Sonification of Medical Images Based on Statistical Descriptors

Rodrigo F. Cadiz\textsuperscript{1,2}, Patricio de la Cuadra\textsuperscript{1}, Aarón Montoya\textsuperscript{1}, Veronica Marín\textsuperscript{1}, Marcelo E. Andia\textsuperscript{2,4}, Cristian Tejos\textsuperscript{2,3}, and Pablo Irarrazaval\textsuperscript{2,3}

\textsuperscript{1}Center for Research in Audio Technologies, Music Institute
\textsuperscript{2}Biomedical Imaging Center
\textsuperscript{3}Department of Electrical Engineering, School of Engineering
\textsuperscript{4}Department of Radiology, School of Medicine
Pontificia Universidad Católica de Chile
\{rcadiz,pcuadra,amontoy\}@uc.cl, veromab@gmail.com, \{mandia,ctejos,pim\}@ing.puc.cl

ABSTRACT

We propose an image sonification approach based on statistical descriptors, with a special emphasis on medical images. These descriptors capture several characteristics of the image data and can be beneficial for a musically interesting sonification of the visual domain. We conducted an experiment where subjects were asked to evaluate the usefulness of an extra layer of sound in tasks of identification of malignant regions in mammograms. Our results show that the addition of a musical domain to the visual data can improve the efficiency and reliability of this task. We believe that this approach could yield interesting and promising results in fields such as the sonification of other kinds of medical images or multidimensional scientific data.

1. INTRODUCTION

Sonification is a subtype of auditory displays that can be defined as the use of non-speech audio to convey information [1, p.9]. One of the most widely known examples for sonification is a Geiger detector, which responds to radiation levels by producing sound.

We are specifically interested in the sonification of images, in particular medical images which present important challenges to radiologists or image experts, where the addition of an extra layer of aural information could be of use in their analyses or diagnostics. We propose a sonification scheme based on one or several statistical descriptors of the images, in the hope of capturing relevant features of the data that could be converted into sound. In certain applications, an auditory display of the data results on a better performance than a visual one [2].

This article is structured as follows. We first proceed to examine the most important aspects of image sonification, with particular interest in the sonification of medical images. Second, we describe the proposed sonification scheme, detailing the selected statistical descriptors and sound synthesis methods. Third, we describe the experiment. Fourth, we discuss the main results and finally, we present the most important conclusions of our work.

2. IMAGE SONIFICATION

The conversion of data between the visual and audio domain has been rapidly evolving and becoming a very active field of research during the last three decades [3]. There are several approaches for the conversion of image data into sound. Mathematically, the process of image sonification can be formulated as the mapping between a two-dimensional dataset (two spatial dimensions) to a one-dimensional one (one temporal dimension) [3]. For this reason, structures such as helicoids [4] or isomorphical mappings [5] are interesting approaches for such a translation and it has been shown that it is possible to implement one-to-one invertible image to sound mappings while preserving visual information [6].

2.1 Image sonification in medicine

In the field of medicine, sounds are sometimes used for diagnosis. A stethoscope is a normal part of a doctor’s equipment. Physicians are trained to listen to tissues rubbing in the lungs, gasses bubbling in the intestines, heart valves opening and closing, and blood pumping through veins. However, many other indicators, such as vital signs, are measured and shown as graphs, which could be distracting during visually demanding tasks in an operation, and it is possible to synthesize sounds to represent these indicators instead [7]. As a result, acoustic rendering has been successfully applied as an additional information channel and/or warning signal in biomedical signal analysis and data presentation [8]. Indeed, medical students performed better in a simulated operation when dynamic variables about the health of the patient were presented as sounds rather than graphs, and better with sounds alone than with both sounds and graphs combined [9]. In real surgery applications, tactical audio as an audio feed-
back have been used as support for precise manual position-
ing of a surgical instrument in the operating room [8].

Magnetic resonance imaging (MRI) has become one of the
most important means of finding symptoms of disease inside
a patient’s body. However, it is very difficult to visually detect
unhealthy regions of the brain in an MRI image, because of
the nature of brain tissue [7]. In [10], damaged regions of the
brain were made more distinguishable by the sonification
of image textures. Listening to the image data, in addition to the
visual display, may help a physician to diagnose a dangerous
illness which might otherwise go undetected [7].

Texture plays an important role in image analysis and has
been widely adopted in medical imaging problems as well
as related areas such as computer vision and pattern recog-
nition [11]. In MRI, texture analysis can be used to identify
diseased tissues, however, with MRI exams of the brain, vi-
sual discrimination between normal and damaged tissue still
remains as a difficult task [10]. Statistical descriptors are a
useful tool to discriminate and classify textures in images.

where the detection of microcalcifications is of extreme diffi-
culty and urgency, as breast cancer is the most prevalent can-
cer that leads to death in women today [11]. Mammograms
are images of the breast obtained by X-rays, and they are cur-
rently the most effective way to detect breast cancer at an
early stage. Figure 1 shows a standard mammogram.

Radiologists look for the presence of tiny accumulations of
calcium, which are seen as white spots in the image mixed
with the background or other tissues present in the breasts.
These microcalcifications have a diameter of less than 0.5 mm
and they are grouped in clusters, as shown in figure 2. Their
visual detection over an homogeneous background or tissues
is very difficult [11].

3. PROPOSED SONIFICATION

We propose a parameter mapping sonification [12] based on
the capture of relevant statistical descriptors of the images and
using them to drive perceptually meaningful sound synthesis
parameters. We are particularly interested in mammograms,

3.1 Statistical descriptors

Statistical descriptors intent to describe different characteris-
tics of an image based on the pixel’s value information. In
our proposal, all these descriptors are calculated on a region
of interest (ROI) of width $n$ and height $m$ selected by the user.

The descriptors that we used were maximum, minimum,
mean, standard deviation, skewness and kurtosis. If we ex-
amine the formulas for standard deviation, skewness and kur-
tosis, it is clear that all these descriptors compare in some
degree the pixels in the ROI with the mean. This is some-
thing desirable in our case, as microcalcifications appear in
small regions of a few pixels with a relatively high intensity
value compared to the rest of the region of interest.

We also considered other kinds of image descriptors. We
studied all descriptors available from the MATLAB Toolbox
Balu [13]. Using this toolbox, we conducted a classification
test with a set of high resolution mammograms obtained at
our university medical center. It turned out that some coefficients of the Discrete Cosine Transform (DCT) were very useful at discriminating images with and without microcalcifications, so we decided to include two of these DCT coefficients (DCT1 and DCT2) into the set of descriptors used in the sonification. Our interpretation of the usefulness of these coefficients is that they adequately capture the relative size of the microcalcifications in relation to their surroundings.

3.2 Synthesis methods

We initially started with ten different synthesis methods. These methods were carefully chosen after a significant amount of experimentation and fine-tuning inside the research team, with a special emphasis in their musicality. Although sonification of data does not have to be musical, we strongly feel that if these sounds are designed by musicians, as in this case, the results would be more appealing. This is of special importance in the case of radiological images, because these professionals can diagnose hundreds of images per day and the sounds they hear must be pleasant and keep their interest. All the frequency and intensity ranges were adjusted based on these criteria.

On a second stage, we discarded five of the initial synthesis, as discussed in section 4, and we re-designed the sound synthesis, as follows:

- Sonification B1. Karplus strong. We used a soft damping effect preventing the string from becoming silent in the presence of a black ROI.

- Sonification B2. Bell granulator. We used kurtosis and skewness to drive the soundfile offset and grain duration parameters.

- Sonification B3. Geiger counter. We decided to use kurtosis as the repetition frequency and the skewness to control the intensity of the sound.

- Sonification B4. Multiparameter 1. We mapped kurtosis to the carrier frequency and skewness to harmonicity. The repetition rate is controlled by the standard deviation. We also incorporated a low pass filter driven by kurtosis and panning given by DCT2.

- Sonification B5. Multiparameter 2. We implemented a Karplus-Strong algorithm driven by kurtosis and skewness. The rest is exactly as described for B4.

4. EXPERIMENT

We conducted a first experiment with twenty-three subjects, with no special training or background in image processing or medical imaging. Eight of them had musical experience. We showed them three randomly chosen mammograms with one microcalcification region marked by a radiologist with a red square. Each subject looked and heard each of the images with each of the ten proposed sonifications (A1 through A10).

In [12], the authors propose that an effective parameter mapping sonification necessarily involves a compromise between intuitive, pleasant and precise display characteristics. Based on this assumption, we asked the subjects the following two questions:

1. Do you hear a difference when the ROI is inside the microcalcification zone compared to the rest of the image?

2. Is the sound you hear pleasant?

We tabulated the subject’s responses using a four point scale: Strongly disagree, Disagree, Agree, Strongly Agree. This first stage allowed us to discard five of the initial sonifications. We removed the ones that were not very useful in telling the microcalcification zones apart and the ones that were not pleasant.

On a second stage, we used sonification B1 through B5. We conducted a new experiment this time with eight subjects: students, researchers, or faculty of the university’s Biomedical Imaging Center, all of them with some degree of experience with medical images. We presented to them two images with one microcalcification region marked by a radiologist with a red square. Each subject looked and heard each of the images with each of the five new sonifications. We asked them the very same questions as before. We tabulated the subject’s responses using a four point scale: Strongly disagree, Disagree, Agree, Strongly Agree.

5. RESULTS

![Figure 3](image_url) Subject’s ratings in regard to the ability of each sonification to sound different in regions with and without microcalcifications.

Figures 3 and 4 show the subject responses to the two questions. In relation to the first one, figure 3 seems to imply that sonifications B2 and B4 allow for an audible distinction between regions of the image that contain microcalcifications compared to regions without microcalcifications. This is promising, considering that the nature of the images we are working with allow for a high number of false positives, so it is normal to expect a relatively low rating for this question.
As our objective is to someday incorporate this audible data channel to radiological practice, and the final decision will still be made by the human radiologist, false positives are not really a problem for us. Figure 4 shows that sonifications B2, B3 and B4 are pleasant to be used on an actual radiology task. As one radiologist could review hundreds of images a day, we think it is very important to provide them with pleasant and musically appealing sounds.

6. CONCLUSIONS AND FUTURE WORK

We can conclude from these preliminary two experiments that sonifications B2 and B4 are good candidates to be incorporated into radiological practice. We need to further test these sonifications with students and professional radiologists, and we hope that the addition of an extra sonic layer for the visual display could allow for a better performance in their diagnosis. As a more general conclusion, we once more have obtained evidence that the human auditory systems is capable of providing a very good alternative and complement to the usual visual way of analysis and understanding complex scientific data, and in particular, mammograms.

Acknowledgments

This project was funded by interdisciplinary grant No.3/2011 of Vicerrectoría de Investigación, Pontificia Universidad Católica de Chile

7. REFERENCES


