Research Article

Spectral Biomimetic Technique for Wood Classification Inspired by Human Echolocation

Juan Antonio Martínez Rojas,1 Santiago Vignote Peña,2 Jesús Alpuente Hermosilla,1 Rocío Sánchez Montero,1 Pablo Luis López Espí,1 and Isaac Martínez Rojas2

1 Department of Signal Theory and Communications, Escuela Politécnica Superior, Universidad de Alcalá, Campus Universitario, Ctra. de Madrid-Barcelona, km 33.600, 28805 Alcalá de Henares, Spain
2 Department of Forest Economy and Management, E.T.S.I. de Montes, Universidad Politécnica de Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain

Correspondence should be addressed to Juan Antonio Martínez Rojas, juanan.martinez@uah.es

Received 16 April 2012; Revised 10 July 2012; Accepted 25 July 2012

Academic Editor: Toru Otsuru

Copyright © 2012 Juan Antonio Martínez Rojas et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Palatal clicks are most interesting for human echolocation. Moreover, these sounds are suitable for other acoustic applications due to their regular mathematical properties and reproducibility. Simple and nondestructive techniques, bioinspired by synthesized pulses whose form reproduces the best features of palatal clicks, can be developed. The use of synthetic palatal pulses also allows detailed studies of the real possibilities of acoustic human echolocation without the problems associated with subjective individual differences. These techniques are being applied to the study of wood. As an example, a comparison of the performance of both natural and synthetic human echolocation to identify three different species of wood is presented. The results show that human echolocation has a vast potential.

1. Introduction

Acoustic techniques have been shown to be an invaluable tool for the study of many types of materials. Wood is an excellent material to test the limitations and potential of noninvasive acoustic methods due to its complexity and availability [1]. However, most currently used methods work in the ultrasonic range, [2, 3]. Studies of biological tissues in the audible range are scarce [4–6].

Measurement of acoustic properties of porous materials with impedance tubes is a powerful common procedure. Top quality commercial systems are available and several international standards are published [7]. Acoustic pulse reflectometry is a more recent technique for reconstructing the profile of ducts and other waveguides using pulse reflections from unknown duct geometries [8]. Other techniques have been developed for determining the properties of porous materials, mainly laser vibrometry and acoustic excitation based methods [9]. However, the use of impedance tubes is not suitable for real time monitoring of wood properties in harsh environments like sawmills or for inspecting large wood structures noninvasively.

Wood species classification using impedance tubes, acoustic pulse reflectometry in the audible range or laser vibrometry is not reported in the literature, as far as we know. As an alternative technique to echolocation pulses, stress wave analysis using acoustic impact has been successfully used by our group to classify some wood species [10]. However, scanning pulse-echo methods are more suitable to study large sections of wood, like pillars and beams.

Human echolocation enables the perception of objects through reflected echoes. Certain sounds are more appropriate for human echolocation as they produce clearer echoes. Previous studies have shown that the waveform of human palatal pulses resembles almost perfect damped sinusoidal functions, very similar to an ordinary Bessel pulse [11, 12]. Biomimetic approaches based on echolocation offer enough flexibility to inspect in a nondestructive way any kind of wood structure with very little computational effort, if only relative measures are needed. If absolute measurements...
of acoustic parameters are necessary or accurate response functions through deconvolution are to be interpreted, this method turns much more complex.

In this case, we are interested in studying the feasibility of human echolocation pulses to identify different wood veneers using as simple as possible calculations both in the time and frequency domain. More complex and powerful approaches to pulse shaping and processing, like chirping and impulse response calculations, have not been tried in order to keep the interpretation of the results and the comparison with real human performance as simple as possible. Only tone and amplitude variations among the reflected echoes from different wood surfaces are considered. Calculation of impulse response functions is not studied because participants in echolocations experiments are only trained to recognize tonal and timbral differences between echoes from different surfaces using a relative and comparative approach.

Our main objective is not to compete with well-tested acoustic nondestructive methods, but to prove that, in principle, there are enough tonal and timbral differences among the echoes from different surfaces, so that human echolocators would be able to use these variations to identify such surfaces. It is not difficult to train humans to distinguish surfaces with very different acoustic properties, like steel, glass, wood, and cork. If only natural clicks are used, there would be doubts about the reproducibility of the pulses being the cause of the timbral variations. To minimize this difficulty we have opted for a dual approach, analyzing echoes from both natural and synthetic pulses. A complete study of human performance using a statistically significant number of participants, both blind and sighted, is not the purpose of the present work and will be the subject of future research. In this work, we want to prove using both artificial and natural clicks that human echolocation could be used in principle to identify certain complex surfaces such as wood. Naturally we have used a limited number of wood species with similar finish, but different physical properties.

Obviously, natural pulses are very difficult to reproduce. Even if we use advanced mathematical techniques to try to mimic the worst behaved oscillations, precisely these pulses are the most affected by loudspeaker distortions. After many trials and different mathematical fitness approaches, we have opted for an optimal smooth function fit to a mean of the best behaved natural clicks.

In this work, the term palatal click is used in a general sense, because for human echolocation a detailed phonetic analysis of the sounds has not been performed. From a strict phonetic point of view, some of the produced sounds are not pure palatal clicks but alveopalatal or even more complex combinations, [13–15]. All participants in our study are native Spanish speakers, so they are asked only to make as clear palatal clicks as possible, trying not to touch the alveolar area with the tip of the tongue. It would be an unnecessary complication to train them to faithfully reproduce the correct palatal sounds found in Khoisan languages. Our guide in selecting “correct palatal” clicks has been the similarity of their waveforms with a damped sinusoidal curve, with minimal noise or perturbations. Although this is not a correct phonetic criterion, it is a practical one for echolocation training purposes, as our previous works found [11, 12].

The studied surfaces where square wood veneers (side = 20 cm) from three different wood species: chestnut (Castanea sativa), cherry (Prunus avium), and sycamore maple (Acer pseudoplatanus). For wood species identification, both natural and artificial palatal pulses have been used. The results show that this acoustical identification technique is theoretically feasible with good performance. Some implications derived from this research applicable to natural human echolocation will be discussed.

2. Methodology and Experimental Procedure

The experimental setup was very similar to those described in previous studies [11, 12], but with some differences. The experiments were performed in a room with dimensions of approximately 15 × 15 × 3 m. However, computer noise was not allowed, due to the objective nature of the measurements. The reverberation constant (30 dB below the level of the direct sound), R(30), was 0.10 ± 0.01 s (average and standard deviation). It was obtained from 10 recordings of the sound of a loud impulsive noise. No sound processing or enhancement has been done to study the echo signals. Sound amplitudes are normalized to facilitate figure comparison.

Three different electret microphones, with nominal frequency response from 20 Hz to 20 kHz, and two sound cards were used to verify the waveform of the recorded sounds. The sound cards, one from Realtek and one from Creative Labs were used and compared. A previous calibration with pure sinusoidal tones from a function generator showed that no important errors were present in the recordings. The sampling frequency was set to 192 kHz in order to achieve the best available temporal resolution. The recording, analysis and figures were obtained with the Praat program [16].

In order to study the echoes from artificial clicks, a 60 W nominal peak power loudspeaker (Creative Labs) was vertically placed just in front of the wood veneer also placed vertically. A millimeter ruler was used to measure the distance between the veneer target and the loudspeaker. A series of test were made in order to determine the best distance for the recordings using the loudspeaker to full power. Distances less than 10 cm produced saturated recordings, but echoes degraded for distances larger than 25 cm. So, it was decided to make the measurements using a distance of 20 cm between the wood veneers and the loudspeaker. Microphones were placed at the same distance and in parallel with the loudspeaker, 5 cm laterally separated from it. Also, a distance of 20 cm was very convenient to make recordings using natural palatal clicks in the same conditions as the artificial ones. Participants’ faces were placed just in front of the wood veneers trying to follow the same geometrical configuration than loudspeakers.

The veneer thickness was 0.5 mm approximately and their dimensions 20 cm × 20 cm. This area was selected because it was a good compromise between size and mechanical stability. Wood experts recommended us to use veneers of this size. Larger samples were prone to defects
and fractures, although it is desirable to have veneers as large as possible. This small thickness was selected to try to reduce the analysis to a 2D problem, as in wood audible sounds penetrate deeply. A full volumetric study in such a complex material as wood would be certainly much more difficult. The measured acoustical parameters are related to the structure and other properties of these different woods.

Wood experts selected a total of 5 veneers from each wood species using visual criteria of lack of defects or fractures and uniformity. Grain direction was also preserved among samples as much as possible to discard any acoustical phenomena related with this property.

Acoustics in wood depends, at least, on wood density, porosity, water content, and quantity and quality of other impregnating substances (resins, tannins, etc.). Cherry and maple, in spite of their evident visual differences, share other relevant physical and chemical properties. Their density is 630 kg/m³ approximately, and their structure, porosity and impregnating substances are similar. Chestnut, however, is quite different. Its wood density is about 550 kg/m³. This implies that it has greater porosity. In addition, it contains more impregnating substances. In this study, the effect of water content was minimized using a controlled drying process.

Water content was controlled keeping the temperature constant during experiments (294 K). Water content was measured using a commercial moisture meter over thicker samples made from the same wood as the veneers and kept in the same conditions. A value of (13 ± 2)% was obtained for all wood samples during the experiments. Wood veneer moisture was not directly measured because they were too thin to be perforated by the moisture meter.

In order to generate artificial clicks, 50 natural palatal clicks were analysed. They were produced by 5 different people. In order to simplify the analysis, they were reduced to 10 using the following criteria.

(i) Natural clicks should resemble a damped sinusoid as much as possible. Very high frequency oscillations superimposed to the main sinusoid are always different. Besides they are not reliable and difficult to reproduce.

(ii) The clicks should be simple, no more than one pulse should be visible in an interval of 500 ms. This minimizes the risk of incorrect performances due to excessively fast tongue motion.

(iii) The whole duration of the clicks should be between 100 and 200 ms. We consider this duration as the standard palatal click duration. Shorter and longer clicks tend to be irregular.

This does not mean that the discarded clicks would not be suitable for echolocation for mobility purposes. In such case, the exact form of the pulses is not critical because echolocators can use several different clicks to obtain information about their surroundings. In fact, due to natural factors, during real long time echolocation many clicks are a combination of palatal and alveopalatal sounds. A simple comparison of the timbral quality of the echoes is often enough for most purposes. However, in this case we should
demand the highest pulse purity for clicks in order to
generate something similar to a standard palatal click.

The synthetic palatal pulses were produced with the aid
of the Praat program. Ten selected clicks were individually
fitted as described below. After each function was calculated,
their peak amplitudes and duration were compared. An
exponent \( t^{1.2} \) was finally the best fit. As a result of this analysis, it was
found that a reasonable fit for well-behaved natural clicks was
described by a function of the following form:

\[
p(t) = -[0.1 \sin(2\pi 500t) + 0.01 \sin(2\pi 400t)] + \sin(2\pi 800t) + 0.002 \sin(2\pi 1400t) \\
+ 0.005 \sin(2\pi 1600t) + 0.002 \sin(2\pi 2000t) + 0.01 \sin(2\pi 2400t) + 0.01 \sin(2\pi 2500t) \\
+ 0.01 \sin(2\pi 2600t)] e^{-1000t^{1.2}},
\]

where the frequencies are in Hz and time in seconds. A
typical natural palatal click can be seen in Figures 1 and 2.
The form of this synthetic pulse can be seen in Figure 3 and
its frequency spectrum is shown in Figure 4.

The final fitting was a compromise between generality
and accuracy, so no particular click is perfectly reproduced.
In fact, the echoes from artificial sounds after being produced
by loudspeakers and recorded by electret microphones showed no differences even when more terms in the series
were introduced, so it was not necessary to increase the
complexity of the fitting function, because the noticeable
gain was negligible.

3. Results and Discussions

In the range from 1500 to 3000 Hz, some natural clicks
have a flatter spectrum than artificial ones. Better fittings
would require more terms in the series, but this accuracy
would not be reflected in the experimental results because the
used loudspeakers were not able to reproduce such accurate
functions. No effort has been made to fit the spectra over
3000 Hz. Most of the relevant information from the pulses
are below 3000 Hz and the intensity drops very quickly after
this point. Moreover, the use of higher frequencies produced
numerical instability and fitting problems, thus the use of
this range of frequencies was the best option.

The following acoustical parameters were measured from
the waveform and spectrum of every pulse: time duration,
number of relative maxima, maximum wave amplitude,
minimum wave amplitude, first frequency band, second
frequency band, third frequency band, raise time from 0
to maximum amplitude and average intensity of the pulse.
Acoustic intensity is calculated by the Praat program using
the mean value of the intensity curve in dB, with values in
dB SPL, dB relative to 2 \( \times 10^{-5} \) Pa. The frequency bands are
defined as the first three clearly resolved peaks in the audible
spectrum. The frequency spectra were finally calculated
using the FFT algorithm provided by the Praat program
with a Hamming window and 1024 discretization points.
The Praat manual specifies that the FFT is calculated after
the time domain signal is zero-padded. The results for the
frequency peaks did not change when different windows were
tried.

The same parameters were used for both natural and
synthetic pulses. The results for the echoes of synthetic
palatal clicks are summarized in Tables 1 and 2. The data are
presented as mean and standard deviation. The repeatability

Table 1: Summary of acoustic parameters of synthetic pulses from three different wood species.

<table>
<thead>
<tr>
<th>Wood</th>
<th>T (ms)</th>
<th>NMax</th>
<th>MaxWA (n.u.)</th>
<th>MinWA (n.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chestnut</td>
<td>16.72 ± 0.30</td>
<td>39.0 ± 1.7</td>
<td>0.070 ± 0.000</td>
<td>0.110 ± 0.000</td>
</tr>
<tr>
<td>Cherry</td>
<td>17.13 ± 0.26</td>
<td>39.2 ± 0.5</td>
<td>0.070 ± 0.000</td>
<td>0.110 ± 0.000</td>
</tr>
<tr>
<td>Maple</td>
<td>17.08 ± 0.46</td>
<td>35.8 ± 0.8</td>
<td>0.060 ± 0.005</td>
<td>0.100 ± 0.000</td>
</tr>
</tbody>
</table>

\( T \): time duration; \( NMax \): number of maxima; \( MaxWA \): maximum wave amplitude (normalized units); \( MinWA \): minimum wave amplitude (normalized units).

Table 2: Summary of acoustic parameters of synthetic pulses from three different wood species.

<table>
<thead>
<tr>
<th>Wood</th>
<th>F1 (Hz)</th>
<th>F2 (Hz)</th>
<th>F3 (Hz)</th>
<th>RT (ms)</th>
<th>I (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chestnut</td>
<td>59.8 ± 1.3</td>
<td>750 ± 0</td>
<td>1410 ± 0</td>
<td>0.16 ± 0.01</td>
<td>53.29 ± 0.09</td>
</tr>
<tr>
<td>Cherry</td>
<td>58.6 ± 0.9</td>
<td>795 ± 0</td>
<td>1455 ± 0</td>
<td>0.18 ± 0.02</td>
<td>52.78 ± 0.07</td>
</tr>
<tr>
<td>Maple</td>
<td>58.6 ± 1.7</td>
<td>750 ± 0</td>
<td>1455 ± 0</td>
<td>0.17 ± 0.02</td>
<td>52.65 ± 0.06</td>
</tr>
</tbody>
</table>

\( F1, F2, F3 \): frequency bands (Hz); \( RT \): rise time (ms); \( I \): intensity (dB).
bands are di
provides accurate identification information. Frequency
of the recorded echoes is not a good approach. Figure 5
most practical applications.
additional computational burden would be unnecessary for
technique is based solely on relative measurements; thus,
damping present in human tissues. Some characterization
structure of echoes produced by artificial clicks are more
first peak, as can be clearly seen in Figure 7. The peak
echolocation pulse, which maintains the maximum at the
main amplitude is displaced
show that all waveforms have a similar appearance and
detailed identification is possible using the fine structure of
the frequency spectra and their relative widths.

difficulties in peak amplitude and
time duration for the same sample were less than 5%. Criteria
for time duration measurement was amplitude of the pulse
being less than 1% over the ambient noise.
The results for the echoes from natural palatal clicks can be
seen in Tables 3 and 4. Values of both reference pulses,
without echoes, can be seen in Table 5.
In the case of synthetic echoes, the best acoustic parameters
for species identification are the main frequencies of the
spectra at frequencies below 1500 Hz. The first frequency
band is almost the same in all cases. This is related to the
to the fact that the underlying material structure of the three
samples is similar, wood in this case. Chestnut and maple
share the same second frequency band, while cherry and
maple share the third one. This permits an unambiguous
identification of the three veneers. However, as Figure 6
clearly shows, the structure of the frequency bands is more
complex than these simple parameters suggest. A much more
detailed identification is possible using the fine structure of
the frequency spectra and their relative widths.
In the synthetic pulse case, the analysis of the waveform
of the recorded echoes is not a good approach. Figure 5
shows that all waveforms have a similar appearance and
that their characteristic parameters do not provide relevant
information. Additionally, the main amplitude is displaced
from the first peak to the fourth, as opposed to the natural
echolocation pulse, which maintains the maximum at the
first peak, as can be clearly seen in Figure 7. The peak
structure of echoes produced by artificial clicks are more
complex than the natural ones, perhaps due to the lack of
damping present in human tissues. Some characterization
and deconvolution of these effects are feasible, but this
technique is based solely on relative measurements; thus,
additional computational burden would be unnecessary for
most practical applications.
Waveforms and spectra from natural palatal click echoes
provides accurate identification information. Frequency
bands are different in each veneer and their fine structure
is more detailed than that of synthetic pulses. For a good
species identification of the wood veneers, information
can be extracted from a wider range, until the 5000 Hz,
although the first main bands are sufficient. Relevant dif-
fferences between natural and artificial clicks are seen in
their waveforms. With natural pulses, all the veneers are
clearly recognized, even not quantitatively measuring their
acoustical parameters. This is explained by the contribution
of a greater number of frequency bands in natural pulses than
in the artificial clicks.
The sound reflected from the studied veneers shows a
more complex behaviour than wood density alone would
suggest. The results obtained with synthetic pulses are more
consistent with the expected density functional dependence.
Pulse duration, first and third frequency bands, and intensity,
agree in both cherry and maple. The other acoustical
parameters show the opposite tendency or are not conclusive.
The results for natural palatal clicks are not easily corre-
lated with the underlying structure. The results differences
between natural and artificial pulses could be due to a
higher reproducibility of the synthetic sounds. However, the
obtained statistical uncertainties show that this should not be
such an important issue. The far greater timbral content of
natural palatal clicks should explain most of the differences.

### 4. Implications of This Study to Human Echolocation

We have only proven, using both natural and artificial click
echoes from thin wood veneers, that, a priori, human beings
would have enough acoustical cues to distinguish several
different wood surfaces by means of echolocation, mainly
by timbral differences below 3000 Hz. Although we have
analyzed only the lowest first three resolved frequency bands,
natural click spectra shown in Figure 8 show that even more
prominent differences exist in the 1000–3000 Hz band, which
correspond to the optimum audible range for humans.
Table 5: Summary of acoustic parameters of natural and synthetic reference palatal clicks without echoes.

<table>
<thead>
<tr>
<th>Click</th>
<th>T (ms)</th>
<th>NMax</th>
<th>MWA (normalized units)</th>
<th>mWA (normalized units)</th>
<th>F1 (Hz)</th>
<th>F2 (Hz)</th>
<th>F3 (Hz)</th>
<th>RT (ms)</th>
<th>I (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural</td>
<td>20</td>
<td>16</td>
<td>0.6</td>
<td>0.5</td>
<td>50</td>
<td>470</td>
<td>800</td>
<td>1</td>
<td>73</td>
</tr>
<tr>
<td>Synthetic</td>
<td>20</td>
<td>16</td>
<td>0.8</td>
<td>0.9</td>
<td>50</td>
<td>400</td>
<td>800</td>
<td>0.93</td>
<td>100</td>
</tr>
</tbody>
</table>

T: time duration; NMax, MWA: maximum wave amplitude (normalized units); mWA: minimum wave amplitude (normalized units); F1, F2, F3: frequency bands (Hz); RT: rise time (ms); I: intensity (dB). Both pulses are considered ideal, so no errors are shown.

Figure 5: Waveforms of echoes from synthetic pulses from different wood veneers.

(a) Synthetic echo waveform of a chestnut veneer from a distance of 20 cm

(b) Synthetic echo waveform of a cherry veneer from a distance of 20 cm

(c) Synthetic echo waveform of a maple veneer from a distance of 20 cm

Figure 6: Frequency spectra of echoes from synthetic pulses from different wood veneers.

(a) Synthetic echo spectrum of a chestnut veneer from a distance of 20 cm

(b) Synthetic echo spectrum of a cherry veneer from a distance of 20 cm

(c) Synthetic echo spectrum of a maple veneer from a distance of 20 cm

Any human being who can perceive these timbral differences should be able to identify certain wood veneers. This finding would be an outstanding achievement for a person acquainted with echolocation techniques. However, the short duration of the pulses, their reproducibility, and ambient noise make it a difficult task in real life situations.

Thus, fairly advanced training would require in order to be able to identify wood veneers using human echolocation. Future studies using a large number of advanced echolocators and recordings will be necessary to test this hypothesis.

5. Conclusions

Palatal clicks could be used for surface characterization in the audible range, due to their mathematical and physical
properties. The use of synthetic pulses with similar properties is an adequate tool for human echolocation research for two main reasons. Firstly, as shown in wood veneer identification, biomimetic applications with a clear and intuitive meaning are possible, so fair conclusions about the performance of an hypothetical ideal echolocator can be deduced. Secondly, an objective comparative study of natural and artificial clicks is possible. Synthetic clicks are accurately reproducible and their main acoustic parameters can be easily tuned and modified. Thus, absolute performance of human echolocation can be objectively studied without subjective biases. Acoustical and psychoacoustical factors can be properly isolated and analyzed using this method.

Both natural and synthetic palatal clicks have the potential of high surface discrimination even in similar and complex materials, such as wood veneers. This encouraging result suggests that with proper training, human echolocation can be a very powerful aid for blind people.

References


