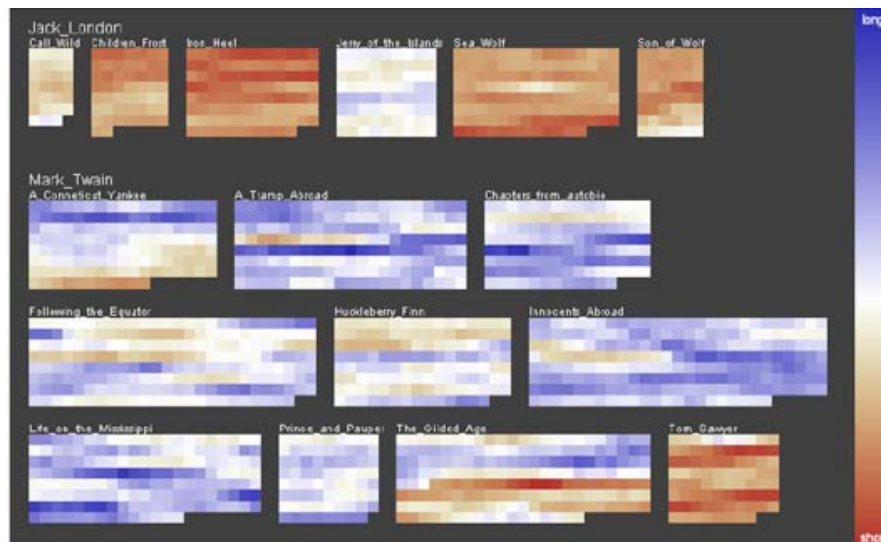


Figure 1.3: Fingerprint generated by using length information based on books of Jack London and Mark Twain, taken from [3]

Thirdly, I developed an efficient visualization tool that represents the wine database and provides the interacting interface, which is used for customer to generate his/her own wine fingerprint. Then, he/she can use this fingerprint to find the right wine by comparing it with the fingerprints in the wine fingerprints dataset. The database fingerprint generating system is used to generating the picture database based on the wine information in the database. After the implementation phase was finished, the user would use it to generate a fingerprint according to his/her preference. For example: If a customer is looking for a red wine that is dry, light taste, ruby red, and from the countries but Australia, he/she will input these features into the tool. Then, he/she will generate the fingerprint according to the four features. He/she can compare my fingerprint with all wine fingerprints in the dataset to find some fingerprints with same or similar parts. These fingerprints of the wines are probably the wines that he/she is looking for.

Lastly, we did an evaluation for the final program. We asked some people to make a use of this program and gave us some comments on it. Then, this system could be improved in the future according to their comments.

## 1.5 Report Structure

In Chapter two, I will explain the search I have studied before implement the final system. It contains the research of finding the wine database schema and proper representation to visualize the wine attribute. Chapter 3 presents the requirements of my final system, which include the requirements of the web-based user interface, user

fingerprints generation system and the database fingerprints generation system. After all requirements are done, I start to explain the process of building the database in Chapter 4. Then, it comes to the essential part of the report, which is a proper visualization representation. Chapter 5 shows two proper representations that we can not decide which one is better. So, we decided to implement both and let the user to decide which one is better in the evaluation part in Chapter 8. Chapter 6 brings out the final system that contains two parts: database fingerprint generation and user fingerprint generation. The implementation details are introduced in Chapter 7. Finally, the conclusion of the thesis and some future works are listed in Chapter 8.

## 2. Research Review

In this chapter, I will present the researches I have done, both on the wine information and related visualization examples. In section 2.1, I will explain the common properties of wine and how to describe them. In section 2.2, I will introduce the related visualization work that inspires the final visualization of this thesis.

### 2.1 Gathering Wine Attributes

Wine is a very common beverage just like coke and juice. Archaeological evidence suggests that the earliest production of wine, made by fermenting grapes, took place in sites in Georgia and Iran, from as early as 6000 BC [4]. Nowadays, wine is becoming a commodity for most people. This thesis concentrates on the consumers' view to describe a wine, and not go to the deep point of the wine experts and specialists.

How do the consumers describe wine? When a consumer want to buy a wine, the name of the wine is firstly concerned if he/she knew it. But if he/she does not know the name in advance, he/she maybe has a personal favor to some wine properties, such as producer, vintage, producing country, region, grape type, color, taste, and price. Moreover, an experienced wine consumer may also consider the storage type, the rating marked by the buyers, the viscosity, and the complexity of a wine. By putting all the properties together, 14 properties are used for building our database.

Among these properties, name, vintage, producer are textual attributes. Grape type, also called grape variety, is classified in two parts in wine world: red grapes and white grapes. Normally, red grapes ferment into red wine and white grape ferment into white wine. However, in some cases, a white wine can be made from red grapes by only taking the flesh and not using the skin of red grapes. For not confusing the potential users of the final tool, the grape type property in the report is simply classified into red grape and white grape. It has 467 grape varieties, red grape and white grape [5]. They are ordered alphabetically.

**Color** is an important property. Wine has three different colors in general, which are red, rose, and white. However, they are not enough for some experienced consumers to describe the detailed color of a wine. Table 2.1 is the sale of wine colors [6]. It has a very proper and detailed description of wine color. By combing it with the other wine color description, and use the final scale for the color attribute in the database of this report. In addition, this description may embed into the final tool to help the users find their desired color.

White wine	Greenish yellow	This is a pale yellow with green highlights, which tend to diminish over time. This color is found in fresh young wines, sometimes from grapes that have been picked rather early
	Straw yellow	This is the most common shade of light yellow, with sometimes quite significant differences in intensity (more or less deep).
	Golden yellow	This is a deep yellow that one finds from certain grape varieties. It is also often to be found in the highlights of full-bodied white wines that have a certain aging potential or have been matured in wood, or those whose grapes have undergone a period of drying
	Amber yellow	This is the typical color of certain wines made from dried grapes, of fortified wines and, in any case, of wines that are made from extremely ripe grapes.
Rose wine	Pale pink	This recalls the color of rose petals or of peach blossom.
	Cherry-red	Fairly deep pink, which recalls the color of certain varieties of cherries.
	“Claret”	More red than pink, but without the intensity of a red wine.
	Onion skin	A deep pink, with highlights that tend towards orange.
Red wine	Purplish red	This is intense and tends towards violet: it is typical of young wines.
	Ruby red	a deep red reminiscent of rubies that is typical of wines that should be drunk relatively young. It in fact indicates the age at which such wines are at their best.
	Garnet red	A color that tends towards blood red. It is the first sign of maturity in a wine: this shade indicates that the wine has been aged for at least a couple of years and has attained a good degree of maturation. Some wines display garnet nuances even in youth, as is the case with those made from Pinot Nero or Nebbiolo grapes.
	Brick red	Its orange tones remind us of the color of bricks. This shade is the typical indicator of aging: it is that of the right stage of development for fine wines that reach their peak after a number of years, but a sign of decrepitude for those that do not have the structure to keep.

Table 2.1: The scale of wine color

**Wine taste** is the sensory examination and evaluation of wine. It is the most interesting and important property to visualize in this thesis. Since taste is very complicated and can not be quantized, the simplest way to describe the taste of wine is sweet or dry. It is not sufficient for most consumers who always want to know more about the taste. We have searched many books and webpage about the taste. Finally,

we found a professional and detailed description, which is the aroma of wine. The human tongue is limited to the primary tastes perceived by taste receptors on the tongue-acidity, bitterness, saltiness, sweetness, and umami. The wide array of fruit, earthy, floral, herbal, mineral and woody flavor perceived in wine is derived from aroma notes interpreted by the olfactory bulb [7]. If we find the description of the aroma of wine, we will find a good structure for this attribute. Fortunately, *the aroma wheel* [7] developed by Ann C. Noble of University of California, Davis, is very helpful for my thesis. Figure 2.1 is a section of the Aroma Wheel invented by Ann C. Noble. The wheel breaks down wine aromas into 12 basic categories and then further sub-divided into different aromas.

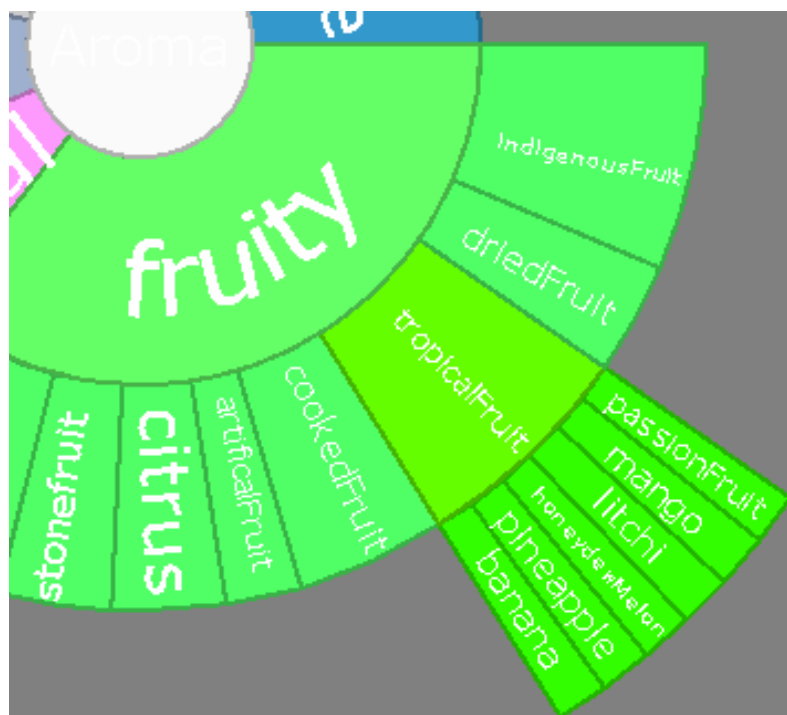


Figure 2.1: A section of *the Aroma Wheel*, taken from the final tool

The German Wine Institute has created a special German language version of the Aroma Wheel meant to be specially adapted to German wines, with one wheel for white wines and one wheel for red wines [9]. This thesis used this *aroma wheel*, because I can not find the original one of invented by Ann C. Noble. The so-called aroma wheel in Figure 2.2 shows the components of smell and taste of red wine.



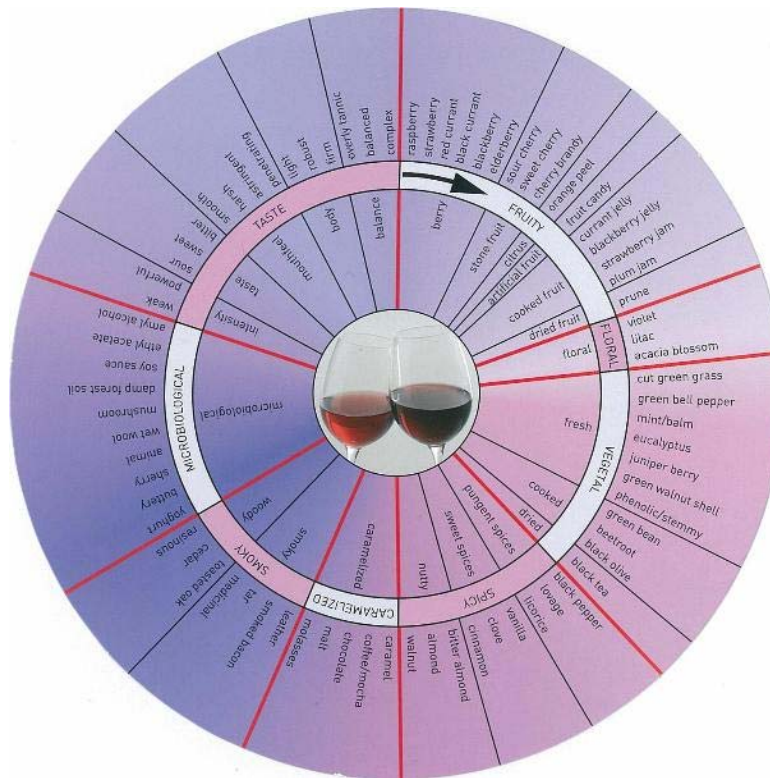


Figure 2.2: Red wine *aroma wheel*, taken from [9]

This red wine *aroma wheel* is subdivided into eight segments, which, in an inner circle, are broken down into twenty broad classifications of smell and taste. Meanwhile, there is a white wine *aroma wheel* (Figure 2.3).

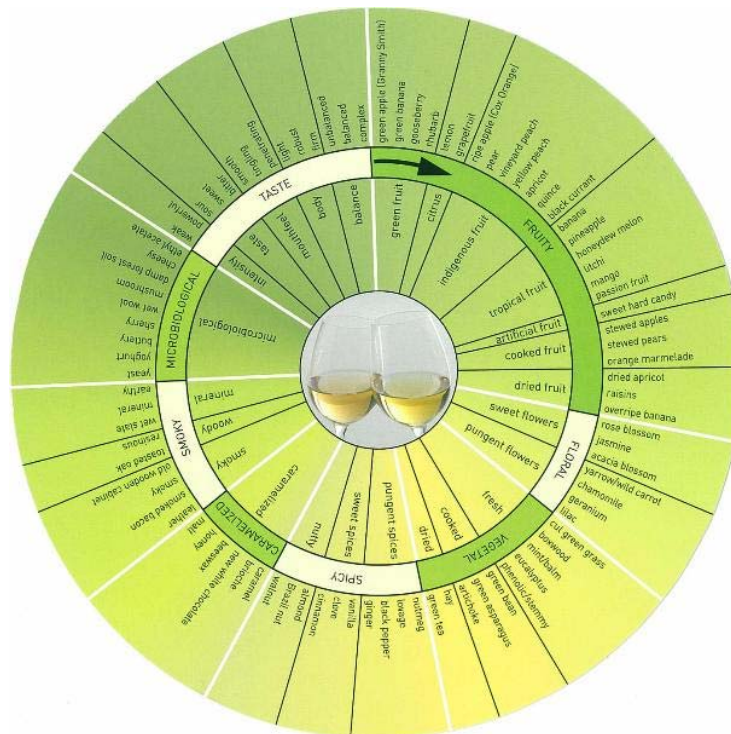


Figure 2.3: White wine *aroma wheel*, taken from [9]



We combined these two wheels together to build a new wine aroma tree, which describes the “aroma” attribute in our database in the next chapter. We converted the wheel into a tree structure. Figure 2.4 is part of the hierarchical tree, three levels, that describes the aroma attribute.

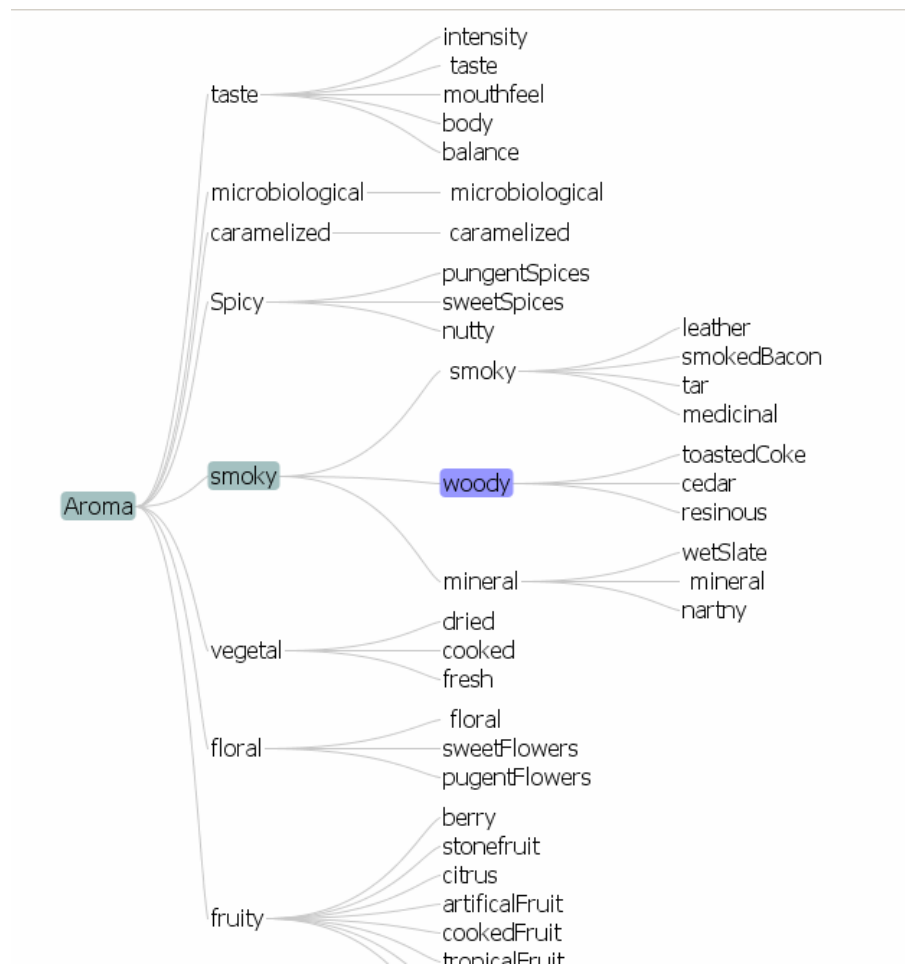


Figure 2.4: Hierarchical tree of aromas

**Country** is also a basic and important property. Wines from different countries have different taste and color. For example, wine from Australia is sweeter than wine from Germany because the illumination intensity in Australia is stronger than in Germany. There are around 70 countries producing wine [10]. The main countries are France, Italy, Germany, Spain, South Africa, USA, Chile, New Zealand. Figure 2.5 is the top 20 wine producing countries from the Food and Agriculture Organization (FAO) [10], which is an agency of the United Nations.

Rank	Country (with link to wine article)	Production (tonnes)
1	 Italy	5,329,449
2	 France	5,056,648
3	 Spain	3,934,140
4	 United States	2,232,000
5	 Argentina	1,564,000
6	 China	1,300,000 <sup>[1]</sup>
7	 Australia	1,274,000
8	 South Africa	1,157,895
9	 Germany	1,014,700
10	 Chile	788,551
11	 Portugal	576,500
12	 Romania	575,000
13	 Russia	512,000
14	 Hungary	485,000
15	 Greece	437,178
16	 Brazil	320,000
17	 Austria	258,000
18	 Ukraine	240,000

Figure 2.5: Wine production by country, taken from [11]

**Storage barrel** is also a factor that the consumers would consider, because the barrel will affect the color, flavor, tannin profile and texture of the wine. So some consumers only prefer the wine where is storing in a wood barrel. This attribute is classified into wood barrel and metal barrel.

**Price** is obvious a significant factor affecting the consumer. Also, some consumer would like to know the rate of wine by the wine reviewers. These two attributes are numerical.

## 2.2 Related Visualization Work

There is not fully related work about visualizing wine information. However, I will introduce some visualization examples that stimulated the final visualization approach.

**Chernoff Faces** were introduced by Herman Chernoff in 1973 [12], which is a visualization technique to illustrate trends in multidimensional data. Figure 2.6 illustrate a Chernoff face where one facial feature on the face could represent one dimension. For example, the face width, the level of the ears, the radius of the ears, the length or curvature of the mouth, the length of the nose, etc.

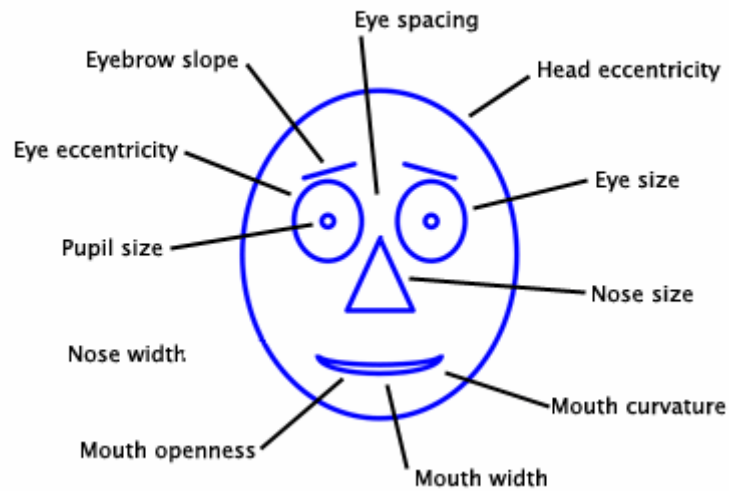


Figure 2.6: Chernoff Face, taken from [13]

It is very lively so that it would interest the observer immediately while he/she look it. In 1979, Dr. Eugene Turner of California State University, Northridge produced a map entitled “Life in Los Angeles with modified Chernoff faces “(Tyner, 1992) [14]. Figure 2.7 shows how Turner uses Chernoff faces to represent the life in Los Angeles. In this example, the life is influenced by four variables, which are affluence, unemployment rate, urban stresses, and percentage of white population. The four variables are mapping into a Chernoff face by facial elements face shape, mouth curvature, eyebrow slope, and face color respectively. We can know the life in difference region of LA easily by reading the faces in that region.

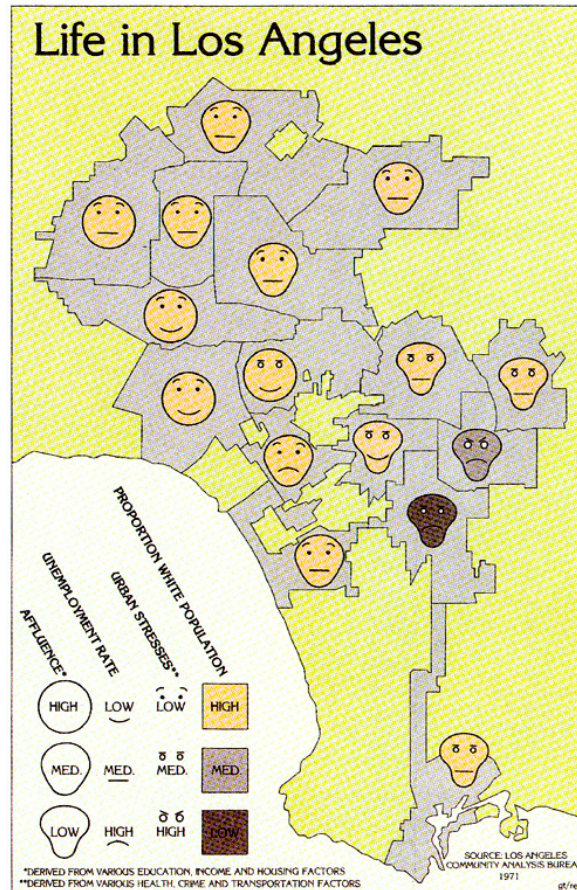


Figure 2.7: Dr. Eugene Turner’s award-winning map

This thesis also has a multi-dimensional database that has 14 variables. At the beginning, I planned to represent some of the variables with Chernoff such as the price, rating, etc. And put the face on a world map that represent the country variable. This idea was intermitted because I have not found the way to visualize the taste variable in Chernoff. This idea could be done in the future work if there is a proper way to combine the taste variable with Chernoff.

*Many eyes* is a website by IBM [15]. It is a place where the users can visualize their own visualization by using the visualization tools in the website. The users just need to define their own data, and choose one visualization type in the website. Then, their visualizations will be done automatically. Figure 2.8 is some map visualization types. I updated my own data which is a country list, then chose the visualization type “World Map” in Figure 2.8. Then, the visualization was done automatically in Figure 2.9.



Figure 2.8: Three visualization type about map, taken from [15]

In Figure 2.9, there is a circle on the location of USA with USA be selected in the data sheet. This idea is suitable for visualizing the country attribute in this thesis. However, the tools of *Many Eyes* can not be used in local application. This idea motivates the final application, which will be introduced in Chapter 5 and Chapter 6.



Figure 2.9: World Map visualization locating USA

Another visualization example is a **balloon tree layout** example in Figure 2.10. It is based on JUNG, the Java Universal Network/Graph Framework, which is a software library that provides a common and extendible language for modeling, analysis, and visualization of data that can be represented as a graph or network. It is written in Java,

which allows JUNG-based applications to make use of the extensive built-in capabilities of the Java API, as well as those of other existing third-party Java libraries [16]. The reason to use it is because it has a good representation on hierarchical structure, which is the hierarchical tree of aroma in my thesis. The tree in Figure 2.10 has four levels. The node in the centre is the root, and the nodes in the same level are clustered inside a circle, which is called balloon here. We perceive the structure of this tree very easy. What is more, it saves much space comparing to a normal tree. In our case, we have a three level aroma tree, and we inspiringly found that the nodes in balloon layout looks just like branches of grapes. So it is used to visualize the aroma tree in one of the representation approaches in Chapter 5.

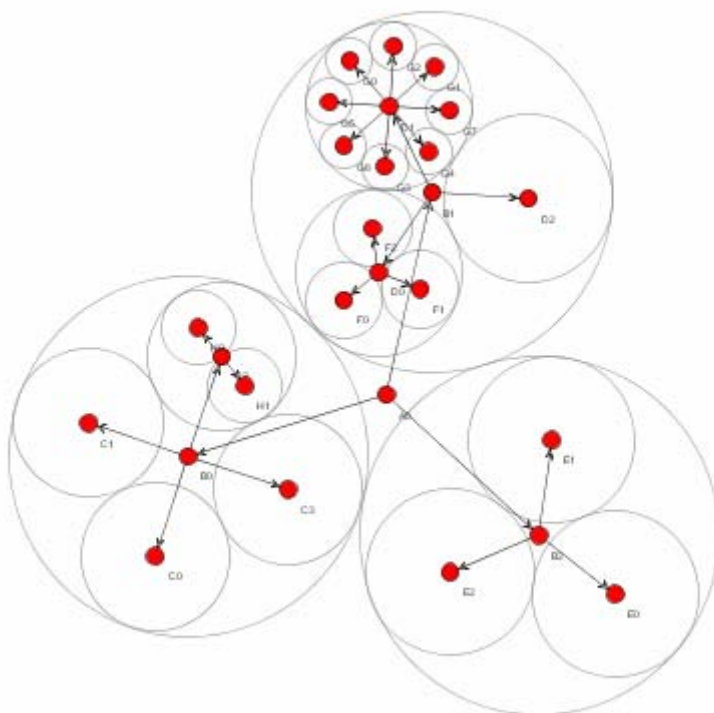


Figure 2.10: Balloon tree layout visualization, taken from JUNG [17]

**Radial Space-Filling** has a similar advantage to the balloon tree, which means they both can visualize hierarchical structure and meanwhile using less space. Figure 2.11 shows the principle of radial space-filling: circle segments areas for the representation of the tree nodes. Since it is similar to the balloon tree layout, but it does not have the property of grapes shape, so the balloon tree is better. If there is a new proves that radial space-filling is better than balloon tree, this idea would be implemented in the future.

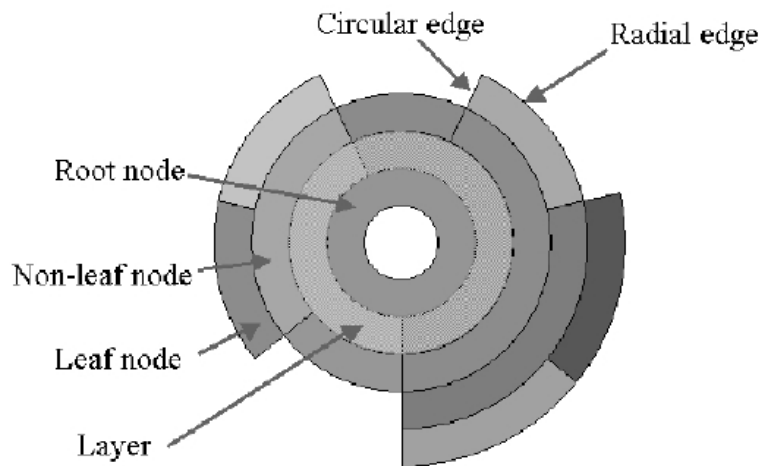


Figure 2.11: Radial Space-Filling, taken from [18]

**Treemap** is a method for displaying tree-structured data using nested rectangles. Each branch of the tree is given a rectangle, which is then tiled with smaller rectangles representing sub-branches. A leaf node's rectangle has an area proportional to a specified dimension on the data [19]. Figure 2.12 shows an example to transfer a hierarchical tree to a tree-map. In the visualization approach 2 in Section 5.2, treemap is the way visualizing the aroma tree.

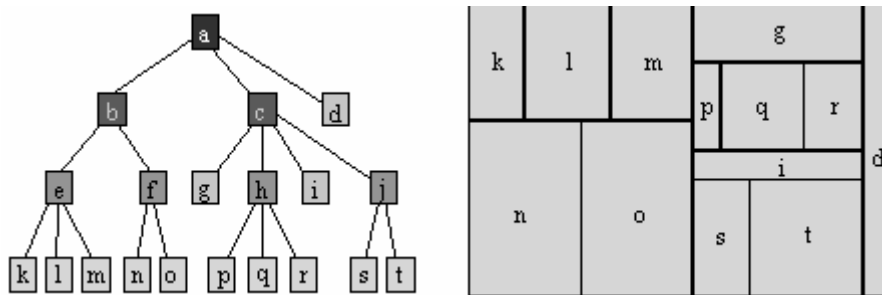


Figure 2.12: Tree-map example, taken from [18]

### 2.3 Summary

In this chapter, I introduced my researches about this thesis at the beginning. The first section is about how to define the property of wine. The significant attribute is aroma that stores in an aroma-wheel. This research was used to build the database in chapter 3. Section 2.2 is about the search on the visualization part, which contains the visualization examples on geographic entity “country” and hierarchical entity “aroma”. Though some of these visualization ideas will not be used to create the representation at this moment, they are interesting and maybe useful for future work.

## 3. Analysis and Design

In this chapter, I present the requirements of the final system, which include the requirements of the web-based user interface, user fingerprints generation tool and the database fingerprints generation tool. Then, I will list what the user can do with our system.

### 3.1 Requirements

Defining the requirements of the application in advance is very important. Thus, one can get a general overview about the application to distribute the work properly according to these requirements. In general, there are two types of requirements: functional requirement and non-functional requirement. According to the scenarios in the previous chapter, I will make a list of requirements that our system should fulfill.

#### 3.1.1 Functional Requirements

In software engineering, a functional requirement defines a function of a software system or its component. A function is described as a set of inputs, the behavior, and outputs [25].

<b>FR1</b>	<b>Add new wine item into database</b>
Importance	Essential
Description	The administrator should be able to import new wine items into the database.
<b>FR2</b>	<b>Connect the database with the system</b>
Importance	Essential
Description	The fingerprint generation component should be able to connect to the database in order to visualize the items inside.
<b>FR3</b>	<b>Visualize the wine items in the database</b>
Importance	Essential
Description	The fingerprint generation component should have an interface which receive the information from database and distribute them into wine variables respectively.
<b>FR4</b>	<b>Input text information for wine name and producer</b>



Importance	Essential
Description	The user interface should have the function to allow the user input the wine name and producer by keyboard.
<b>FR5</b>	<b>Choose grape type from a list</b>
Importance	Essential
Description	The user interface should provide a grapes list from that the user can choose grape type.
<b>FR6</b>	<b>Choose wine color from a list</b>
Importance	Essential
Description	The user interface should provide a color list from that the user can choose specific color.
<b>FR7</b>	<b>Choose country from a world map</b>
Importance	Desirable
Description	The user could choose producing countries of wine visually. More specific, the user interface should provide a world map on that the user can choose the countries with mouse.
<b>FR8</b>	<b>Connect the interface with the fingerprint generation component dynamically</b>
Importance	Essential
Description	The fingerprint generation component can interact with the user interface, it should response quickly while the option in the interface is changing
<b>FR9</b>	<b>Visualizing all wine properties from</b>
Importance	Essential
Description	The fingerprint generation component should be able to visualize all 14 wine properties.
<b>FF10</b>	<b>Saving the database fingerprints as images</b>
Importance	Essential

Description	The system should save the fingerprints of all wine items in the database as images.
<b>FR11</b>	<b>Comparing the fingerprint with the database images</b>
Importance	Desirable
Description	The user can compare his/her fingerprint with all images based on the information of the database.
<b>FR12</b>	<b>Saving the user's fingerprint in the disk</b>
Importance	Desirable
Description	After the user input his/her options through the user interface, his/her fingerprint can save as an image in the local computer.
<b>FR13</b>	<b>Printing the user's fingerprint</b>
Importance	Desirable
Description	After the user input his/her options through the user interface, he can print out the fingerprint by the system.
<b>FR14</b>	<b>Saving the user's preference options</b>
Importance	Desirable
Description	The system can save the user's preference options as a file. The user can generate the fingerprint by loading the file.

Table 3.1: Functional requirements

### 3.1.2 Non-functional Requirements

A Non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors [25]. It is often called quality of a system, such as security, usability, testability, maintainability, extensibility, and scalability etc.

<b>N-FR1</b>	<b>The system should work fast</b>
Importance	Essential
Description	The system should work fast, once the user select the options, these options should appear in the fingerprint immediately.

<b>N-FR2</b>	<b>The system should be web-based application</b>
Importance	Essential
Description	The user can use the system on the Internet. More specific, the user can download the system from a webpage automatically.
<b>N-FR3</b>	<b>The system should work on different browsers</b>
Importance	Essential
Description	The system should be able to work on different browsers (Internet Explorer, Firefox, Opera, Maxthon, etc).
<b>N-FR4</b>	<b>The system should have a good extensibility.</b>
Importance	Desirable
Description	The system should be able to able to add new features and functions in the future.

Table 3.2: Non-functional requirements

### 3.2 Use Cases

Use case is a description of a system's behavior as it responds to a request that originates from outside of that system [27]. In other words, a use case describes "who" can do "what" with the system [27].

<b>UC1</b>	<b>Visualizing the database and save them as images</b>
Description	After the administrator run database fingerprint generation component, the application will connect with the database, navigating the database and return related information to the fingerprint generation component. Then, the application generates all fingerprints and save them as images in local disk.
Precondition	Database is built and fulfilled.
Postcondition	Images are generated in local disk.
Sequence Diagram	Figure 3.1

Table 3.3: Use Case 1

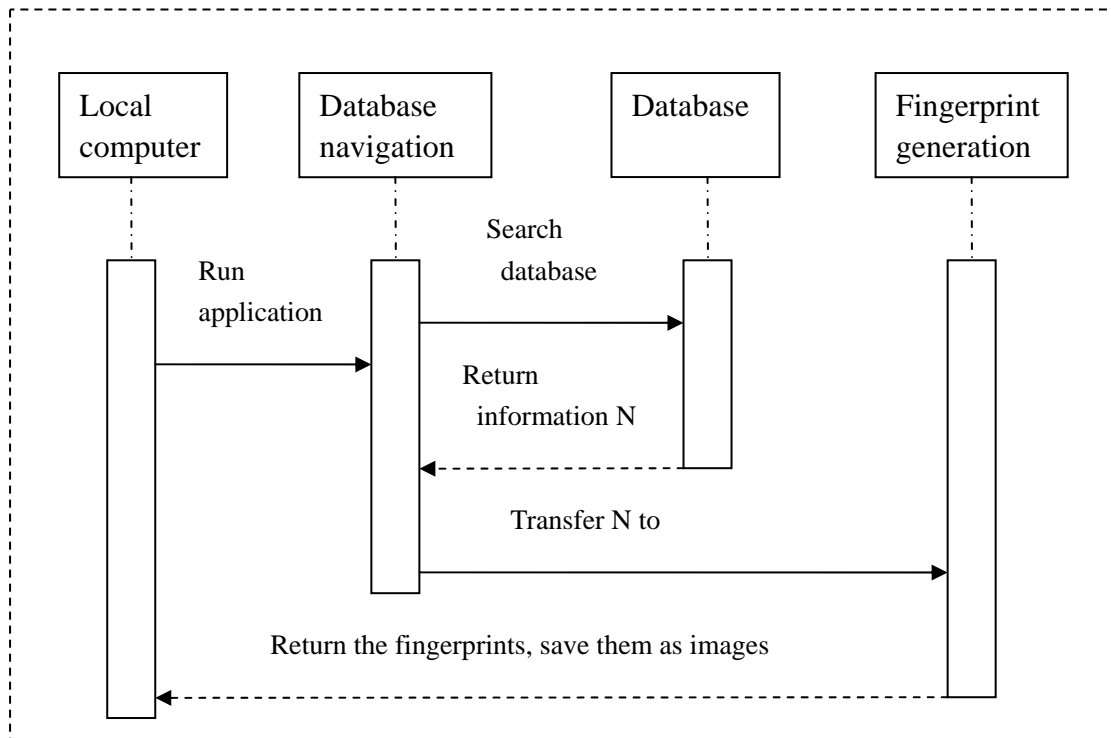


Figure 3.1: Sequence of Use Case 1

UC2	Visualizing the options that the user chose
Description	The user open the webpage of the application, it will download the applet application. Then, the user input the data from user interface. The fingerprint generation component will visualize the data that the user chose.
Precondition	Downloading the application into the user's computer
Postcondition	Generating the fingerprint using the application.
Sequence Diagram	Figure 3.2

Table 3.4: Use Case 2

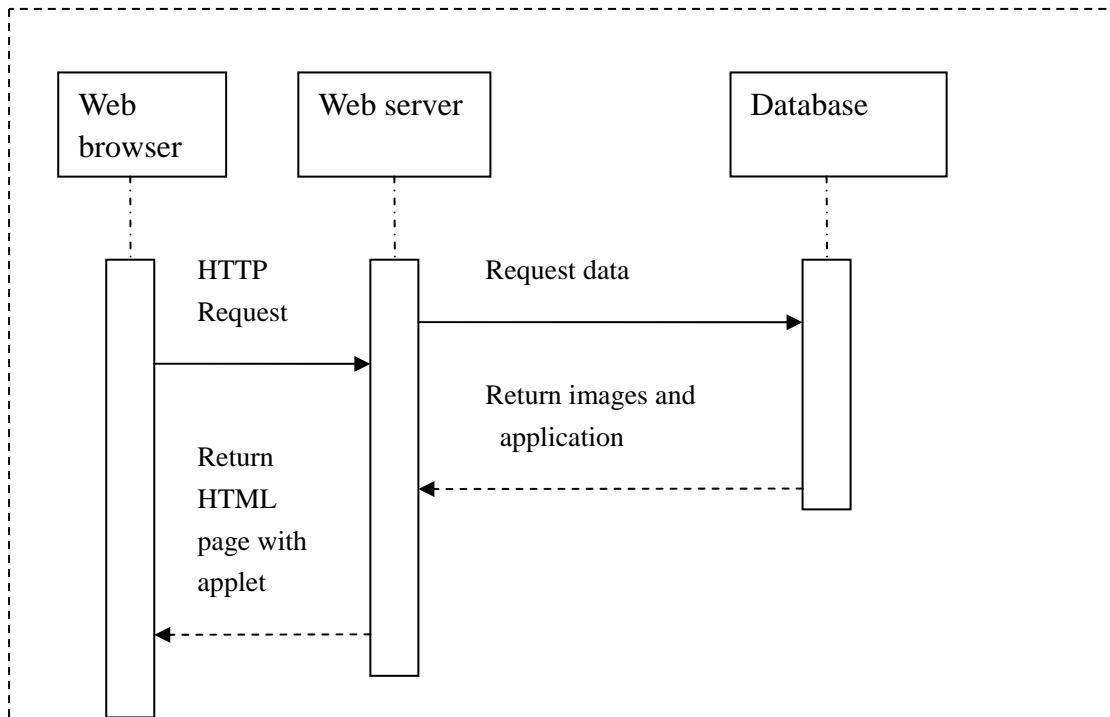


Figure 3.2: Sequence of Use Case 2

<b>UC3</b>	<b>Save the user's fingerprint as image</b>
Description	After the user download the applet application and run it, the application will generate a fingerprint according the user's input. Then, the user could save the fingerprint as an image.
Precondition	Downloading the application into the user's computer and generate the fingerprint using this application.
Postcondition	Save the fingerprint as an image in the user's computer.
<b>UC4</b>	<b>Print the user's fingerprint as image</b>
Description	After the user download the applet application and run it, the application will generate a fingerprint according the user's input. Then, the user could print the fingerprint.
Precondition	Downloading the application into the user's computer and generate the fingerprint using this application.
Postcondition	Print the fingerprint in user's local printer
<b>UC5</b>	<b>Compare the fingerprint, find out the right wine</b>
Description	After the user run the application and generates his/her fingerprint, he/she compare the fingerprint with the images of database, and may find the right wine.
Precondition	Run application and generate the user's fingerprint.
Postcondition	The user may find the right wine.

Table 3.5: Use Case 3, 4, and 5

## 4. Build the Database

I use a MySQL database management system to build my database on a Microsoft windows System. It is easy to use and suitable for a small database. You can find that from the official website of MySQL [28]. I use the Community Server, as that is the freely available version.

### 4.1 Data Attributes

The database is created according to the research in Section 2.1

1	2	3	4	5	6	7
Name	Vintage	Producer	Country	Region	Grape Type	Complexity
8	9	10	11	12	13	14
Category	Color	Aroma	Barrel	Viscosity	Price	Rate

Table 4.1: Fourteen wine attributes

Table 4.1 shows the fourteen attributes. As I mentioned in Section 2.1, the **Aroma** attribute is the most important one, since the wine consumers put a significant consideration on the taste of wine. The focus point is visualizing the aroma attribute. A hierarchical tree (Figure2.4) contains the most common and important wine taste items based on the two wine aroma wheels.

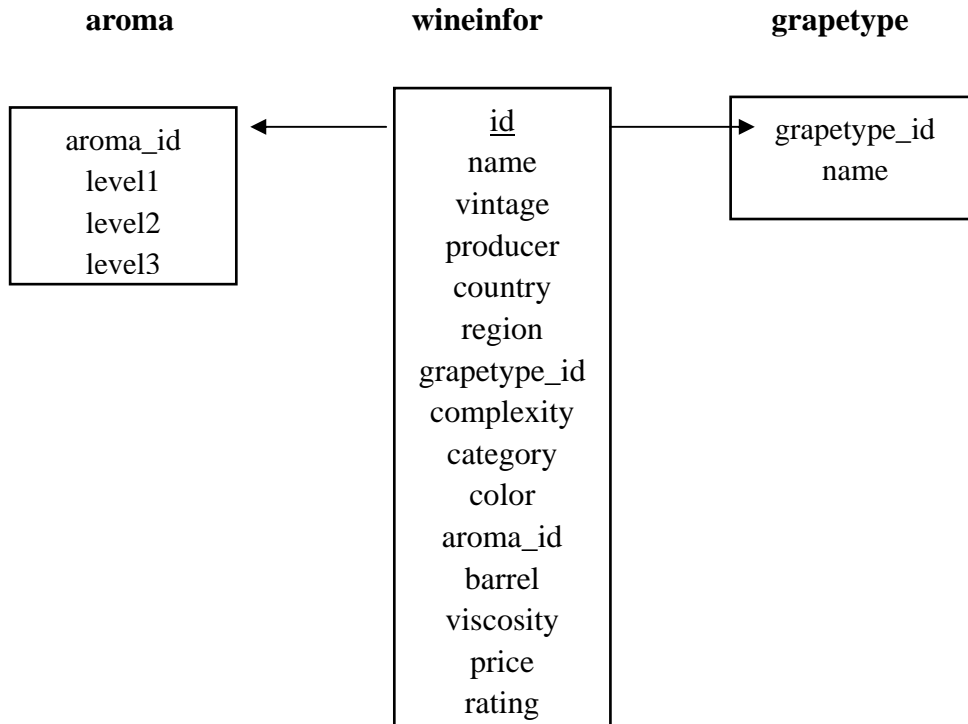


Figure 4.1: Database structure

## 4.2 Database Details

Figure 4.2 is the wine database filled with information. I list the attributes followed with the explanation.

1. Complexity means the intensity of wine, which is divided into light, medium, and heavy.
2. According to the color classification in Section 2.1, in the database Color: 1.Greenish yellow, 2 Straw Yellow, 3 Lime Yellow, 4 Golden Yellow, 5 Old Golden Yellow, 6 Amber Yellow, 7 Pale pink, 8 Cherry red ,9 Clret,10 Onion Skin ,11 Black red, 12 Ruby Red, 13 Chary Red, 14 Purplish Red , 15 Brick Red , 16 Garnet Red.
3. Category: 1.white, 2. rose, 3. Red.
4. Barrel: 1. wood, 2. metal.
5. Viscosity is currently empty, it could be fulfill into the database in the future.
6. Aroma attribute has a list of three-digit number and its related tree has three levels (Figure 2.4), where we ignore the root node. The 100-digit is mapping to the first level of aroma tree, the 10-digit is mapping to the second level, and so on. For example, number 521, 5 means the fifth node of first level that is “smoky”, 2 means the second node of second level that is “woody”, 1 means the first node of the third level that is “toasted cake”. Meanwhile, it is also possible to show only the first and second level by set up the 100-digit to “0”. Then, 520 mean smoky-woody.
7. Grapetype is stored in a table, where a negative value of the grape type id means white grape, positive value of id means red grape. The list of grape varieties is taken from [20], see Figure 4.3.

id	name	Vintage	Producer	Country	Region	Grape Types
1	Gewurztraminer - VT SGN	2002	Hugel	France	Alsace	-91
2	Vega Sicilia - Unico	1998	Bodegas Vega Sicilia	Spain	Spain	197
3	Guigal La Turque Cote Rotie	2001	Etienne Guigal	France	France	152
4	Hermitage	1980	Chave	France	France	170
5	Penfolds Grange	1999	Penfolds	Australia	Victoria	35
6	LAFLEUR	1990	Chateau La Fleur de Gay	France	Bordeaux	29
7	Pinot Grigio Valdadige D.O.C	2007	SANTA MARGHERITA	Italy	Italy	-227
8	Luna dei Faldi Vigneti delle Dolomiti	2000	Cantina La Vis S.c. a r.l.	Italy	Italy	-45
9	Cabernet Franc Lison Pramaggiore	2005	Azienda Agricola Livon	Italy	Italy	34
10	Cuvé di Rosé Veneto	2007	Veneto	Italy	Italy	106

Complexity	Category	Color	Aroma	Barrel	Viscosity	Price per	Rating
3	1	1	862,122,530	null	null	57	9
1	3	13	815	wood	null	780	9
2	3	13	122	wood	null	189	8
1	3	15	null	null	null	375	10
2	3	11	521,142	null	null	568	10
3	3	14	null	null	null	1000	10
2	1	2	410	metal	null	20	6
3	1	11	613	metal	null	68	7
1	3	12	420	metal	null	60	null
1	2	7	713	metal	null	null	null

Figure 4.2: The final database, filled with ten items

grapetype: id	name
-200	(negative value means white grape, positive value means red grape)
.....	
200	

Figure 4.3: Grape Type table in the database



## 5. Two Visualization Approaches

For most visualization applications, finding a best visual representation is the core problem. We have been doing the research to find a suitable the representation on the Internet and some visualization libraries for nearly two month. Eventually, two different representations approaches come out to be suitable. So, I implemented both of them that are going to be evaluated in the end of the thesis.

### 5.1 Circular Approach

As I sated in Section 2.1, the “aroma” attribute is the most important information in this thesis. We focus on visualizing “aroma”, which is in the scope of a sensory perception. The balloon layout tree is the representation of the aroma tree. Figure 5.1 is corresponding balloon layout tree with three nodes highlighted. It gives us a good hierarchical structure, and we may perceive the particular nodes quickly.

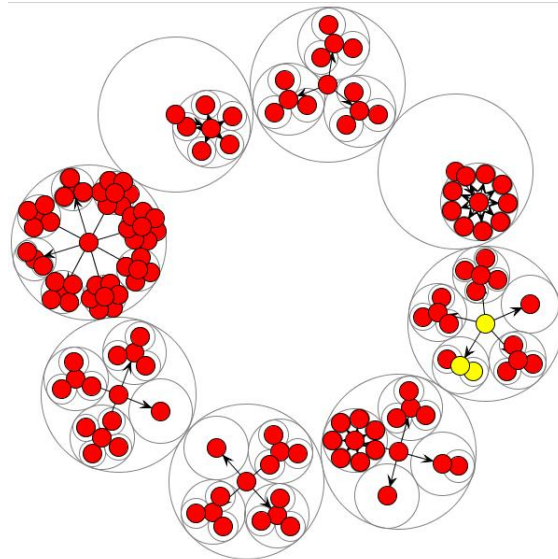


Figure 5.1: Balloon Tree Layout with the same nodes and structure of Figure 2.4

After choosing the basic structure, we need to embed the other attributes into it. There are 13 attributes left. One demand is using the space effectively. Since there is space left in the centre of this *balloon tree layout*, we decided to represent the other attributes in this area. Figure 5.2 is a visualization example from Many Eyes [15]. This idea is used to represent the other attributes, and put them in the centre of Figure 5.1. More specific, we use size and color of one circle to double code each attribute.

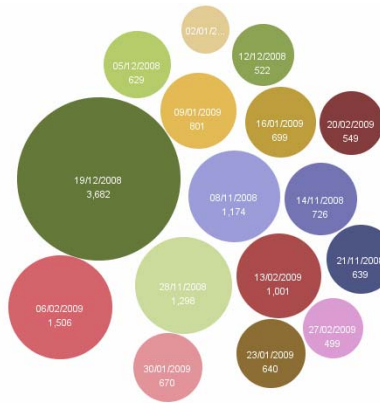


Figure 5.2: A visualization example, taken from [21]

Among the other thirteen attributes, “category” and “color” have similar property. We combine them, which will reduce the dimension. “Country” is a geographical attribute. We find two ways to represent it: one is representing it on a map image, the second one is laying a number of dots around the balloon tree layout and each dot represents one country. Here, we use the second one because it not just save space but also fit well with the balloon layout. There are about 30 countries that produce wine. Among them, 13 countries are the most part. Figure 5.3 shows this representation.

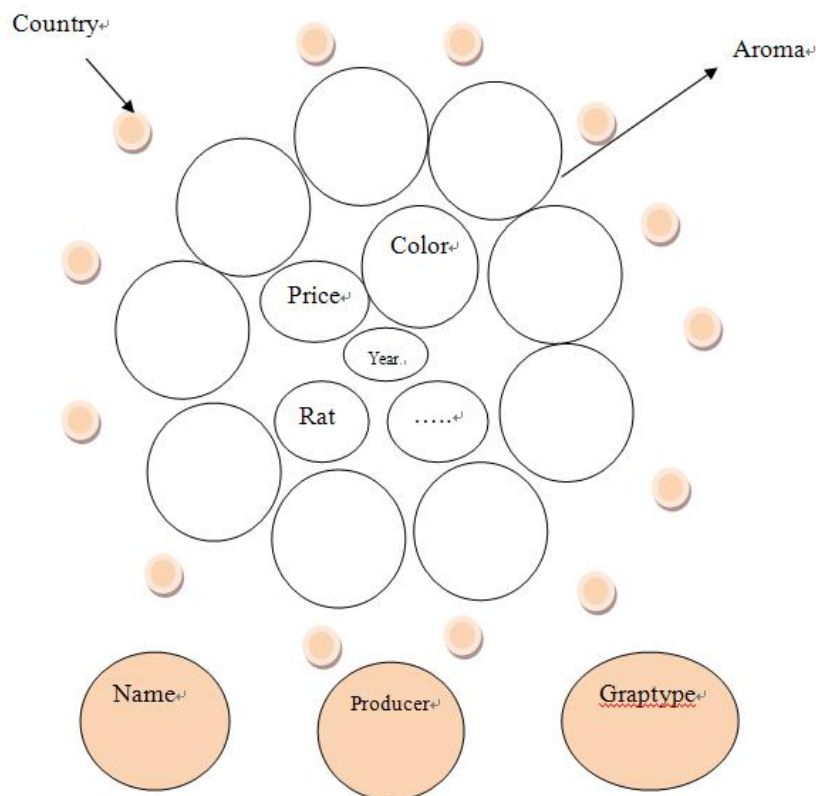


Figure 5.3: Final representation based on balloon layout

## 5.2 Treemap Approach

The other representation is based on tree-map layout, which represents the “aroma” hierarchical structure. We will use a world image to represent the “country” attribute. It fits tree-map because they are both rectangle shape. For saving the space, we decided to use part of Figure 5.4 according to the fact that wine grapes mostly grow between the 30th and the 50th degree of latitude, in both the Northern and Southern hemisphere [22].



Figure 5.4: A world image, generating by using visualization of Many Eyes [21]

Figure 5.5 shows this representation. We put the north part of a world map image in the up side, the south part of a world map image in the bottom side. The middle space will be used for the treemap and other attributes. In general, this representation has a rectangle shape, while the previous representation in Section 5.1 has circle shape.

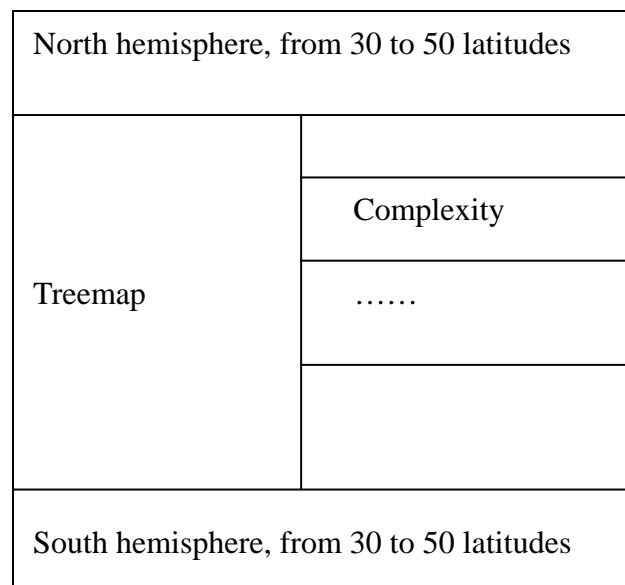


Figure 5.5: Final representation based on a treemap

### 5.3 Interface Scenarios

As I described in Section 1.3, I implemented a user interface, which could interact with the fingerprint generation tool. By using this interface, the users could input their options of wine. What's more, the interface and fingerprint generation tool should be dynamically connected, which means the fingerprint should change dynamically according to the operation on the interface.

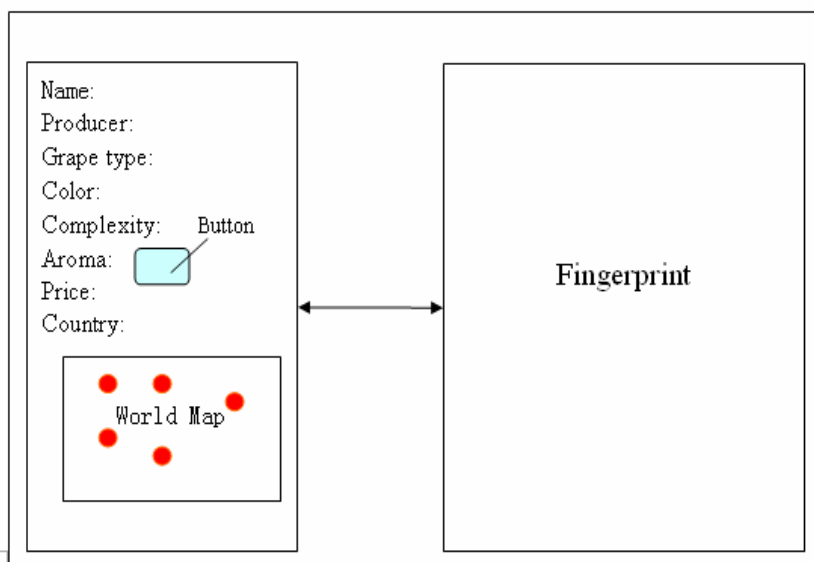


Figure 5.6: User Interface

Figure 5.6 is the proposed user interface. By using this interface, the user should be able to input their options about the wine name, producer, grape type, color, complexity, aroma, price, and country.

There are two challenges about the interface. One is how to input the option of aroma. Because aroma is a tree structure in my thesis, one possibility is using a list which contains all nodes of the tree. However, it is not easy for the user to select from the list, especially when the list has a lot of items. We would like to use the Radial Space-Filling approach in Figure 2.11. Then, the user can select a particular node by clicking on the circle segment in radial space-filling. Figure 5.7 illustrates the proposed visual aroma selection interface. There is a button for aroma, once the user click on it, it will pop-out a radial space-filling visualization whose circle segments is selectable to the users.

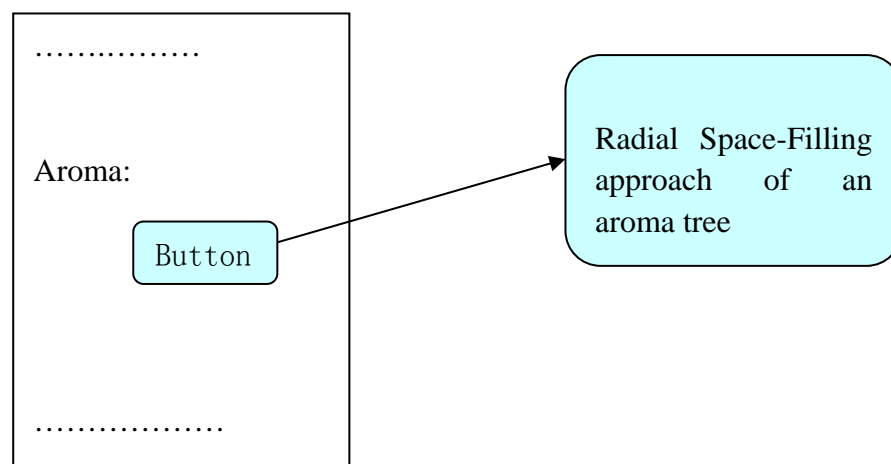


Figure 5.7: Aroma selection interface

The other challenge is how to input the country attribute. Of course, it is possible to use a list of the country, but it is not user-friendly, too. We would like to use the idea of Figure 2.8, which is using a world map image as the background, putting a dot on the location of each country. Then, if the user clicks on a dot, which is related to one country, the color of the dot will change and the country is the one which the user wants to select, see Figure 5.6.

## 6. Visualization and Interaction

After setting up the representation pattern in Chapter 5, we implemented the system. In this chapter, I will depict our wine visualization system including its interaction part. Since we have two representation approaches, we have implemented two approaches. I will introduce the approach based on the balloon tree layout in Section 6.1, the approach based on the treemap layout in Section 6.2. At last, I will show you how the user interface interacts with the fingerprint generation component.

### 6.1 Circular Approach: Generate Fingerprint Based on Balloon Tree Layout

By combing the two layouts, a balloon layout and a circle layout, it got the final fingerprint in Figure 6.1. This fingerprint visualizes 11 wine variables among 14. Because I did not find a way to represent the “region” attribute, and the “viscosity” attribute is empty in the database right now. Also, “color” and “category” are made into one because “color” is an extended variable of “category”.

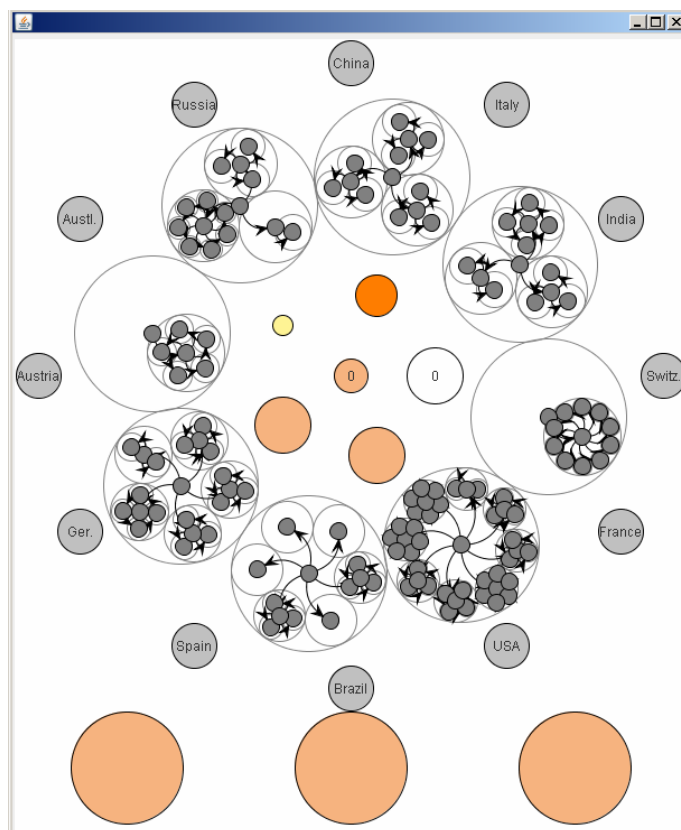


Figure 6.1: Fingerprint of our Circular Approach, based on a balloon tree layout

The circle in Figure 6.2 depicts the representation of the “country” attribute. Every country has a stable position in the circle, once one country was chosen, the color will change to red.

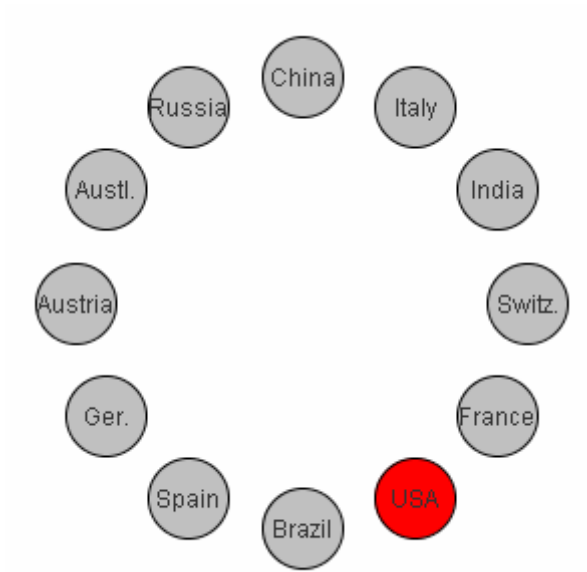


Figure 6.2: Circle layout represent “country” variable

The balloon layout in Figure 6.3 represents the “aroma” attribute. The outside 8 balloons are the first level of nodes in the aroma tree. The red color nodes mean the information of these nodes is the aroma information of a wine.

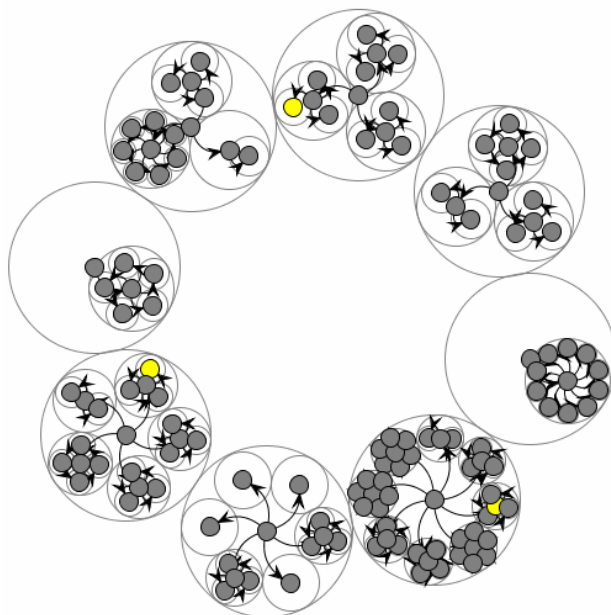


Figure 6.3: Balloon layout represent “aroma” variable

In Figure 6.4, there are 6 circles. Among them, “price” and “rating” are double coding by color and size, “color & category” are double coded by color and number.

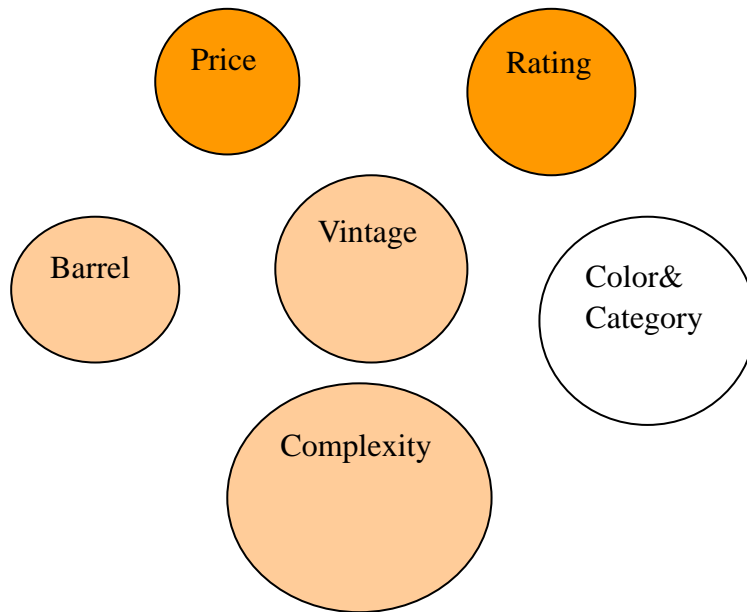


Figure 6.4: 6 circles represent 7 variables

The left three text attributes are represented as text in the fingerprint. For fixing them well in this circle approach, the text information is printed inside three circles, see Figure 6.5.

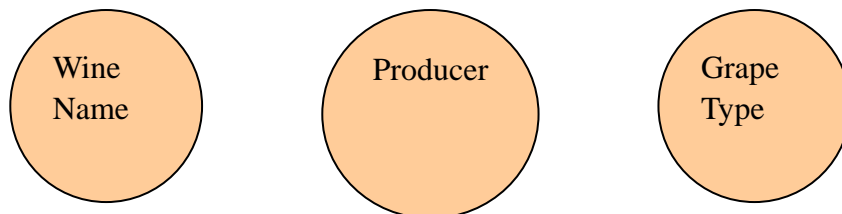


Figure 6.5: Represent “name”, ”producer”, ”grape type” as text information inside the circle



## 6.2 Treemap Approach: Generate Fingerprint Based on Treemap Layout

The Treemap Approach is the second representation that I introduced in Section 5.2. Figure 6.6 shows a fingerprint that “country” is France, “year” is 1969, price is 331, rating is 5, and so on.

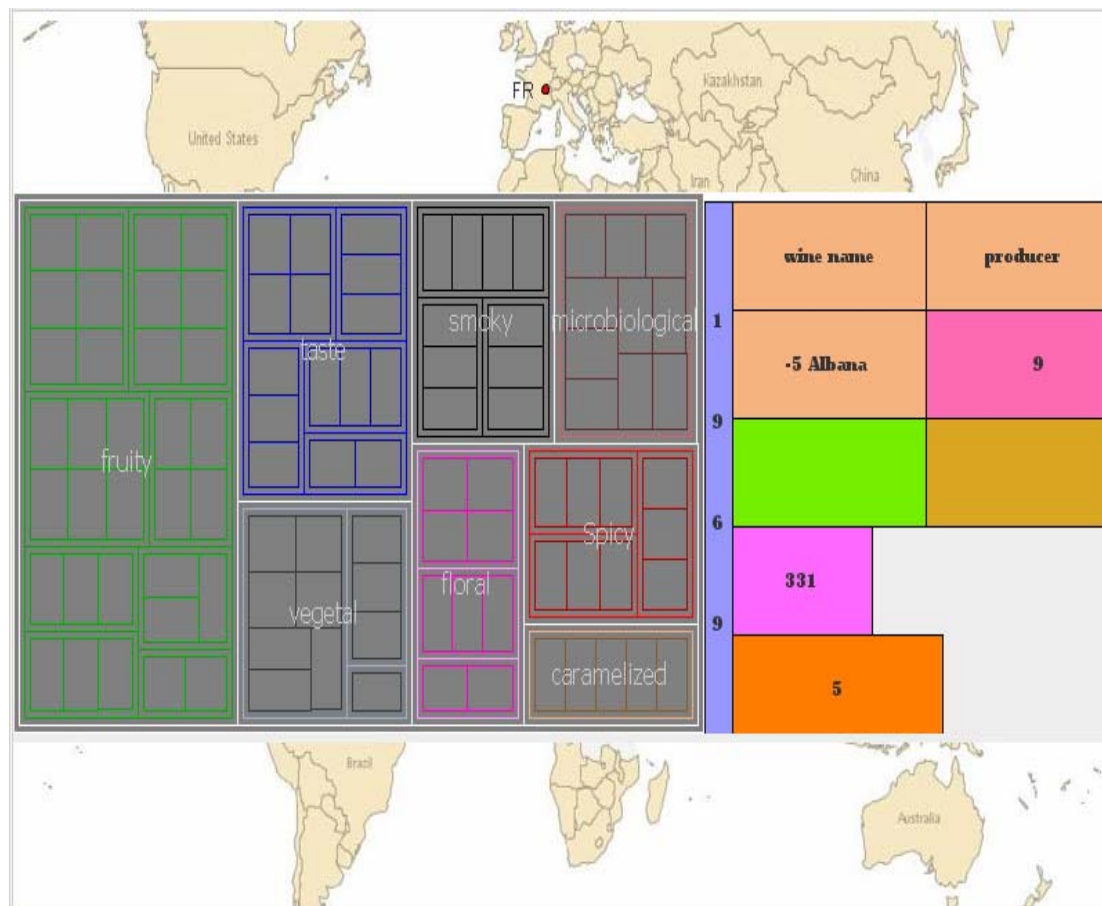


Figure 6.6: Fingerprint generated by treemap approach

The treemap visualizes the aroma tree. We defined different border colors and border thicknesses for the treemap to make the hierarchical structure clearer. The two images in the top and bottom of this fingerprint represent the “country” geographically. The red dot with a text label on the map locates the special country. The right side of treemap contains many rectangles that are used to visualize the left variables. Figure 6.7 depicts the other 10 variables. The blue rectangle on the left is “year”, which is 1980 in this example. The two rectangles on the top represent “name”, “producer” and “grapetype” respectively. Here, the name of the wine is *hermitage*, the producer of hermitage is *chave* and the grape type of hermitage is *albana*. The right rectangle of “grapetype” represents the “color & category” attributes. “category” contains white, rose and red. “color” variable list the specific color in white, rose and red. There are 16 kinds of color in the database where each color has a unique number. In this example, the color of the wine is *15 Brick Red*. Bellow the “grapetype” rectangle is “complexity” one. It is coded by color: light green, green, and dark green which represent light, medium and strong respectively.

The next rectangle is also coded by color: yellow and light gray which represent the wood barrel and mental barrel of “barrel” respectively. The rectangle below “complexity” and “barrel” represents the “price” variable. It is double coded by color and length. When the price is higher, the length of the rectangle is longer and the color of the rectangle is darker (from light pink to dark pink). The last rectangle in the bottom represents the “rating” variable which is similar with “price” in spite of the color range is from light yellow to red.

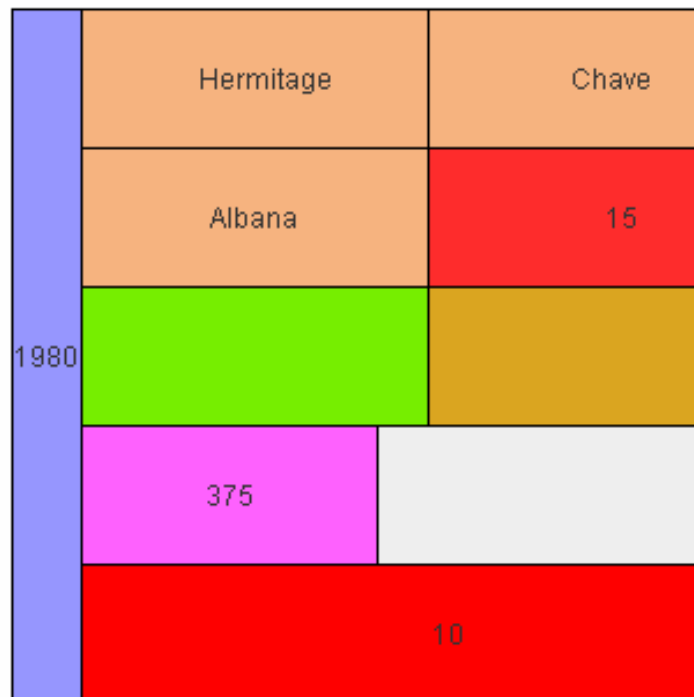


Figure 6.7: A number of rectangles representations

### 6.3 User Interface

After the fingerprints generating system was implemented, I needed a GUI (Graphical User Interface) that provides interactive functions between the users and the system. The users should use this GUI to input their preferences on wine. Figure 5.8 shows the interface after running the application.

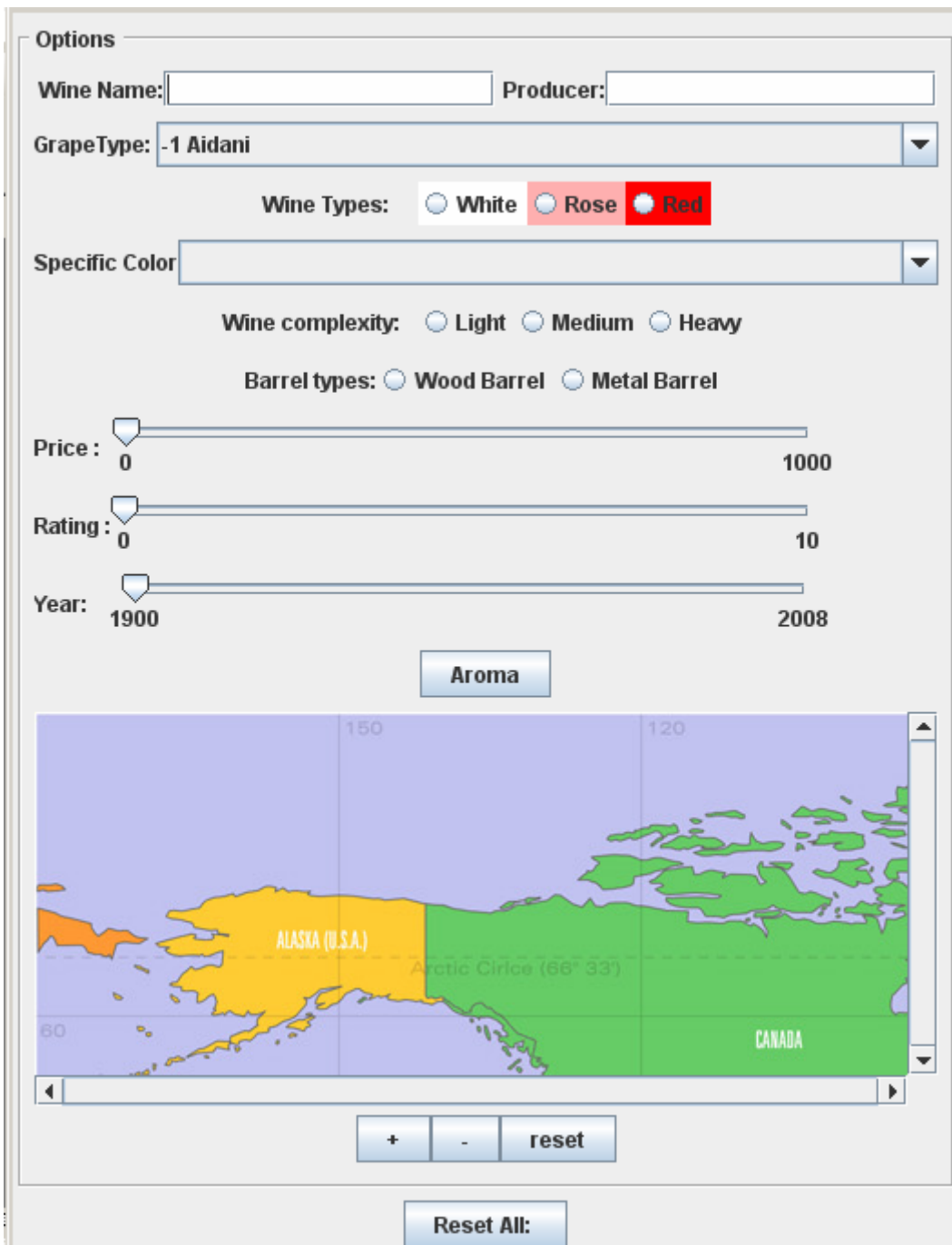


Figure 6.8: User Interface

The wine GUI mainly uses *Java Swing*, which is a widget toolkit for Java. It is part of Sun Microsystems' Java Foundation Classes (JFC) — an API for providing a GUI for Java programs. By changing the option of the GUI, the fingerprint in Figure 5.1 and 5.6 will change too. We see a world map image in Figure 5.8, where each country integrated with a blue dot. It is used to choosing the country for the users. For example, if I want to buy a wine from Spain, I simply right mouse click on the dot on Spain, then the blue dot will become red. Meanwhile, the corresponding “country” attribute on the fingerprint will change, too. On the other side, if I do not like the wine from Australia, I will left click on the blue dot on Australia in the map, and then the

color of the dot will change from blue to black. The corresponding fingerprint will change, too.

Figure 6.9 is the sub-interface when a user clicks the aroma button. It will pop-out a new frame with a *Radial Space-Filling* technique visualizing the aroma tree. Each circle segment has a unique color that also reflects as the same color in the fingerprint in Figure 6.6. The users can choose a specific node by clicking on the segment of circle. This interface also provides a multi-selection function. The users could select more than one aroma description by clicking “Add” button in the bottom. They could delete one selection by clicking “Delete selected”. They could cancel all selections by clicking “Reset” button.

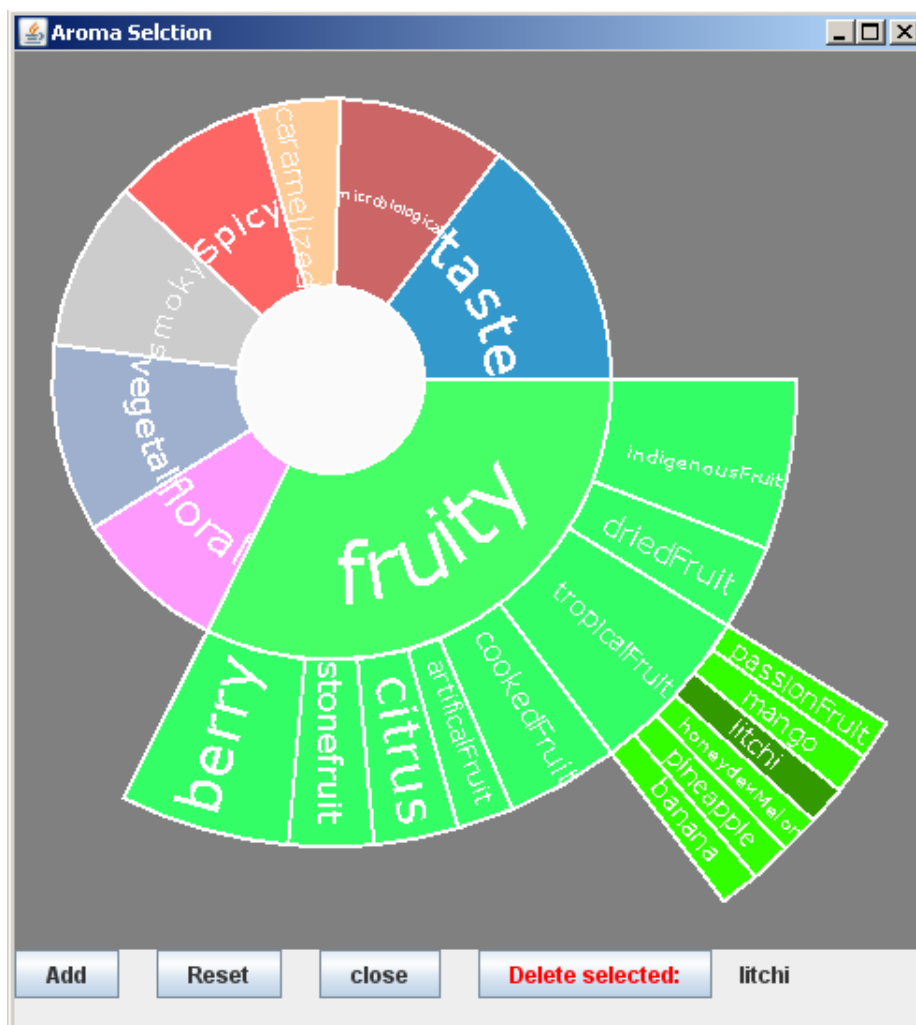


Figure 6.9: Pop-out frame for selecting aroma after clicking aroma button

#### 6.4 Database Fingerprints Generating

In this section, I will depict how we connected the database with our application, get the data and use it in the fingerprint generating system. It is also possible to store all fingerprints as JPG image files on the disk.

This tool follows three phases. Firstly, connect to the database and get the searching information. Secondly, import the data into the core fingerprint generating

code and generate the fingerprints of the database. Lastly, store the fingerprints as images in local disk. Figure 6.10 shows the codes used to connect the database “wine”.

```
try {  
    Class.forName("com.mysql.jdbc.Driver");  
    con = DriverManager.getConnection(  
        "jdbc:mysql://localhost/wine?use", "root", "123");
```

Figure 6.10: Codes used to connect the database

Figure 6.11 is the interface that the administrator uses to connect the database and generate the fingerprint. The administrator could change the store location in the disk, generating the fingerprints of all items in the database and generate the fingerprints of a set of items. The green color means this wine’s fingerprint has already been generated. The red color means this wine’s fingerprint has not been generated.

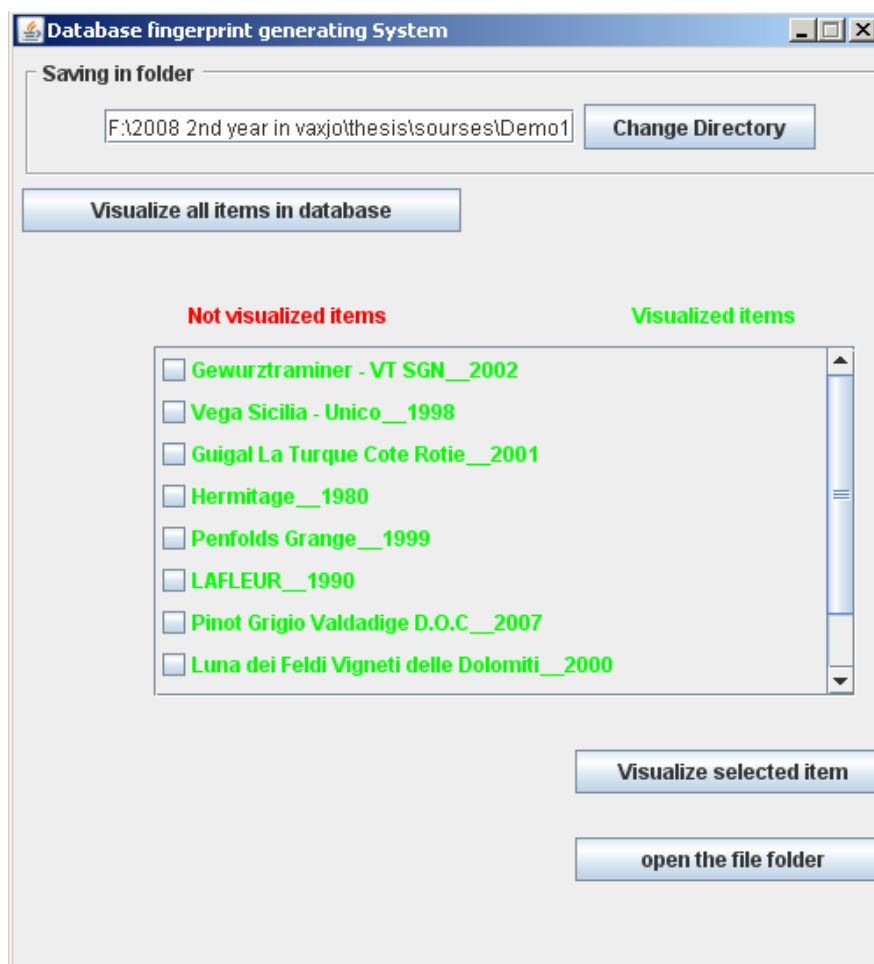


Figure 6.11: Database fingerprint generating system

## 7. Implementation

In this chapter, I would like to represent the programming language, external libraries, database system, integrated development environment (IDE), and operating system, and continuing with the description of the classes in my application.

### 7.1 Tools, Language, and Libraries

<b>Programming language</b>	<b>JAVA applet</b>
Description	The program that is developed on can be attached with a HTML page to make it a web-based program.
<b>Library1</b>	<b>JUNG</b>
Description	JUNG, the Java Universal Network/Graph Framework, is a software library that provides a common and extendible language for the modeling, analysis, and visualization of data that can be represented as a graph or network.
<b>Library2</b>	<b>Prefuse</b>
Description	Prefuse is a set of software tools for creating rich interactive data visualizations. The original prefuse toolkit provides a visualization framework for the Java programming language [23].
<b>Library3</b>	<b>JDBC</b>
Description	JDBC is the industry standard for database-independent connectivity between the Java programming language and a wide range of databases – SQL databases and other tabular data sources, such as spreadsheets or flat files. The JDBC API provides a call-level API for SQL-based database access [24].
<b>Database management system</b>	<b>MySQL Community Edition</b>
Description	MySQL Community Edition is a freely downloadable version of open source database that is supported by an active community of open source developers and enthusiasts.
<b>Operating system</b>	<b>Windows XP</b>
<b>IDE</b>	<b>Eclipse Platform Version: 3.4.0</b>
Description	Eclipse is a multi-language software development platform comprising an IDE and a plug-in system to extend it. It is written primarily in Java and is used to develop applications in this language.

Table 7.1: Tools, language, libraries

## 7.2 Class Description

In this section, I will present the detailed implementation by explain the core classes description. As I introduced in Section 5.1 and 5.2, I have two kinds of representation idea, so there are two applications at last. Figure 7.1 is the architecture of my applications. I will explain the classes of Approach 1, Approach 2, and database fingerprint generating system respectively.

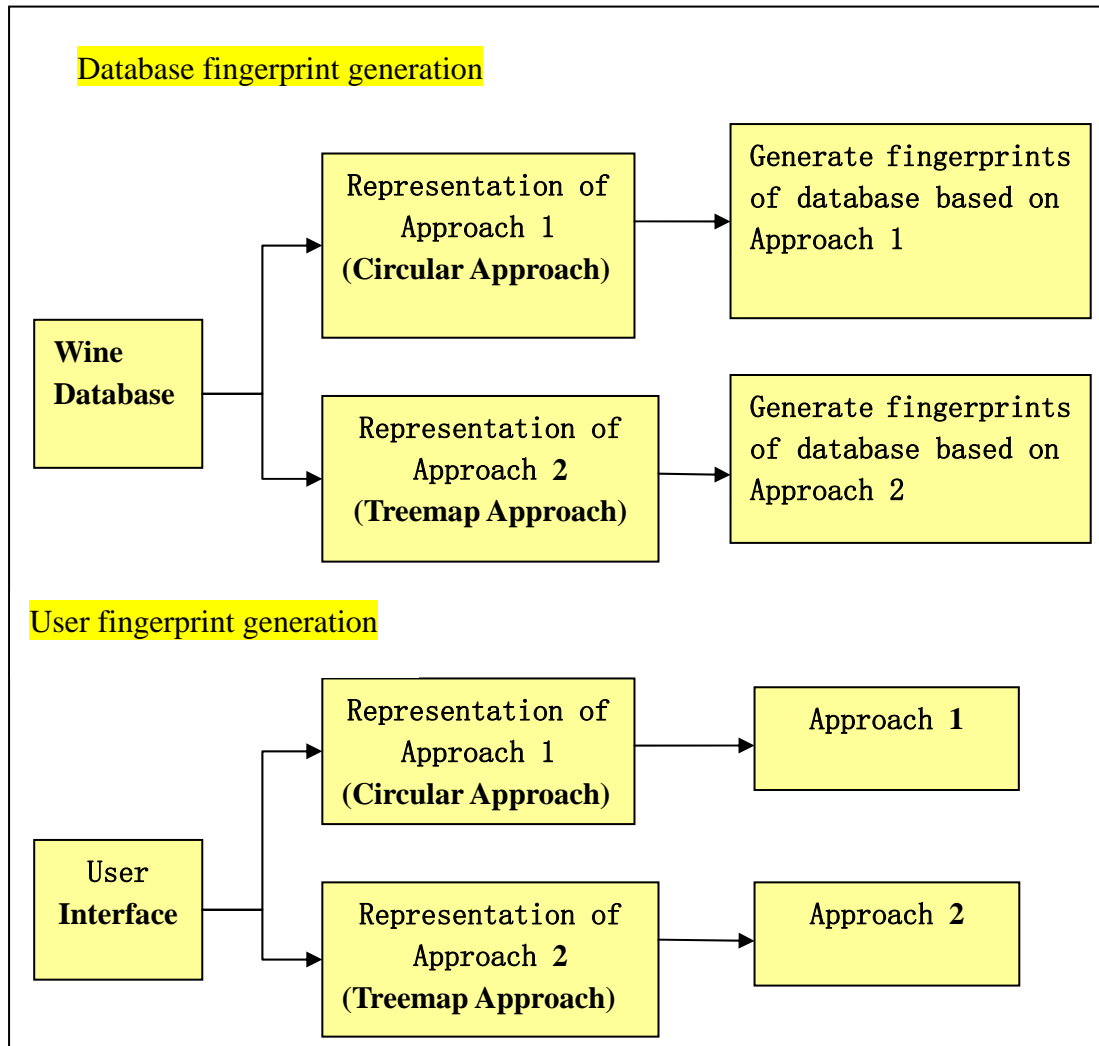


Figure 7.1: Architecture of wine visualization system

### 7.2.1 Class Description of Circular Approach

Class	BalloonLayoutDemo
Description	This class is used to visualize the wine information by combining balloonlayout, circlelayout, and some circles with text.
Inner classes	Rings: Generate the ellipse that constrains the sub-tree (see Figure 6.3). VertexShapeSizeAspect: Set the size of the vertexes in the whole fingerprint.

Methods	Paint (Graphics): Draw the ellipse of the balloonlyout. BalloonLayoutDemo(): Generate multiple-layouts by combing balloonlyout and circlelayout together.
Structure diagram	Figure 7.2

Table 7.2: Class-BalloonLayoutDemo description

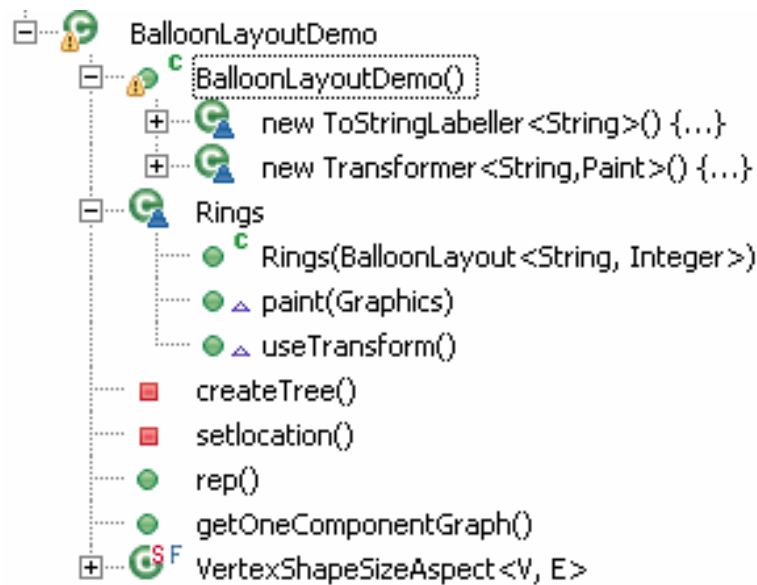


Figure 7.2: BalloonLayoutDemo-class structure

Class	Initial
Description	This class is used to store the initial information of all vertices including vertices of balloonlyout, vertices of circlelayout, and vertices of circle nodes.
Methods	Initial (): return the radial of eclipse of the nodes in Balloonlyout, which will be used in Paint () method in Table 7.1. CreateTree(): return the tree graph of all vertices, which is used in BalloonLayoutDemo() method in Table 7.1 GrapeList(): Store the list of the grape types into a string array, which is used to initial the grape list selection in the user interface.
Structure diagram	Figure 7.3

Table 7.3: Class-Initial description



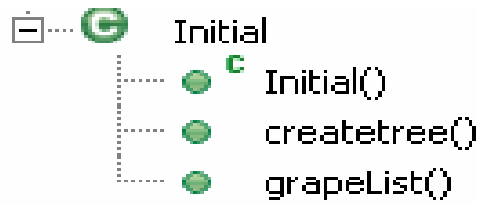


Figure 7.3: Initial-class structure

### 7.2.2 Classes Description of Treemap Approach

Class	TreeMap
Description	This class is based on the example from Prefuse [8]. I modified some parts. It is used to visualize the essential variable of wine, aroma. The input is a XML file of aroma information and the output is the tree map in Figure 6.6.
Methods	BorderColorAction(String) : set the color of the border of a node FillColorAction(String): set the filling color of a node LabelLayout(String): set the label of a node
Structure diagram	Figure 7.4

Table 7.4: Class-TreeMap description

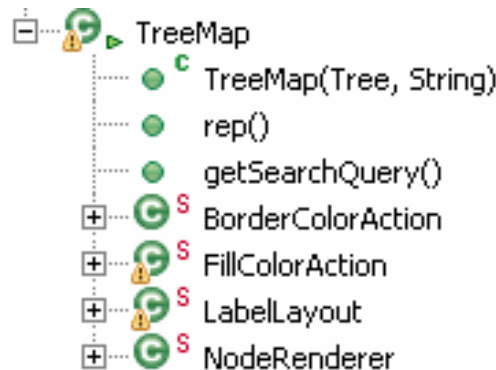


Figure 7.4: The structure of class-TreeMap

Class	SourthMap
Description	This class visualizes the country location as a dot on the south part of word map.
Methods	LatLonPixelTransformer (Dimension): Transform the Latitude and Longitude of a country into the pixel location on the screen. For example, the input location of France is {"46 40 N", "5 0 E"}.

	After transforming, the location is: (376.1111116409302,-145.55555608537452).
Structure diagram	Figure 7.5

Table 7.5: Class- SourthMap description

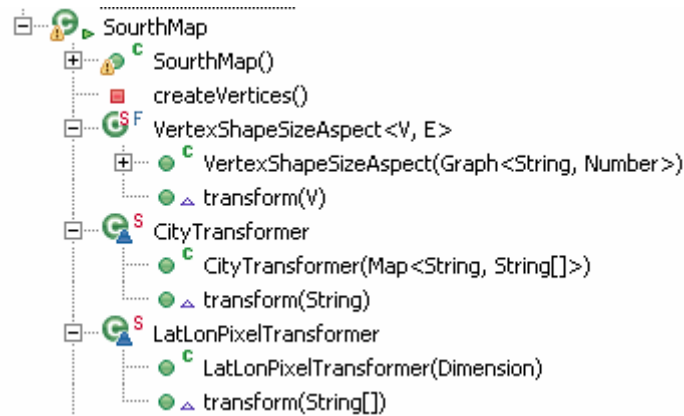


Figure 7.5: The structure of class-SourthMap

The class, NorthMap, is similar very with SourthMap. The only difference is the algorithm of location transformation method. So I will not list this class any more.

Class	Rectangle
Description	This class visualizes 10 variables, which are “name”, ”producer”, ”vintage”, “color &category”, “grapetype”, “barrel”, “complexity”, “price”, ”rating”.
Output	Figure 6.7

Table 7.6: Class- Rectangle description

### 7.2.3 Classes of the User Interface

Class	Basicframe
Description	This class implements the frame of the user interface. It mainly uses the Java swing toolkit [12] embedded with the StarburstDemo class and WorldMapDemo class.
Methods	addActionListener(new ActionListener() { public void actionPerformed(ActionEvent e) {} }): listening the action of a java swing component, and perform defined function to this action.

Table 7.7: Class- Basicframe description

Class	StarburstDemo
Description	This class is based on the example from Prefuse [8], I modified some parts. It visualizes the aroma tree of Figure 2.4 into radial-space filling. Then, the user can choose the nodes in the starburst and store or reset them (Figure 7.9).
Methods	NodeColorAction(String) : set the color of a node TextNodeAction(String): set the text information on the node
Structure diagram	Figure 7.6

Table 7.8: Class-StarburstDemo description

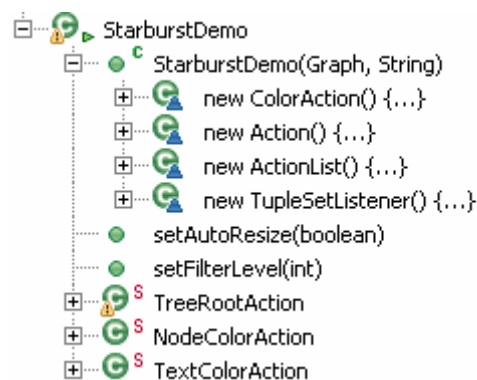


Figure 7.6: The structure of class- StarburstDemo

#### 7.2.4 Classes of Database Generating System

Class	Database
Description	This class connects the representation of approach 1 and approach 2 to the database. Then, it navigates the database and returns the re-formatted wine information to the approach 1 and approach 2 for generating the fingerprints.
Methods	Database(): Connect database. extracttype(): return the grape type information in the database. gettype(List<List<Integer>>): modifying the data from extracttype() to match the data required grape type format in demo1 and demo2. extractaroma(): return the aroma information in the database getaroma(List<List<Integer>>): modifying the data from extractaroma () to match the data required aroma format in demo1 and demo2.
Structure diagram	Figure 7.7

Table 7.9: Class- Database description

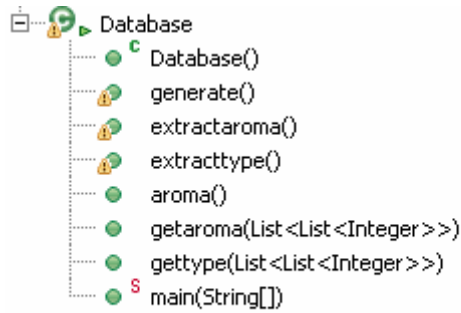


Figure 7.7: The structure of class- Database

Class	ScreenImage
Description	This class will store the generated fingerprints as images in the local disk.
Methods	<p>createImage(JComponent component, String fileName): Create a BufferedImage for Swing components. The entire component will be captured to an image.</p> <p>createImage(JComponent component, Rectangle region, String fileName): Create a BufferedImage for Swing components. All or part of the component can be captured to an image.</p> <p>BufferedImage createImage(JComponent component, Rectangle region, String fileName): Create a BufferedImage for Swing components.</p> <p>writeImage(BufferedImage image, String fileName): Write a BufferedImage to a File.</p>
Output	Figure 6.12
Structure diagram	Figure 7.8

Table 7.10: Class- ScreenImage description

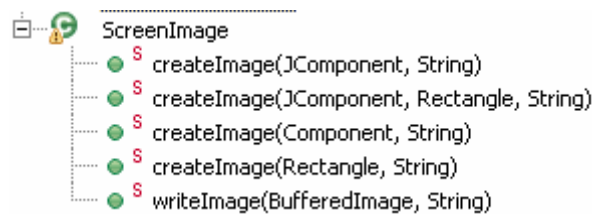


Figure 7.8: The structure of class- ScreenImage

## 8. Evaluation

After the implementation, we did an evaluation of this system. We made a JAVA applet, a program written in the Java programming language, which can be included in an HTML page, so that the interviewees do not need to install any software. We uploaded the system on a web server, which you can find in the Appendix B. Then, the users can use this tool on the Internet. The system would be improved according the users' evaluation in the future.

We have made three questionnaires, which you could find in Appendix C. Questionnaire A is about the background of the interviewee, questionnaire B is testing the user interface and the two approaches in general, and questionnaire C is about testing the fingerprints generated by the two approaches.

When all the preparative are done, we send an invitation email to a group of interviewees by email. Unfortunately, we only got ten replies. Actually, we can get more only because some interviewees can not open the system from the website. We noticed that they were using the wireless network of the campus. However, we could not fix this problem. Table 8.1 is the final analysis result from the eight answered questionnaires. NOAQ represents the Number of Answered Questionnaires. From B1 to C6, they represent the questions of the evaluation, which you could find in the Appendix C.

NOAQ	B1	B2	B3	B4	B5	B6	B7
8	3.625	3.75	3.125	3.375	3.25	2.333	7/3
C1	C2	C3	C4	C5	C6		
alcoholicity	3.125	75% correct	3.25	62.5% correct	5/3		

Table 8.1: The Analysis result

The results in the table are evaluated from 1 to 5 except B7, C1, C3, C5, and C6. 1 means very bad and 5 means very good. Among these scores, the lowest one is B6, which is a question about how the interviewee understands Approach 2 (treemap approach). Obviously, the test persons do not understand treemap approach very much. B7 is about asking the interviewees which approach they prefer to use, whose result, 7/3, means 7 persons chose Approach 1 (circular approach) and 3 persons chose approach 2. We count one vote for each approach when the interviewees chose both. The score of C3 is bigger than the score of C5; also means it is better that the users find the correct fingerprint by using Approach 1. Also, there are some text comments for the system. I combine them with the other result. Then, I get the further conclusions:

1. In general, the users prefer to use Approach 1. Because they understand it easier and also the circular shapes give some users the feeling of grapes.
2. There are some spelling mistakes in the aroma selection window.
3. The navigation of the price slider is confusing, since there is not unit of price

4. It will be helpful if there is a manual or user help button in the user interface to introduce the function of each item.
5. There are about 200 grape types, navigating them is a little bit hard.
6. More user find that fingerprint of the Approach 1 is easier to understand than the fingerprint of Approach 2.
7. One wine property could be added: alcoholicity
8. We need to embed a sample into the system to introduce the meaning of each item of the fingerprint.

## 9. Conclusions

In this Chapter, I will recall the goal of the thesis and give a conclusion regarding to the final system. Then, I will present the experience that I learned from the thesis. Also, some future work will be listed to make the function of the system easier to use.

### 9.1 Achievement

The goal of the thesis is: *Exploring and developing a suitable database schema of wine and a system which can visualize the wine items of the database and the user's preference wine. The system could help the user to find the right wine.*

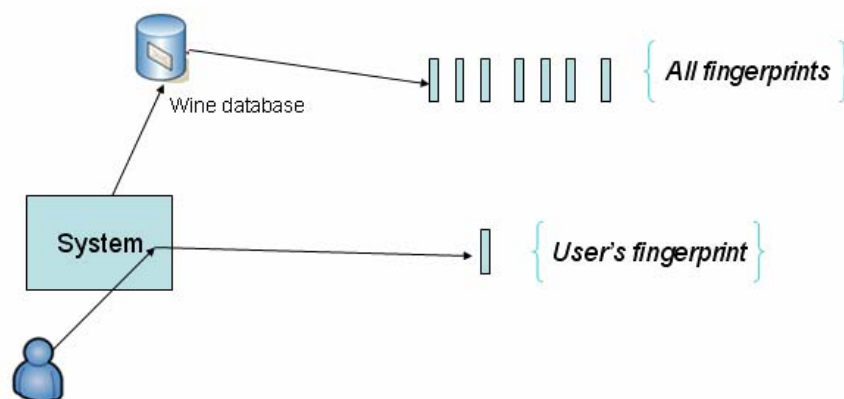


Figure 9.1: The system simulation

There are several phases during achieving this goal. First of all, define the wine attributes and made the database. For doing this, we needed to find the answer of the question. “How many properties of a wine have?”. The answer of this question is the prerequisite of our database. After doing some researches and analysis, 14 attributes are extracted out, see Table 4.1. These variables have different formats: “name”, “producer”, “vintage”, “grape type”, “country” is textual information. “Aroma” variable is a hierarchical tree structure (Figure 2.4). “price” and “rating” are integer variables. “color” is classified into 16 different kinds of color. “complexity” is classified into: strong, medium and light. Barrel has two options: wood or metal. “category” has three options: red, rose and white. After extracting the variables, we build the database based on them. Then, 10 different kinds of wine are filled into the database for testing.

Secondly, it comes to the following phase, which is finding a suitable visual representation for these attributes. For finding a good representation idea, we have looked through the visualization examples from the Internet and other visualization libraries. Finally, two representation ideas were set up, see Section 5.1 and 5.2. After the representation ideas were finalized, I analyzed and designed the requirements for constructing the blueprint for the implementation. It is presented in Chapter 7.

After the entire prerequisites are prepared, I could finally start to code the system. I use the Java programming language and some visualization libraries such as JUNG and Perfuse. I have spent two and half month to implement the system, see Figure 9.1.

The total system has about five thousands lines. Chapter 7 explains the implementation in detail. The program is basically UI design which is based on Java library swing and visualization technique supported by Prefuse and JUNG.

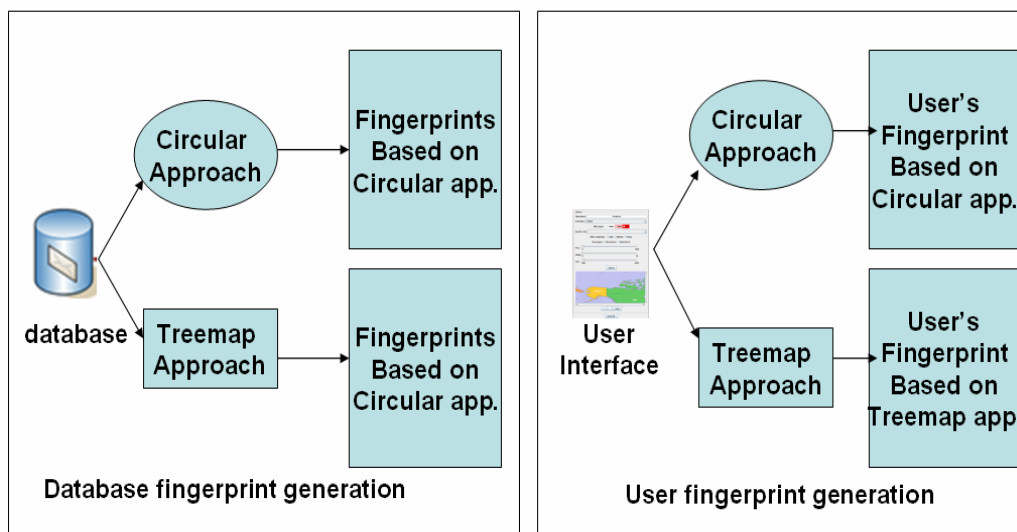


Figure 9.2: The Final System

So far, this visualization system has met the goal that I stated in the beginning of this section. First, the wine database is made. And, it has 14 wine attributes which provide the basic and detail wine information. The interesting aroma attribute is described by an aroma wheel that provides hierarchical description, see Figure 2.4. Second, the final system was implemented, see Figure 9.2. It contains two parts: database fingerprint generation and user fingerprint generation. The database fingerprint generation part will generate all fingerprints in the database. The user fingerprint generation part will generate a fingerprint according to the user's input in the User Interface. Lastly, we tested the system in Chapter 8. For testing if the user can use the system to find the right wine, we listed a target fingerprint and several other fingerprints. Among these fingerprints, there is one fingerprint which is very similar with the target one. We asked the testers to quickly read them and find the fingerprint which they think is similar with the target one. Finally, most of them find the right one. You could find this question from Questionnaire C in the appendix. By now, the system achieved the find goal which is helping the users find the right wine.

## 9.2 Experience

After done this thesis, I got a deeper understanding on information visualization. The core problem of visualization is finding a good visual representation of the given information so that the user can perceive the information easily and quickly. Since it is a practical and subjective topic, I have learned how to transfer a motivation into practice and have trained my programming skill, especially on developing a GUI tool and a visualization application.



### 9.3 Future work

Due to the limited time, there are some works left. And, the final tools could be improved too. Thus, there are some works that could be done in the future.

1. The world map choosing part on the GUI is not perfect right now. It is better that the user choose one country, this country could be highlighted. In addition, the user cannot choose a specific region where the wine is produced. In the future, we hope that these two points about the location choosing could be improved.
2. Right now, we only have ten kinds of wine in the database. It is not enough for a system. We need to fill more wine into it. Moreover, if we find some more interesting properties of wine, we can put it in the database and add them into the representation.
3. If we find a better representation of the wine attributes, which means this new representation could represent the wine information quicker and better, we can implement it and compare it with the two presentation ideas in this thesis.
4. Also, the bugs which I presented the Chapter 8 could be fixed.

## References

- [1] What font to use when and where, [www.wijnne.com/w/d4p.html](http://www.wijnne.com/w/d4p.html), Access data: 03-2009
- [2] "Literature Fingerprinting: A New Method for Visual Literary Analysis" by Daniel A. Keim, University of Konstanz and Daniela Oelke, University of Konstanz, ISBN: 978-1-4244-1659-2, Publication Date: Oct. 30 2007
- [3] A. Z. Broder. Some applications of Rabin's fingerprinting method. In *Sequences II: Methods in Communications, Security, and Computer Science*, pages 143--152. Springer-Verlag, 1993
- [4] Berkowitz, Mark (1996). "[World's Earliest Wine](#)" *Archaeology* (Archaeological Institute of America) , vol. 49 (1996), Retrieved 24 February 2004. .
- [5] Robinson, Jancis (2002). *Vines, Grapes and Wines*. Mitchell Beazley. ISBN 1-85732-999-6
- [6] [http://www.santamargherita.com/en/The\\_scale\\_of\\_wine\\_colors](http://www.santamargherita.com/en/The_scale_of_wine_colors), Access data: 04-2009
- [7] J. Robinson (ed) "*The Oxford Companion to Wine*" Third Edition pg 683 Oxford University Press 2006 [ISBN 0198609906](#)
- [8] Prefuse, A Java library, [www.prefuse.org](http://www.prefuse.org) , Access data: 04-2009
- [9] Wine Wheel of Red Wine and White Wine, German Wine Institute, <http://www.deutscheweine.de> Access Data: 04-2009
- [10] Wine Producing Country, FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS <http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#anchor>, Access data: 04-2009
- [11] List of Wine Producing Country, Wikipedia, [http://en.wikipedia.org/wiki/List\\_of\\_wine-producing\\_countries](http://en.wikipedia.org/wiki/List_of_wine-producing_countries), Access data: 04-2009
- [12] Chernoff, H., & Rizvi, M. H. (1975). Effect on classification error of random permutations of features in representing multivariate data by faces. *Journal of American Statistical Association*, 70, 548-554.
- [13] Andreas Kerren, Information Visualization I: Course Slide 2, 2009, Växjö University
- [14] Mapping Quality of Life with Chernoff Faces , Joseph G. Spinelli and Yu Zhou, <http://proceedings.esri.com/library/userconf/educ04/papers/pap5000.pdf> Access data: 04-2009
- [15] Many Eyes, a website visualizing data by IBM, <http://manyeyes.alphaworks.ibm.com/manyeyes/> Access data: 04-2009
- [16] JUNG, A Java library, <http://jung.sourceforge.net/>, Access data: 04-2009
- [17] Balloon layout, JUNG website, <http://jung.sourceforge.net/applet/balloonlayout.html>, Access data: 04-2009
- [18] Andreas Kerren, Information Visualization I: Course Slide 4, 2009, Växjö University
- [19] Beamtrees: Compact Visualization of Large Hierarchies Ham, Frank van and Wijk,

- Jarke J. van (2002): In: InfoVis 2002 - 2002 IEEE Symposium on Information Visualization 27 October - 1 November
- [20] List of grape varieties, Wikipedia,  
[http://en.wikipedia.org/wiki/List\\_of\\_grape\\_varieties](http://en.wikipedia.org/wiki/List_of_grape_varieties), Access data: 05-2009
- [21] Visualization Example, Many Eyes,  
<http://manyeyes.alphaworks.ibm.com/manyeyes/>, Access data: 04-2009
- [22] List of wine-producing regions, Wikipedia  
[http://en.wikipedia.org/wiki/Wine\\_region](http://en.wikipedia.org/wiki/Wine_region), Access data: 05-2009
- [23] Prefuse: information visualization toolkit,  
<http://prefuse.org/>, Access data: 04-2009
- [24] Java SE Technologies- database, Sun,  
<http://java.sun.com/javase/technologies/database/>, Access data: 03-200
- [25] Wiegers, Karl E. (2003). Software Requirements 2: Practical techniques for gathering and managing requirements throughout the product development cycle (2nd ed.). Redmond: Microsoft Press. ISBN 978-0735618794.
- [26] Spence, R., 2001. Information Visualization. Barcelona: Graphos S. A., Arte sobre papel.
- [27] Alexander I., Maiden N. Scenarios, Stories, Use Cases. Wiley 2004. ISBN: 0470861940.
- [28] MySQL database management system, [www.mysql.org](http://www.mysql.org) , Access data: 03-2009
- [29] Spence, R., 2001. *Information Visualization*. Barcelona: Graphos S. A., Arte sobre papel.

# Appendix

## A. Invitation Email

### Invitation letter

Hello,

The ISOVIS group at MSI has developed a prototype implementation of a Visualization system for wine attributes. By using our system, a potential customer could get a better understanding of specific wines and might find out his/her favorite wine easier.

To get a first feedback from users, we would like to ask you to attend a short usability study. You don't need to install any software as the system is web-based. We expect that you don't need more than 15-20 minutes for testing our approaches and filling out three questionnaires.

With your help, we can improve our system, and we appreciate your willingness to attend this brief evaluation. **Please, regard the following list of instructions:**

- 1. Read the attached information sheet carefully: "info.pdf"**
- 2. Follow the links in that sheet to test our two visualization approaches**
- 3. Fill out the three (!) questionnaires in "questionnaires.doc"**
- 4. Reply this email to Yuanxun Mei and Ilir Jusufi together with the completed questionnaires by next Monday 24/8-2009 at the latest.**

Thank you very much!

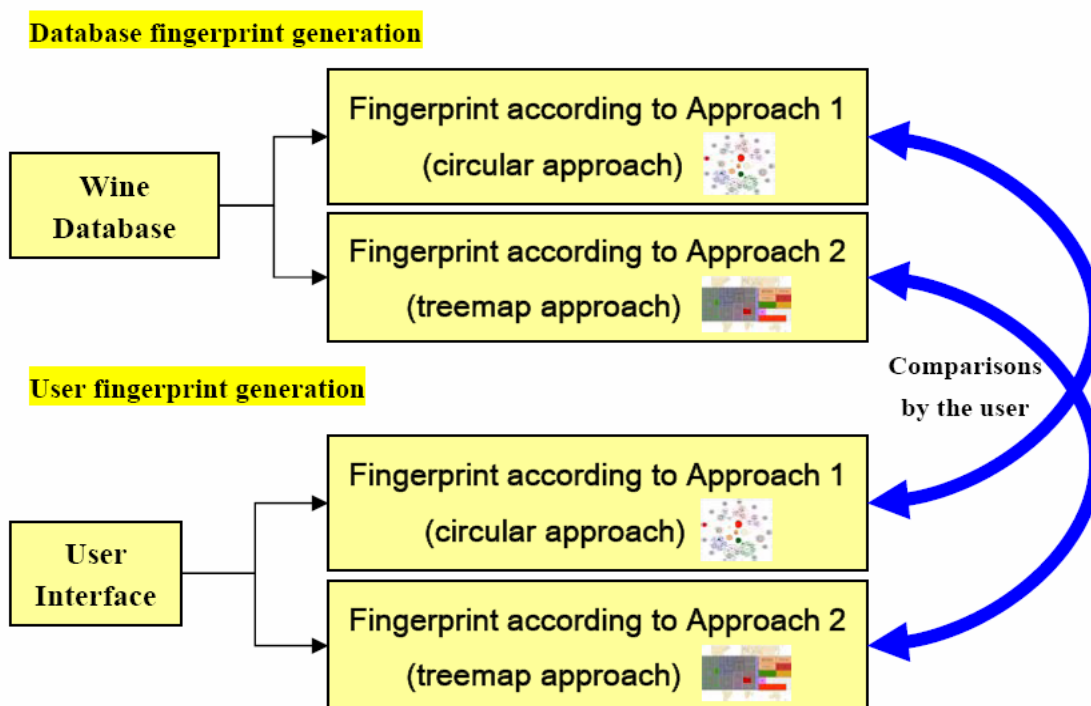
Best wishes,

Yuanxun Mei, Ilir Jusufi, and Andreas Kerren

## B. Information Sheet

### Evaluation of a Wine Attribute Visualization System

Our wine attribute visualization system represents the main wine attributes in form of a picture that we call fingerprint. You can use this system to generate your personal wine fingerprint and compare it with the fingerprints in a wine database. Then, you might be able to find the right wine for your own individual taste. For example, if you are looking for a red wine that is dry, has a light taste, is ruby red, and comes from any country but Australia, you can enter these properties or conditions into the system. Then, it will generate a fingerprint according to these four features. In a second step, you can compare your fingerprint with wine fingerprints in a wine database to find some fingerprints with same or similar parts. These fingerprints of specific wines are probably the wines that you are looking for.



As we currently have two representation approaches, our system consists of two versions. The diagram above shows the simplified architecture of the system. These two approaches share the “graphical user interface” (GUI) and the “wine database”. The difference is that they follow distinct representation ideas, which means that the fingerprints have a different design. The links of our two approaches are:

<http://cs.msi.vxu.se/isovis/wine/Applet1/> (so-called circular approach)

and

<http://cs.msi.vxu.se/isovis/wine/Applet2/> (so-called treemap approach).

**Please, try to generate a wine fingerprint of your currently preferred wine using both Java applets with identical GUIs.** The user interface is self-explaining: by changing the controls you can specify wine attributes that are visually represented in

the fingerprint. Note that the button “Aroma” opens a new window with a more complex control to select one or more hierarchically organized types of aromas.

## C. Questionnaires

### A. Demographics Questionnaire

*Instructions:* for each question, please write the answer or choose the option that mostly reflects your impression/opinion. If there are some questions you do not know the answers or do not want to answer, you are free to leave them.

**Name (voluntary):**

**Gender:**    Male     Female

**Age:**    18-20     21-29     30-39     40-49     50-59     60+

**Education (highest level completed):**

High School     Bachelor     Master     Ph.D

**Job title / Field of study:**

**Your experience of drinking wine:**

Beginner     Average customer     Expert

## B. Questionnaire for the usability of the user interface and two approaches

*Instructions:* for each question, please choose the number of the answer that is mostly reflected to your impression of this question. If there are some questions you do not know the answers or do not want to answer, you are free to leave them.

### 1. What do you think the usability of the user interface in general?

Very hard    1     2     3     4     5     Excellent

### 2. There is an “aroma” selection button in the user interface. If you click on it, it will pop out a circle constructed by some circle segments. You can choose your preference wine aroma with it. After using it, how well do you understand the usage of this “Aroma” selection function?

Very bad    1     2     3     4     5     Very good

### 3. After using Approach 1, how well does the generated fingerprint reflect the chosen options in the user interface?

Very confusing    1     2     3     4     5     Very clear

### 4. After using Approach 1, were you able to get familiar with all the options in the user interface and understand their relation with visualization in the fingerprint?

Very confusing    1     2     3     4     5     Very clear

### 5. After using Approach 2, how well does the generated fingerprint reflect the chosen options in the user interface?

Very confusing    1     2     3     4     5     Very clear



**6. After using Approach 2, were you able to get familiar with all the options in the user interface and understand their relation with visualization in the fingerprint?**

Very confusing 1  2  3  4  5  Very clear

**7. Which approach do you prefer to use?**

Approach 1  Approach 2  Both  Neither

**8. Comments regarding the user interface:**

### C. Questionnaire for the fingerprints

*Instructions:* for each question, please write the answer below the question or choose the number of the answer that is mostly reflected to your impression to this question. If there are some questions you do not know the answers or do not want to answer, you are free to leave them.

**1. Based on the two approaches, are there any other wine attributes that are important to you, however, they are not showed in this system?**

No  Yes  (if yes which ones?)

**2. How much do you understand the fingerprints of Approach 1?**

Nothing  1  2  3  4  5  Completely

**3. By watching the “testing fingerprint” and the “No.1”, “No.2”, “No.3”, “No.4” on the link <http://cs.msi.vxu.se/isovis/wine/evalapp1/>, which of the fingerprints corresponds to the “testing fingerprint”? Choose the one you think matches the best?**

No.1  No.2  No.3  No.4  None

**4. How much do you understand the fingerprints of Approach 2?**

Nothing  1  2  3  4  5  Completely

**5. By watching the “testing fingerprint” and the “No.1”, “No.2”, “No.3”, “No.4” on the link <http://cs.msi.vxu.se/isovis/wine/evalapp2/>, which of the fingerprints corresponds to the “testing fingerprint”? Choose the one you think matches the best?**

No.1  No.2  No.3  No.4  None

**6. Which fingerprint is easier to understand?**

Fingerprint of Approach 1  Fingerprint of Approach 2  Both  Neither

**7. Comments regarding the fingerprints:**



**Matematiska och systemtekniska institutionen**

SE-351 95 Växjö

Tel. +46 (0)470 70 80 00, fax +46 (0)470 840 04

<http://www.vxu.se/msi/>