Crowd-Sourcing the Sounds of Places with a Web-Based Evolutionary Algorithm (Detailed Implementation and Results)

Alexander E.I. Brownlee
Computing Science and Mathematics
University of Stirling
Stirling, UK
sbr@cs.stir.ac.uk

Suk-Jun Kim
School of Language, Literature, Music and Visual Culture
University of Aberdeen
Aberdeen, UK
s.kim@abdn.ac.uk

Szu-Han Wang
Centre for Clinical Brain Sciences
(Section of Synaptic Biology)
University of Edinburgh
city
s.wang@ed.ac.uk

Stella Chan
School of Health in Social Science
(Section of Clinical Psychology)
University of Edinburgh
stella.chan@ed.ac.uk

Jamie A. Lawson
School of Language, Literature, Music and Visual Culture
University of Aberdeen
Aberdeen, UK
r01jal16@abdn.ac.uk

ABSTRACT

The sounds that we associate with particular places are tightly interwoven with our memories and sense of belonging. It is assumed that such an association is a complex process, and much of its mechanism is hidden from analytical examination. The association of sound to place has been much explored and examined by artistic approaches. For example, soundscape composition, which makes great use of recorded and barely-processed sounds from places in the compositional practice, highlights the power of the association. However, it does not offer us a scientific insight into its process, particularly, the role of familiarity of sounds people hear and their association to specific places. We describe a platform designed to assist in gathering the sounds that a group of people associate with a place. A web-based evolutionary algorithm, with human-in-the-loop fitness evaluations, ranks and recombines sounds to find collections that the group rates as familiar. An experiment involving independent groups of people associated with four geographical locations shows that the process does indeed find sounds deemed familiar by participants.

CCS CONCEPTS

• **Computing methodologies → Genetic algorithms**; • **Applied computing → Sound and music computing**;

KEYWORDS

sound, soundscapes, human-in-the-loop, mood

This technical report is provided in support of the paper "Crowd-Sourcing the Sounds of Places with a Web-Based Evolutionary Algorithm" published at GECCO 2019, Prague, Czech Republic.

1 INTRODUCTION

The collective memories of a place and its people are tightly interwoven with identity and a sense of belonging. Sound is a fundamental part of memory. Teasing out the aural identity of a place is a powerful way to understand the collective memory of that place [3]. Artists have explored the identity of place through sound, but this is challenging because the way we recognise a place through sound is a complex process, and elicited sounds are inevitably conditioned by preconceptions of the artists. Existing strategies to explore sounds typically involve listening to long recordings from the place (soundscapes) and gathering information subjectively or using frequency analysis [11, 12, 14, 15, 22]. No tool exists to assist a community to formulate their collective sounds for a given place of interest (POI).

We demonstrate the Distributed Evolutionary Algorithm for collecting the Sounds of Places (DEASP) as a tool for the formulating such soundscapes for a POI. DEASP uses short samples as the basic building block of each sound. An evolutionary algorithm (EA) is run periodically to assemble samples into groups. A web-based interface then allows participants to evaluate these groups, and these evaluations are stored for use as a contribution to fitness when the EA runs. The interface also allows participants to upload their own samples. The value of this system over a simpler voting-based approach for samples is that the iterative aspect of an EA allows participants to hear combinations of sounds chosen by others and rank them too. This means that the final choices are arrived at collaboratively.

In a case study focusing on four geographical POIs, with 82 participants, we used DEASP to generate collections of sounds representative of each POI. An independent measure of familiarity was taken during the experiment, and this demonstrates that the platform is effective in finding sounds that the group collectively feels to be familiar. Participants were also asked to indicate their moods before and after evaluating the sounds. We note some further observations made as a result of this process.

The main contributions of this paper are: a new application of EAs to a previously untackled problem in art and design; the use of crowds as part of a fitness function; and a new fitness metric designed to cope with the varying response rates for each solution.

We begin by noting some related work in soundscapes and human-in-the-loop EAs. Then we describe the DEASP platform and methodology, followed by details of our experiment with participants drawn from the general public. Finally we present results

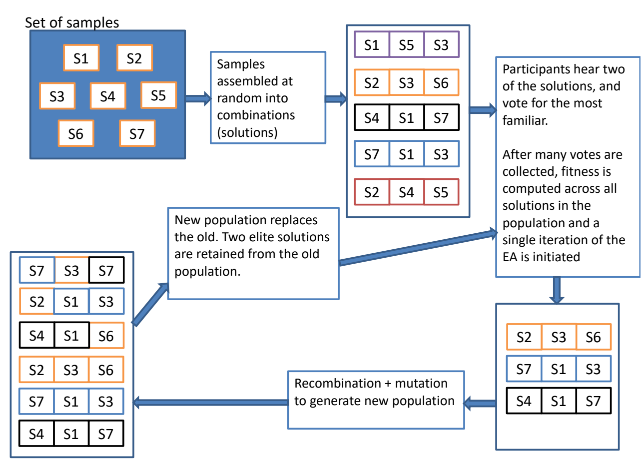


Figure 1: The overall framework of DEASP. Sx refers to a sound sample; solutions here are shown with three samples but the solutions in our study were formed of ten.

of the experiment including the artistic outputs, and draw our conclusions.

2 RELATED WORK

2.1 Soundscapes

The role of sound in how we recognise places — a sense of place — has been extensively examined and explored in acoustic ecology, ethnomusicology, sound studies and artistic practice [3, 10, 13, 18, 20, 24, 25, 27]. Particularly, most of soundscape compositions and many electroacoustic compositions have focussed on the effect of sounds recorded from specific places [11, 12, 14, 15, 22]. Such artistic practice, however, rely much on a belief that in people's awareness of places, sounds somehow play a role, a personal and subjective assumption that has not been fully examined.

2.2 Human-in-the-loop evaluation

Human-in-the-loop and interactive EAs are far from new. They are commonly used in situations where there is no clear fitness function [1, 2]. Examples include numerous applications in search-based software engineering [19] and image segmentation [17]. Creative applications are an obvious target for such an approach, including evolving animations to fit aesthetic preferences [28], assembling components of images imitating works of the artist M.C. Escher [7], constructing scenes of medieval towns [23], generating music [16], and the "Mondriaan Evolver" [29]. Direct relationships between sounds and mood have also been explored by researchers in evolutionary computation [21].

A major distinction is that each of these approaches are single-user. DEASP attempts to capture the preferences of a large group of individuals, eliciting a "collective" sound associate with a given place.

3 METHODOLOGY

The DEASP software platform was implemented as a web application, running on PHP and MySQL. The overall framework of DEASP is presented in Figure 1. Clearly the main loop is a familiar structure of an EA. The representation has each solution being a list of samples (in our application, a solution is 10x 5 second samples). The code driving the application has been designed to be easily adapted to serve sounds for other places or topics than those in the present study.

Overall, the EA generates a population of solutions (each solution being a sound made up of a sequence of samples). Participants evaluate pairs of these solutions, choosing the one they find most familiar with respect to the POI. When enough of these choices (votes) are collected, they are converted to fitnesses, and a single iteration of the EA runs to generate a new population. The process then repeats. We now give more detail on each of the major components of DEASP.

3.1 Evaluations by human participants

When a new population is ready, all participants are emailed with a link to hear some sounds. Participants then engage with the system via a web interface, part of which is shown in Figure 2. The web application was also developed to be mobile-compatible, using a responsive layout and HTML5. Upon logging in (email address + password, or a Google account login), participants' workflow is as follows:

1. Login. The participant logs in, and completes a form to test their mood. This follows the I-PANAS-SF [26] template, a standard method for evaluating moods in psychology. Participants rate their mood from 1-5 [1] for each of upset/ hostile/ alert/ ashamed/ inspired/ nervous/ determined/ attentive/ afraid/ active.

2. Evaluation – main. The participant is presented with a pair of solutions, which can be played in turn (multiple times if desired). A screenshot of this form is given in Figure 2. These two solutions are chosen at random from the population corresponding to the places that they have previously indicated that they know well (participant self-identified their known places on joining the study). Once both the sounds have been played, the participant is able to choose which one they found most familiar.

3. Evaluation – control. As a control, for each place in the study, a separate population was maintained, for which fitnesses would be provided by participants who did not know the place well. At this point, the participant is invited to hear two sounds from one of these "control" populations. (so, if the participant indicated that they knew Stirling well, at this point they would be played sounds for Aberdeen)

4. Additional sounds. The participant is then given the option to evaluate more sounds for their "familiar" place if they wish.

5. Finishing. The participant then completes a second evaluation of their moods prior to logging out.

The whole process above takes around 5-7 minutes; or longer if the participant elects to evaluate additional sounds.

People have a natural limited attention span, and, like fatigue, a loss of engagement with the process has been shown to lead to

[1] Specifically, 1: Very slightly or not at all; 2: A little; 3: Moderately; 4: Quite a bit; 5: Extremely

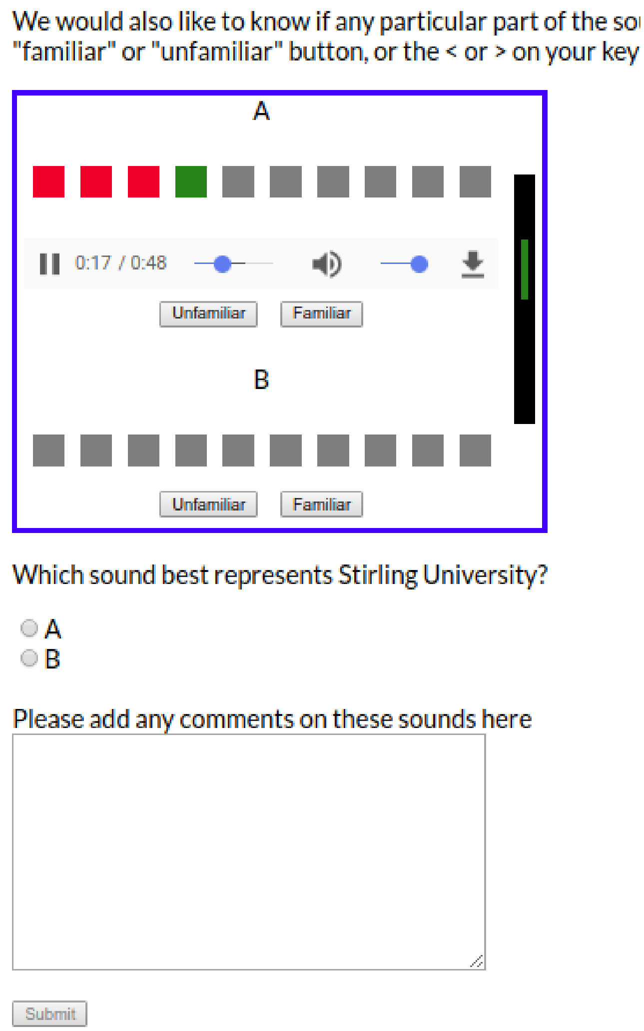


Figure 2: The web interface for evaluating sounds.

increasingly erratic decisions as the user performs more and more evaluations [5]. This is the reason we elected to have participants only listening to two pairs of sounds (one familiar place and one unfamiliar place) in each visit, with the option to listen to more. We were then able to aggregate these votes into a fitness representing the preferences of all participants for each place. Even so, the EA generating the solutions was designed to keep a small population size with higher variation and selective pressure than in conventional applications, to keep the solutions varied and to ensure that participants could see clear progress from one visit to the next.

3.2 Evolutionary algorithm

The evolutionary algorithm follows a conventional fixed-length integer encoding. Assembling the components of a solution in a design problem in this way is referred to by [8] as *component-based representation*. Solutions have ten integers, each of which is an ID representing a sound sample in the database. As noted above the population size is small, with 10 solutions. These are generated at random, with the only constraint being that no sample may appear in the same solution more than once.

The EA is paused after each generation while participants evaluate the sounds. Once this phase is complete, their votes are converted to fitnesses, as described in the next section.

The 2 fittest solutions were maintained as elites from one generation to the next. The remaining 8 solutions are generated via crossover and mutation. To achieve rapid progress, 3-tournament selection (rather than 2-tournament) is used to select parents. Furthermore, every offspring is subjected to crossover and mutation: crossover is standard 2-point, and mutation is simply the replacement of one sample in the solution with another chosen at random from the database (restricted to prevent repeats in one solution).

3.3 Fitness evaluation

The most novel component of the EA we implemented lies in the fitness evaluation. During the evaluation phase, participants are evaluating solutions by comparing one with another. A count is kept of the number of times each solution was chosen by a participant. This continues until a number of evaluations have been completed by participants: in our study this was simply when no participant had evaluated a solution for more than 7 days. The votes of participants are used to compute a fitness for each solution.

Our fitness measure is designed to accommodate three factors:

(1) solutions which where chosen in a higher fraction of evaluations (i.e. were more often deemed to be more familiar than another solution) should have higher fitness
(2) the pairs of solutions presented to each participant for evaluation are chosen at random: so not all solutions have had the same number of chances to be voted for; furthermore some might not have received any votes
(3) of solutions which have won the same fraction of votes they have been involved in, we would prefer those that have been evaluated most often as we can have more confidence that they reflect the preference of the crowd. i.e. if A was evaluated 6 times and won 3 times, and B was evaluated 8 times and won 4 times, we prefer B.

Consequently, the function takes the following form. The fitness f of a solution x is given by:

$$f(x) = \begin{cases} v(x)/t(x) + d * t(x)/10000, & \text{if } t(x) > 0 \\ 0.5, & \text{otherwise} \end{cases} \tag{1}$$

where $v(x)$ is the number of times that solution was voted for by a participant, $t(x)$ is the total number of times is was presented to a participant for evaluation, and:

$$d = \begin{cases} 1, & \text{if } v(x)/t(x) \geq 0.5 \\ -1, & \text{otherwise} \end{cases} \tag{2}$$

$v(x)/t(x)$ reflects the proportion of times that the solution was deemed to be more familiar than another solution. This will be in the interval $[0, 1]$, with any $v(x)/t(x) > 0.5$ indicating a solution deemed to be familiar more often than unfamiliar. $t(x)/10000$ acts as a tiebreaker: where two solutions have the same $c(x)/t(x)$ ratio, the one that has been evaluated more often will be given a strong weighting. That is, given either a higher fitness (for familiar solutions where $v(x)/t(x) \geq 0.5$) or a lower fitness for unfamiliar solutions. The 10000 is arbitrary; it would need to be larger if very large numbers of votes are anticipated. In the rare situation where a solution has not been evaluated by any participant, it is given a fitness of 0.5.

3.4 Independent validation of solutions

We were interested to know whether the approach would produce sounds that were deemed to be more familiar, but simply using the fitnesses or votes returned by participants is obviously not a suitable measure: by definition the EA will be producing solutions that maximise these measures. As a more independent test, we added a *clicker*. As the sounds played, the participant could register specific samples as familiar or unfamiliar by pressing < or > on their keyboard, or clicking buttons on-screen. This would trigger the visual response of a green or red bar growing on-screen when a button was pressed, then shrinking back to zero if no further presses were made (this is visible to the right of Figure 2). The total familiar/unfamiliar clicks for each sample were logged for each partipant's interactions.

4 EXPERIMENT

We now describe a live run of the DEASP platform with volunteer participants, covering four target places of interest. These POIs consist of four locations: University Campus and City Centre in both Stirling and Aberdeen.

4.1 Collecting the sound samples

16325 sound samples of 3-5s length were collected, spread evenly over the four POIs, by making longer recordings and dividing them up manually. These sounds are a useful artistic output in themselves, complementing existing soundscapes (long recordings of places). These sounds were used as the basic building blocks for the optimisation process.

Participants were also given the option of recording and uploading their own sounds, but during the live trial no participant ever did this.

4.2 Participant Recruitment

Participants were created by advertising within the Universities of Stirling and Aberdeen, through the local press, and via social media. Each participant was able to sign up by him- or herself, and self-identified with a subset of the places in the study. In addition to this, we collected some basic demographic information (age, gender, occupation) and the moods that each participant associated with their selected places in general. These latter points of information have been anonymised: we intend to use these in later work to measure any impact that these factors may have had on mood and sound evaluations.

4.3 Implementation

The platform was run using 8 independent EAs, each with its own population. This was, for each place in the study, one EA for those expressing some connection to the place (the "main" population), and a "control" population for people unfamiliar with the place. The web platform presented the evaluations for the separate EAs seamlessly together to partipants, so they could, for instance, evaluate sounds for their familiar and an unfamiliar place in one short session.

5 RESULTS

82 volunteer participants were recruited and engaged with the DEASP platform over October 2017 to January 2018. Broken down, of these the numbers expressing some connection to each place were:

(1) Stirling City Centre: 30
(2) Stirling University: 39
(3) Aberdeen City Centre: 65
(4) Aberdeen University: 60

Each volunteer participated in 1 to 5 sessions, during which they listened to sounds, producing a total of 194 responses. Each response included mood, a preference of groups of sounds for one familiar and one unfamiliar place, and the independent ratings of the individual samples they had heard.

As might be expected, all volunteers were invited to participate in all iterations of the EAs for their familiar places, but many did not follow the invitation every time. Consequently, a major issue with the process was a rapid drop off in engagement, which led to the EA only running for 1-2 generations in some of the places, and only reaching 5 generations at most. Despite this, there did appear to be some convergence: the samples appearing in the final populations were surprisingly consistent. The fittest sounds found by each run can be found here: https://www.whatisthegrid.co.uk/FinalSounds.html

The number of generations reached by each EA run was as follows:

(1) Stirling City Centre: 2 (main) 4 (control)
(2) Stirling University: 4 (main) 3 (control)
(3) Aberdeen City Centre: 5 (main) 2 (control)
(4) Aberdeen University: 4 (main) 3 (control)

5.1 Improved familiarity

While the final generated sounds are interesting in themselves from an artistic perspective, we were interested to see if they were rated as more familiar by the participants. For each sample present in the starting population for each place, we took the total number of "familiar" clicks and subtracted from this the number of "unfamiliar" clicks. This gave a measure of familiarity for each sample played to participants. We repeated this process for the samples present in the best solution in the final population for each of the EA runs.

The absolute values, even among the random starting population, are generally positive (more familiar than not). This makes sense because the sample were all recorded in the broad geographical areas, so will in some way be familiar with respect to those places, although the point is that the level of familiarity will vary greatly with each sample. We note that most participants would only click "familiar" once on each sample, and for many samples there were neither familiar or unfamiliar clicks. e.g. for the samples present in the initial population for Aberdeen city centre, in total 800 playbacks of samples were heard by participants, but "familiar" was only clicked during 300 of them. The distribution showing

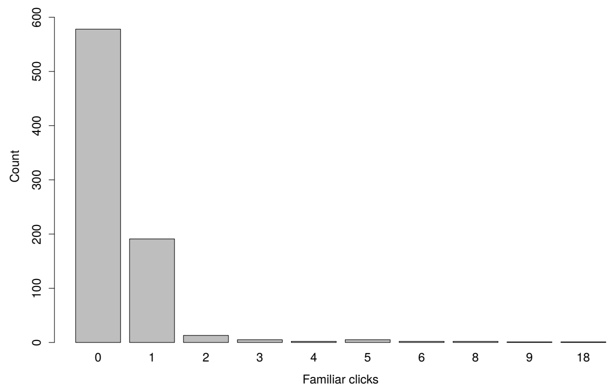


Figure 3: Distribution of "familiar" clicks for Aberdeen city centre during first generations' evaluations: the number of samples that received a given number of "familiar" clicks

how often samples received n clicks for this location is shown in Figure 3; the same was seen for the other populations. After zero clicks, samples often received one click, then rarely received more.

Figures 4 and 5 show boxplots of these figures for each location, for the main and control populations. We focus on Figure 4c shows boxplots of these figures for Aberdeen City Centre, the place which had enough participation to reach 5 generations and has the clearest result. It is clear there is an overall uplift in familiarity for the best solution compared to the starting population (also note the noise in the starting population shown by several outliers). A two-tailed Wilcoxon rank-sum test[2] on these two groups found $p = 0.001416$ (i.e. < 0.05), showing the difference to be statistically significant. However, while significant, the absolute increase in value does not appear to be large (median increasing by 1-5 depending on location). In respect of this we note that, given the distribution of "familiar" clicks in Figure 3 for random solutions, actually an increase of even one click per sample shows a substantial increase in familiarity.

In contrast, for the control population (Figure 5c) there is no significant uplift in the familiarity ($p = 0.2619$), which is reasonable because the participants had no familiarity with the place.

We note that of the other locations, the same trend is shown, except for the main population of Stirling city centre (which only reached enough evaluations for one generation) and the control population for Aberdeen university. The latter records a strong familiarity rating in the best solutions found, even though the participants were unfamiliar with the place: though it does appear that one participant was responsible for a large number of "familiar" clicks in this case. The corresponding Wilcoxon rank-sum test p-values for all places is inset in the each of figures. We take these results as evidence that the platform is indeed able to find a consensus of familiar sounds for a group of people that share a connection to a given place, at least where the participation is high enough for more than two generations to take place.

[2]the distributions were found to be non-normal

5.2 Plotting the convergence

An visualisation of the study has been generated through plotting the convergence of the populations. A bump chart[3] showing the geographical locations over the course of evolution for Aberdeen City Centre is given in Figure 6. This neatly reflects the organic process that the evolution has followed and can be viewed as an additional artistic output in its own right.

The bands represents groups of samples collected from different places around the City Centre. The width of each band reflects the proportion of the samples in each generation that are drawn from that location. At the start of evolution, the distribution of samples is roughly uniform; in the final generation four locations account for more than half the samples: these account for the busiest parts of the City Centre streets, in contrast to the remaining locations which are largely quieter streets and parks.

5.3 Decrease in moods

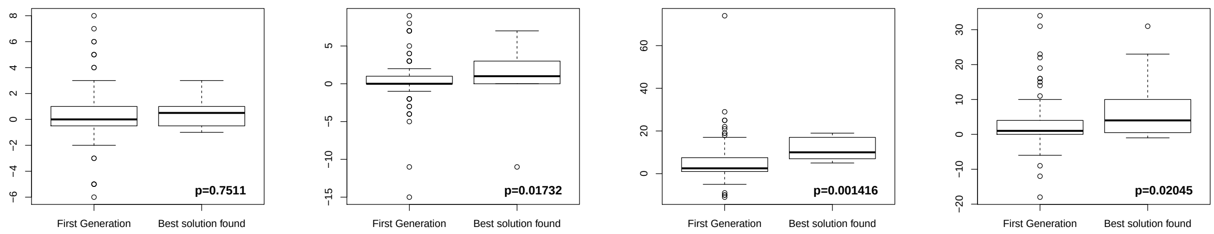
We also observed that across the samples where participants fully completed their mood before and after evaluating the sounds (69 evaluations), several of the moods reduced in strength. Table 1 summarises this result. *mean change* is the arithmetic mean, over all the evaluations, of the score given for each mood at the end of the evaluation minus the score at the start. So, negative values indicate a drop in the value assigned to that mood over the course of the evaluation (moods were all rated on a scale of 1-5). A two-tail paired t-test was computed between the groups of "before" and "after" scores for each mood, and the p-value for this test is also given in the row marked p (note that the scores for each mood across all evaluations was found to be normally distributed using a Shapiro-Wilk Test). The table shows that the scores for *determined*, *attentive* and *active* all decreased by a statistically significant margin during the evaluations. We believe that this is a potentially important factor to consider for any human-in-the-loop evaluations if further evaluations are carried out in quick succession (this is an additional confounding factor to the known problem of user fatigue [5]).

6 CONCLUSION

We have proposed DEASP, a platform for finding the sounds that a large group of people associate with a given place. DEASP uses an EA to take the feedback from human participants and generate new sounds for their evaluation. We proposed a new fitness metric designed to cope with the varying response rates for each solution. The platform successfully generated interesting sounds that are valuable on an artistic basis. We have been able to show statistically that the system finds sounds regarded by the group as more familiar than sounds chosen at random, and analysed the change in moods experienced as part of the process.

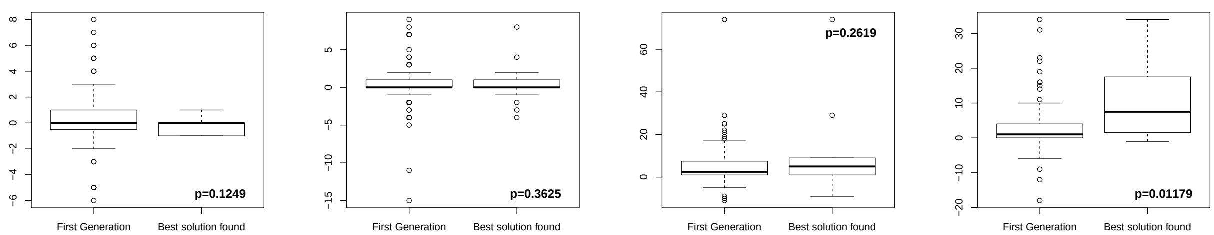
We propose that a similar platform could also be deployed for other applications such as collaborative design or decision making where a group needs to choose components based on qualitative evaluations of them in combination (for example, choosing the content of a standard toolkit). Study on this topic is one planned direction for future work. It would be interesting in investigate suitable surrogate fitness models [4, 9] to extend the search without increasing the number of human sound evaluations.

[3]https://rawgraphs.io/learning/how-to-make-a-bump-chart/



(a) Stirling city centre (b) Stirling university (c) Aberdeen city centre (d) Aberdeen university

Figure 4: Independent familiarity test results for main population for each location (i.e. people who identified as knowing the place). The corresponding Wilcoxon rank-sum test p-values is inset on each plot.



(a) Stirling city centre (b) Stirling university (c) Aberdeen city centre (d) Aberdeen university

Figure 5: Independent familiarity test results for control population for each location (i.e. people who did not know the place). The corresponding Wilcoxon rank-sum test p-values is inset on each plot.

mood	upset	hostile	alert	ashamed	inspired	nervous	determined	attentive	afraid	active
mean change	-0.01	0	-0.14	-0.03	-0.1	-0.12	-0.36	-0.36	-0.04	-0.28
p	0.69	0.84	0.15	0.62	0.45	0.18	0	0	0.45	0.01

Table 1: Changes in mood between start and end of each evaluation

Several studies point the importance of the spatialisation of sounds in the characterisation of a place [6]. Each place presents a specific acoustic signature (reverberation, diffusion, etc.), and while the work here implicitly considers acoustics for each place by virtue of using recordings taken there, it would be value to explore methods to tease this out separately. Finally, we also plan to investigate further the connections between participant demographics, mood and each place.

ACKNOWLEDGMENTS

This work was supported by a grant from the Royal Society of Edinburgh (Scottish Crucible). Also thanks to Prof Peter Stollery who assisted in the early stages of the project, particularly planning of sound collections; and to David Clark, a BSc student who assisted with the web application implementation.

REFERENCES

[1] Hamzeh J Aljawawdeh, Christopher L Simons, and Mohammed Odeh. 2015. Metaheuristic design pattern: Preference. In *Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation*. ACM, 1257–1260.

[2] Peter J Bentley and David W Corne. 2002. -An introduction to Creative Evolutionary Systems. In *Creative evolutionary systems*. Elsevier, 1–75.

[3] Karin Bijsterveld. 2013. *Soundscapes of the Urban Past : Staged Sound as Mediated Cultural Heritage*. Transcript Verlag.

[4] A. E. I. Brownlee, J.R. Woodward, and J. Swan. 2015. Metaheuristic Design Pattern: Surrogate Fitness Functions. In *MetaDeeP Workshop, Proc. GECCO Companion*. ACM Press, Madrid, Spain, 1261–1264.

[5] Praminda Caleb-Solly and Jim E Smith. 2007. Adaptive surface inspection via interactive evolution. *Image and Vision Computing* 25, 7 (2007), 1058–1072.

[6] Elvira Di Bona. 2017. Listening to the Space of Music. *Rivista di estetica* 66 (2017), 93–105.

[7] AE Eiben, R Nabuurs, and I Booij. 2002. The Escher evolver: Evolution to the people. In *Creative Evolutionary Systems*. Elsevier, 425–439.

[8] A. E. Eiben and J. E. Smith. 2015. *Interactive Evolutionary Algorithms*. Springer Berlin Heidelberg, Berlin, Heidelberg, 215–222. https://doi.org/10.1007/978-3-662-44874-8_14

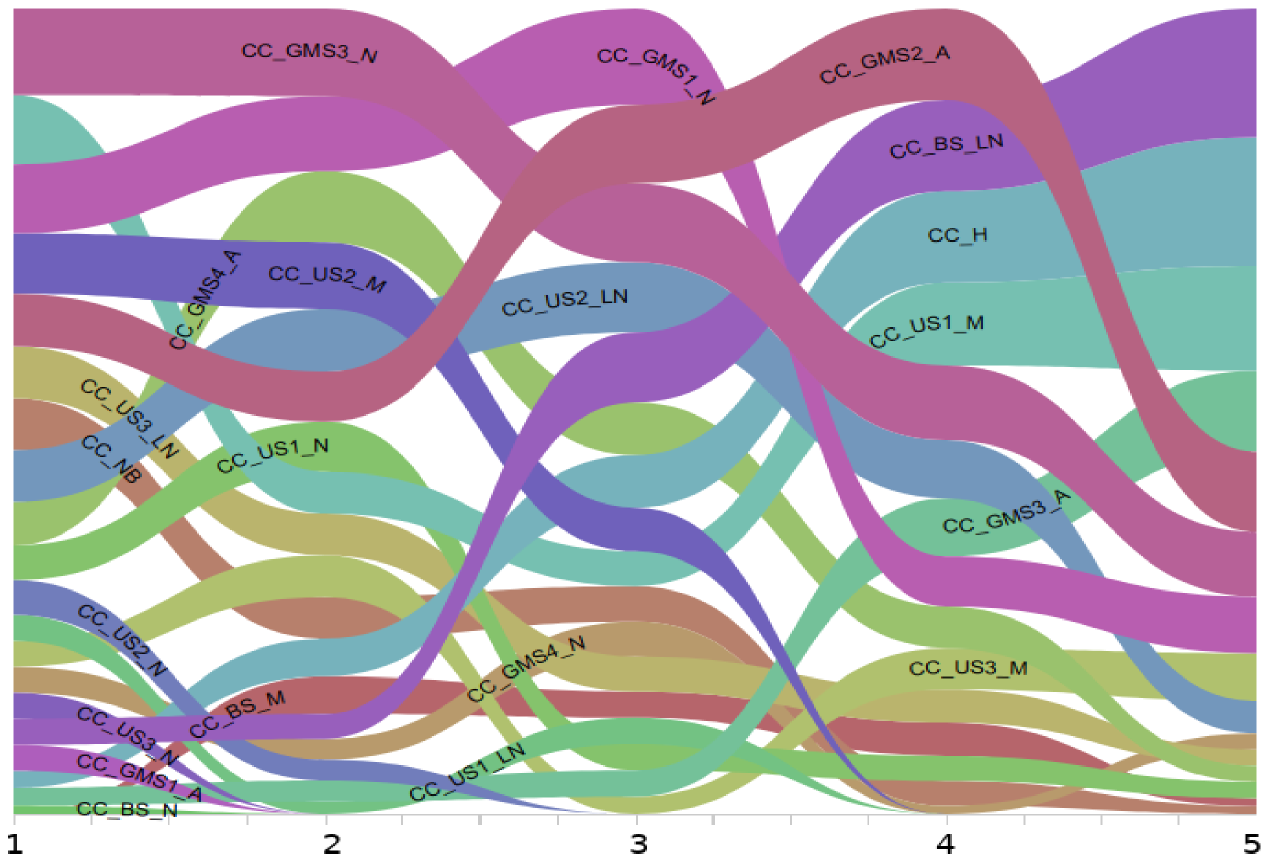


Figure 6: Visualisation of the evolutionary process.

[9] Yaochu Jin. 2011. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm Evol. Comput.* 1, 2 (2011), 61–70. https://doi.org/10.1016/j.swevo.2011.05.001

[10] Ari Y. Kelman. 2010. Rethinking the Soundscape: A Critical Genealogy of a Key Term in Sound Studies. *The Senses and Society* 5, 2 (2010), 212–34.

[11] Suk-Jun Kim. 2010. Imaginal Listening: a quaternary framework for listening to electroacoustic music and phenomena of sound-images. *Organised Sound* 15, 1 (2010), 43–53.

[12] David Kolber. 2002. Hildegard Westerkamp's Kits Beach Soundwalk: shifting perspectives in real world music. *Organised Sound* 7, 1 (2002), 41–43.

[13] Bernie Krause. 2008. Anatomy of the soundscape: evolving perspectives. *Journal of the Audio Engineering Society* 56, 1/2 (2008), 73–80.

[14] Cathy Lane. 2000. Space, Sound, and Music: Using Embodied Experiences of Space to Produce Multiple and Interconnecting Experiences of Space in Acousmatic Music. In *ACTES/PROCEEDINGS*. ISEA, 1–5.

[15] Andra McCartney. 2003. In and out of the sound studio. *Organised Sound* 8, 1 (2003), 89–96.

[16] Artemis Moroni, Jônatas Manzolli, Fernando Von Zuben, and Ricardo Gudwin. 2002. Vox Populi: Evolutionary computation for music evolution. In *Creative evolutionary systems*. Elsevier, 205–221.

[17] Olivier Pauplin, Praminda Caleb-Solly, and Jim Smith. 2010. User-centric image segmentation using an interactive parameter adaptation tool. *Pattern Recognition* 43, 2 (2010), 519–529.

[18] Bryan C. Pijanowski, Luis J. Villanueva-Rivera, Sarah L. Dumyahn, Almo Farina, Bernie L. Krause, Brian M. Napoletano, Stuart H. Gage, and Pieretti Nadia. 2011. Soundscape Ecology: The Science of Sound in the Landscape. *BioScience* 61, 3 (2011), 203–216.

[19] Aurora Ramirez, Jose Raul Romero, and Christopher Simons. 2018. A systematic review of interaction in search-based software engineering. *IEEE Transactions on Software Engineering* (2018).

[20] R Murray Schafer. 1994. *The soundscape: Our sonic environment and the tuning of the world*. Simon and Schuster.

[21] Marco Scirea, Peter Eklund, Julian Togelius, and Sebastian Risi. 2017. Can you feel it?: evaluation of affective expression in music generated by MetaCompose. In *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, 211–218.

[22] Denis Smalley. 2007. Space-form and the acousmatic image. *Organised sound* 12, 1 (2007), 35–58.

[23] Celestino Soddu. 2002. Recognizability of the idea: the evolutionary process of Argenia. In *Creative evolutionary systems*. Elsevier, 109–I.

[24] Feld Steven. 1996. Waterfalls of Song: An Acustemology of Place Resounding in Bosawi, Papua New Guinea. In *Senses of Place*. School of American Research, Santa Fe, 91–135.

[25] Emily Ann Thompson. 2002. *The soundscape of modernity: architectural acoustics and the culture of listening in America, 1900-1933*. MIT Press, Cambridge, Mass.

[26] Edmund R Thompson. 2007. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of cross-cultural psychology* 38, 2 (2007), 227–242.

[27] Barry Truax. 1999. *Handbook for Acoustic Ecology*. World Soundscape Project, Simon Fraser University and ARC Publications.

[28] Isabel Tweraser, Lauren E. Gillespie, and Jacob Schrum. 2018. Querying Across Time to Interactively Evolve Animations. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '18)*. ACM, New York, NY, USA, 213–220. https://doi.org/10.1145/3205455.3205460

[29] JI van Hemert and AE Eiben. 1999. Mondriaan art by evolution. In *Proceedings of the Eleventh Belgium/Netherlands Conference on Artificial Intelligence*. 291–292.