

# A FUZZY MODEL FOR THE PITCH PERCEPTION OF COMPLEX TONES

D. K. Fragoulis, C. N. Papaodysseus and J. N. Avaritsiotis

National Technical University of Athens  
Department of Electrical and Computer Engineering  
9 Heron Polytechniou St., 157 73 Zographou, Athens, Greece  
E-mail: frag@mlab.ece.ntua.gr

## ABSTRACT

In this work, we consider the problem of pitch perception and we propose a fuzzy model that can account, in a qualitative way, for all the possible cases of complex tones. The model includes two distinct stages. The first stage is a frequency analyzer that determines the frequencies of some of the individual sinusoidal components of the complex tone. This information proceeds to the second stage of the model, which is a fuzzy processor. The processor receives the frequencies of the components and performs all the necessary computations in order to derive the pitch. The use of a fuzzy processor that represents and manipulates information, which is not precise, makes the model capable to predict the ambiguities of pitch that have been observed.

## 1. INTRODUCTION

Pitch may be defined as that attribute of auditory sensation in terms of which sounds may be ordered in a musical scale. Like loudness and timbre, it is a subjective attribute that cannot be expressed in physical units. Essentially, pitch is related to the repetition rate of the waveform of a sound. In the case of a pure tone, it is primarily correlated with tone's frequency, despite the fact that intensity, duration and temporal envelope of the tone may have an influence on pitch. Assigning a pitch value to a complex tone is generally understood to mean determination of the frequency of a pure tone having the same subjective pitch as the complex tone. It is noticed that even sounds not formed of well-defined discrete sinusoids may evoke a pitch sensation.

In music, pitch is related with the features of melody and harmony. Since a simple tone evokes a pitch, then a sequence of tones with appropriate frequencies should evoke the percept of melody. However, it seems that a sequence of tones evokes a sense of melody only when tones lie below 45 kHz [1]. Also, experimental results

on pitch identification have shown that the pitch of a complex tone can be ambiguous, especially if low order components are weak or missing, or if only a few components are present.

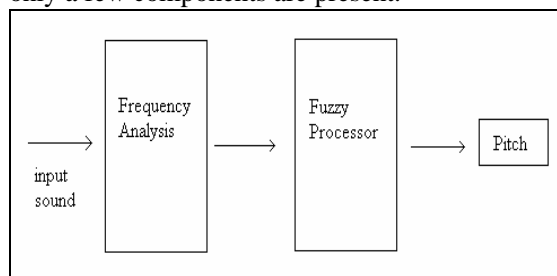


Figure 1. A schematic model for a pitch perception system

The pitch of a complex tone could be derived from pitches of the individual components using a pattern recognition model [2]. Generally, such a model involves two stages, as illustrated in figure 1. The first stage is a frequency analyzer, which simulates the procedure of spectral analysis performed in the inner ear, so that different frequencies are separated. The second stage is a pattern recognizer, which determines the pitch of the complex tone from the frequencies of the resolved components. This is obtained by searching for a fundamental frequency whose "harmonics" match the frequencies of the resolved components of the stimulus as closely as possible.

The pattern recognizer should be able to work properly in all possible types of complex tones. Consider as an example a sound consisting of the frequencies 200Hz, 400Hz, 600Hz, 800Hz, etc. This sound has a low pitch, which is the same to the pitch of a 200Hz pure tone. However, it is possible to filter the sound in order to remove the 200Hz component and we will find that the pitch does not alter. The same result will be observed if we eliminate all except a small group of mid frequency harmonics, say 1800Hz, 2000Hz, 2200Hz. This phenomenon is well known as the "missing fundamental". As an other example we could consider a sound that consists

of three sinusoidal components at 1840Hz, 2040Hz and 2240Hz. The perceived pitch of this sound is about that of a 204Hz pure tone. Actually, there is an ambiguity of the exact pitch value since it seems to lie around 180Hz, 204Hz and 227Hz.

In order to have an effective pitch identification system the pattern recognizer should be able to manage with pitch ambiguities. To achieve this, a fuzzy processor is used which performs all the manipulations using fuzzy numbers. This feature of the processor provides the ability to adjust the calculation precision. Thus, if we represent the frequency values as fuzzy numbers, the system is able to derive vague values of pitch.

## 2. FUZZY NUMBERS

Fuzzy numbers are a special form of fuzzy sets defined in the space of real numbers [3], [4]. They possess some additional properties relative with the shape of their membership functions.

*Definition:* A fuzzy number  $F$  is a fuzzy set defined in  $\mathbf{R}$  such that:

- (i)  $F$  is a normal fuzzy set, i.e. there exists at least one element  $x$  of  $\mathbf{R}$  which  $F(x)=1$ .
- (ii)  $F$  is convex.
- (iii)  $F$  is upper semicontinuous  $\Leftrightarrow$  All  $\alpha$ -cuts of  $F$  are convex.
- (iv)  $F$  has a bounded support  $\Leftrightarrow$  All  $\alpha$ -cuts are closed intervals of  $\mathbf{R}$ .

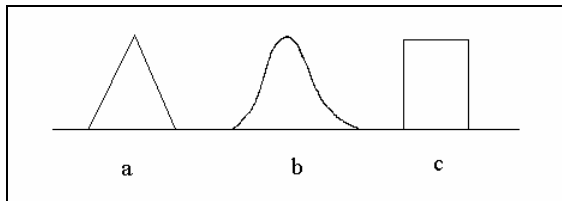


Figure 2. Examples of fuzzy numbers

A fuzzy number may be considered as a suitable model of approximate notions, as for example, near zero, about 5, etc. Some shapes of fuzzy numbers are illustrated in figure 2. The membership function of the fuzzy number visualizes a grade membership of a given element of a concept (near, about, etc.). Unimodality of  $F$  assures us that exactly one region exists, in which the relevant elements have the highest grades of membership. Of course, the shape of a membership function reflects different situations in which a variety of shapes can be observed. In figure 2a, a triangle-like fuzzy number is presented, where three characteristic points are interpreted as upper and lower values of the

range, and the element of  $x$  corresponding to the highest grade of membership is called a modal value of the fuzzy number. In figure 2b, a so-called bell shaped fuzzy number is given, while in figure 2c an interval that is a particular case of fuzzy number, is presented.

We observe that the shapes of the membership functions are similar in that one modal and two extreme limits are evident. This enables us to use a parametric representation of the fuzzy numbers. For this purpose, L(left side) and R(right side) fuzzy numbers are introduced. By an L fuzzy number we mean a fuzzy number of this membership function:

- (i)  $L(-x)=L(x)$
- (ii)  $L(0)=1$
- (iii)  $L$  is increasing in  $[0,+\infty)$

The same definition holds for the R fuzzy number. An L-R fuzzy number possesses the following membership function:

$$F(x) = \begin{cases} L\left(\frac{m-x}{a}\right), & \text{if } x \leq m \\ R\left(\frac{x-m}{b}\right), & \text{if } x \geq m \end{cases}$$

where  $a, b > 0$  are parameters controlling fuzziness of the fuzzy number. For the case  $a=b=0$  we get a genuine real number. Including the above parameters  $(m,a,b)$  the fuzzy number is denoted by  $A(m,a,b)$ .

## 3. THE FUZZY PROCESSOR

The fuzzy processor is composed of two blocks, as illustrated in figure 3. The first block takes as input the frequencies of the individual components of the sound. We notice that the phases and amplitudes of the components are ignored. Next, every component is resolved in order all the potential subcomponents to be derived. Essentially, computing all the submultiples of the frequency of all the components performs the resolution procedure. Subsequently, the fuzzyfication of all the subcomponents is performed, by assigning a fuzzy number at each subcomponent. The shape of the membership function of the fuzzy numbers is given below:

$$F(x) = e^{-b(x-x_0)^2}$$

The value of  $b$  determines how narrow the membership function will be. We notice that the width of the membership function is related to the description explicitness of the subcomponent frequency. The frequency discrimination ability

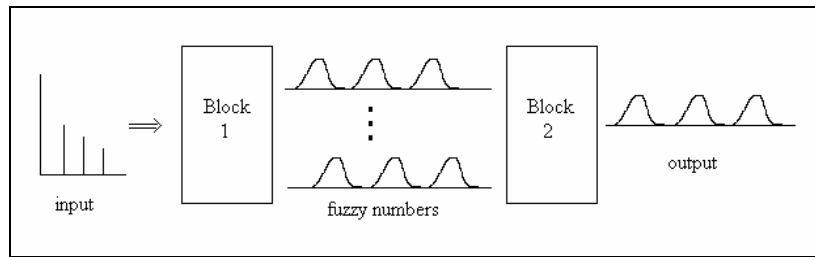


Figure 3. The fuzzy processor

of the human ear appears a logarithmic dependence on frequency. Thus we should select an appropriate value for  $b$ , according to the harmonic interval where each fuzzy number belongs.

After assigning a membership function at each subcomponent, the union operation is applied on the groups of subcomponents that have been generated by the same component of the sound. Therefore, some non-unimodal fuzzy numbers derived, which correspond to the components of the input sound. Finally, the intersection operation is applied on the aforementioned fuzzy numbers. The result of this operation is a fuzzy number whose membership function appear some local peaks, and the greatest of them determines the pitch. If there exist more than one peaks having an almost common value, then all of them contribute to the formation of the pitch.

The union and the intersection operations are defined as follows:

$$A_1 \cap A_2 = \min\{A_1, A_2\}$$

$$A_1 \cup A_2 = \max\{A_1, A_2\}$$

#### 4. SIMULATION RESULTS

The developed fuzzy model has been used to derive the pitch in three special cases of complex tones, as discussed in following. As a first example, we considered a harmonic sound consisted of three components at 200Hz, 300Hz and 400Hz. Figure 4 shows the resolved subcomponents, while in figure 5 the membership

Frequency of component			Value of integer divisor
200Hz	300Hz	400Hz	
100	150	200	2
rejected	100	133.3	3
rejected	rejected	100	4
rejected	rejected	rejected	5

Figure 4. The derived subcomponents for the first example

functions of the corresponding fuzzy numbers are presented. The output of the model is shown in figure 6.

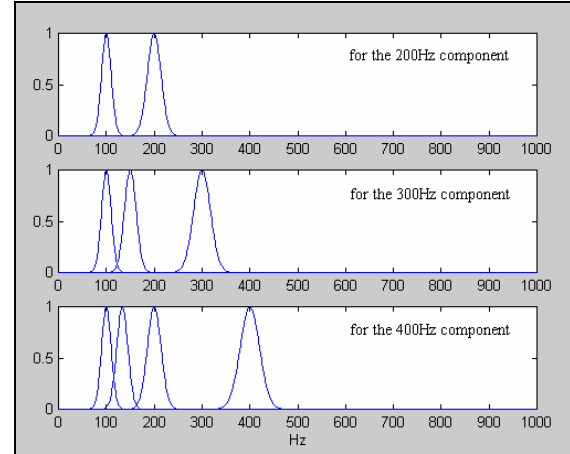


Figure 5. The derived fuzzy numbers for the first example

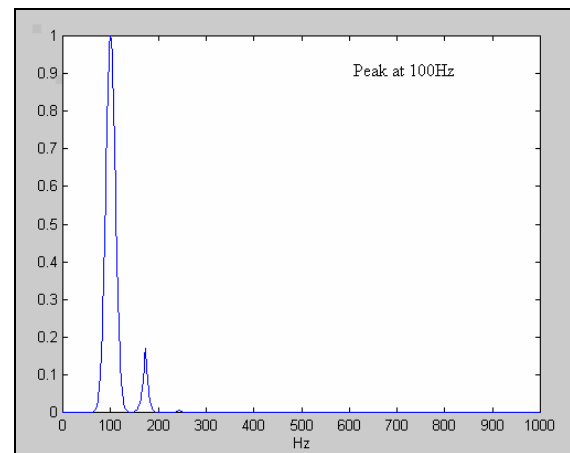


Figure 6. Output of the model for the first example

In a second example, a non-harmonic sound with three components at 240Hz, 340Hz and 440Hz has been used. The results are presented in figures 7, 8 and 9. It is apparent that the proposed model derives a pitch despite the fact that the subcomponents do not coincide (see fig. 7). As a third example a non-harmonic sound with components at 1840Hz, 2040Hz and 2240Hz has been treated. The values of sound components

chosen, have been used by researchers [5] to show experimentally that pitch cannot be determined univocally in some cases. The output of the model is illustrated in figure 10. It is apparent that the output appears in the form of three dominant peaks. We notice that, in all the examples, the subcomponents which do not lie in the interval [100Hz, 5000Hz] have been rejected.

Frequency of component			Value of integer divisor
240Hz	340Hz	440Hz	
120	170	220	2
rejected	113.3	146.6	3
rejected	rejected	110	4
rejected	rejected	rejected	5

Figure 7. The derived subcomponents for the second example

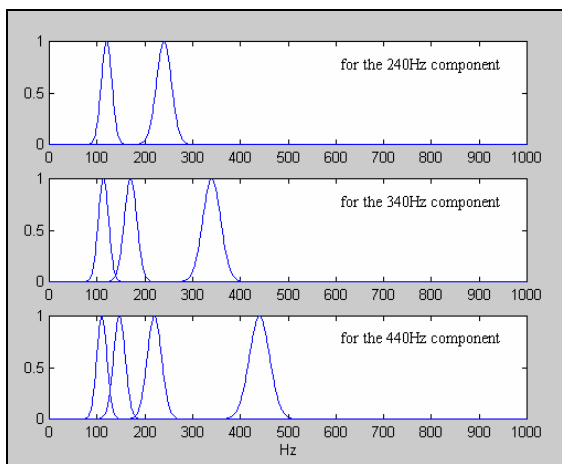


Figure 8. The derived fuzzy numbers for the second example

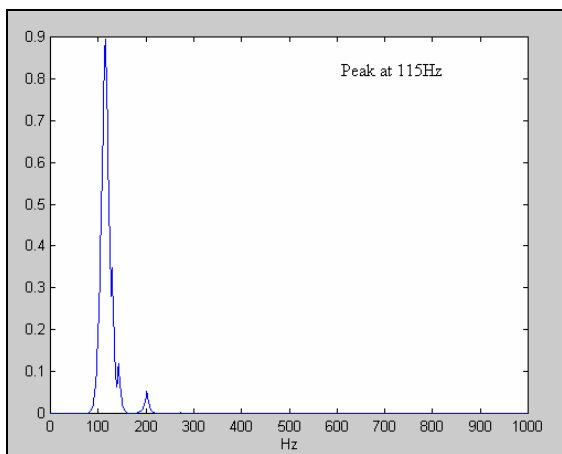


Figure 9. Output of the model for the second example

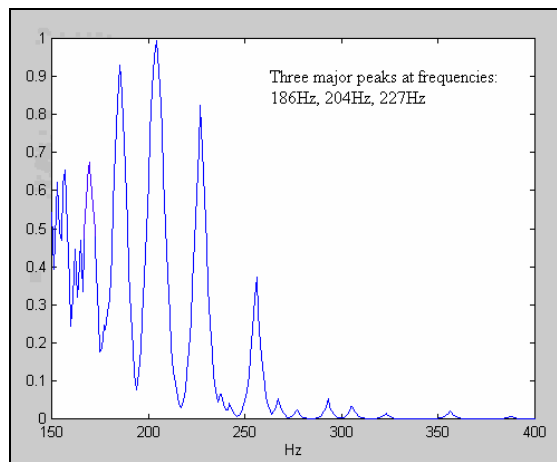


Figure 10. Output of the model for the third example

## 5. CONCLUSION

A fuzzy model for pitch identification has been developed, that is able to derive the pitch of a complex tone using the information of individual components. The basic feature of the model is that it performs all the manipulations using fuzzy numbers. Considering frequency components as fuzzy numbers with adjustable membership function width, provides the ability to impose to the manipulation a precision which is relative to the logarithm of components frequency. Thus, the necessary distance between the central frequencies of two subcomponents in order the intersection of them to cause a local peak to the output of the system is much smaller when components correspond at low frequencies than when correspond at high. Generally, the greatest peak of the derived output corresponds to the pitch of the complex tone. In the case of more than one peak with close values, the pitch can't be determined univocally.

## REFERENCES

- [1] W. D. Ward, *Subjective musical pitch*, J. Acoust. Soc. Am. 26, 369-380, (1954).
- [2] E. Terhardt, *Pitch, consonance and harmony*, J. Acoust. Soc. Am. 55, 1061-1069, (1974).
- [3] Witold Pedrycz, "Fuzzy Control and fuzzy Systems", Research Studies Press LTD, (1993).
- [4] Earl Cox, "The fuzzy systems handbook", Academic Press Professional, (1994).
- [5] J. F. Schouten, R. J. Ritsma and B. L. Cardozo, *Pitch of the residue*, J. Acoust. Soc. Am. 34, 1418-1424, (1962).