SYNAESTHETIC SANCTUARY

WILLIAM SHAW

2021

Synaesthetic Sanctuary

90 point thesis submitted to
Victoria University of Wellington
in fulfillment of the requirements
for the degree of
Master of Design Innovation

William Shaw

2021

◀ ***Figure 1.*** Title.

▶ ***Figure 2.*** Circular audio-visualisations.

ABSTRACT

Traditional scientific methods of visualising sound data have focused on techniques that attempt to capture distinct elements of the audio signal, such as volume and length. However, existing methods such as spectrograms and waveform analysis are limited in their expression of the characteristics associated with complex sounds such as bird song. This research explores strategies to visualise sound in an aesthetically engaging manner. It uses sound data from native New Zealand birds as a design tool for creating an audio-visual design system. The distinct focus on timing and pitch within these songs makes the data suitable for visual comparison. The design techniques explored throughout this research project attempt to express the unique characteristics of a variety of New Zealand bird songs and calls. It investigates how artistic audio-visual methods can be integrated with scientific techniques so that the auditory data can be made more accessible to non-specialists.

More specifically, this research aims to take advantage of the natural phonaesthetic connections people make between sonic and visual elements. The final output of this research consists of a generative design system that uses auditory data to create visualisations of New Zealand bird song. These visualisations have a mathematical basis, as well as being audio-visual artworks in themselves.

ACKNOWLEDGEMENTS

Thank you to Anne Niemetz for your support, guidance and inspiration throughout this research project. I am grateful to have had you as my supervisor, and have always been able to count on you for feedback and assistance.

Thank you to Tom White for your guidance, inspiration, and interest in my work. Your feedback has always been constructive and was instrumental in helping my research project reach its full potential.

Thank you to Zach Challies for your insightful feedback and always being able to help out with technical issues. I have learned a lot from working with you and have appreciated having someone with mutual interests to bounce ideas off.

Thank you to Dugal McKinnon for your inspirational lecture content and providing positivity over a difficult period.

Thank you to Steven Lam, Terry Moore, Byron Mallet and Michael McKinnon for providing critical technical support that made the completion of this project possible.

Finally, thank you to my family for always being there for me. Thank you to Maddi for helping me to remain motivated, inspired and giving me the strength to get through the last year and a half.

CONTENTS

LIST OF FIGURES

INTRODUCTION

The increasing popularity of films, games, websites, and other multimedia formats has brought with it the ongoing development of media technology. As a consequence of this, visual representations of sound known as audio-visualisations are now far more commonplace. While traditional methods of visualising sound data have focused on a scientific methodology that attempts to capture distinct elements of the audio signal, such as volume and length, recent innovations in media technology offer the possibility for designers to capture a more vivid, dynamic representation of sound.

There has been a significant amount of research exploring how sound can be represented artistically. This has usually been focused on an artist's interpretation of how the sound should be visually presented. However, recent innovations in software have made generative tools available to designers which allow for the development of audio-visual design systems capable of generating visual elements from audio files. This allows for the visual comparison of sound data, while also giving designers the flexibility to change the way the data is represented using sound as a design tool.

Technological advancements in soundscape analysis, such as automated recording devices, have produced an increasing amount of sound data which can be used for generating audio-visualisations. Although the quantity of sound data being compiled has continued to increase, methods for visualising these sounds have remained stagnant, limiting the potential for visual engagement with this data.

The native birdlife of New Zealand has a wide variety of distinct songs and calls. This provides a rich archive of sound data to explore using innovative audio-visual techniques. The distinct focus on timing and pitch within these songs means that the data can be made suitable for visual comparison through the manipulation of generative properties.

This research project aims to explore the connections between science and art using audio data as a tool to generate expressive visual elements. Historical artistic audio-visual will be explored alongside more recent technical audio-visual studies to contextualise the study. The final output consists of an audio-visual design system that can generate visualisations of New Zealand bird song using the raw sound data as a generative design tool.

RESEARCH PORTFOLIO
CHAPTER OUTLINE

CHAPTER ONE
BACKGROUND RESEARCH

Provides an overview and analysis of five key themes: Data visualisation, sound visualisation, audio-visual design, soundscape ecology, and New Zealand bird song.

CHAPTER TWO
METHODOLOGY

Outlines the aims and objectives of the project to establish a clear direction for the research, and justifies the research approach which has been undertaken to meet them.

CHAPTER THREE
DESIGN EXPERIMENTATION

Explores a range of design experiments to determine the most suitable software and approach for developing an audio-visual design system.

CHAPTER FOUR
TOUCH DESIGNER

Involves the initial development of the audio-visual design system using Touch Designer.

CHAPTER FIVE
HOUDINI

Consists of further development of the audio-visual design system using Houdini. Concepts explored in the previous chapter are built upon and refined for the development of the final system.

CHAPTER SIX
FINAL DESIGN SYSTEM

Showcases the final audio-visual design system, providing an outline of how the system works and exhibits a range of audio-visualisations which have been generated using the sound data of New Zealand bird song as a design tool.

CHAPTER SEVEN
DISCUSSION & CONCLUSION

Provides a critical reflection of the final audio-visual system, and further showcases how it can be applied to compare the sound data of New Zealand bird song. Implications, limitations, and further opportunities for developing the techniques explored in this research are also discussed.

I

CHAPTER ONE

BACKGROUND RESEARCH

THEME ONE :
DATA VISUALISATION

Contemporary society is often described as the age of big data, and there are hundreds of data visualisation methods that can be organised into categories such as tables, charts, graphs, diagrams, and maps (Reas et al., 2010). The western concept of data visualisation can be traced back to the invention of the line graph, pie chart, and bar chart invented by William Playfair in 1769 (Kirk, 2016). As the amount of data available to us continues to grow due to rapid developments in technology, there has been innovation in the way we process data as well as how we visually display it. Basic visualisation methods such as line graph, pie chart, and bar charts remain the most commonly used methods today for representing simple data. However, the increasing accessibility of simple, easy-to-use software such as Microsoft Excel and Adobe Illustrator has somewhat automated the process of visualising data, making it possible for non-specialists to easily create and develop their own data visualisations. This has transformed what was once a specialist activity into a part of mass culture, which is made evident through the increasing amount of visual data which is contained and displayed on digital devices.

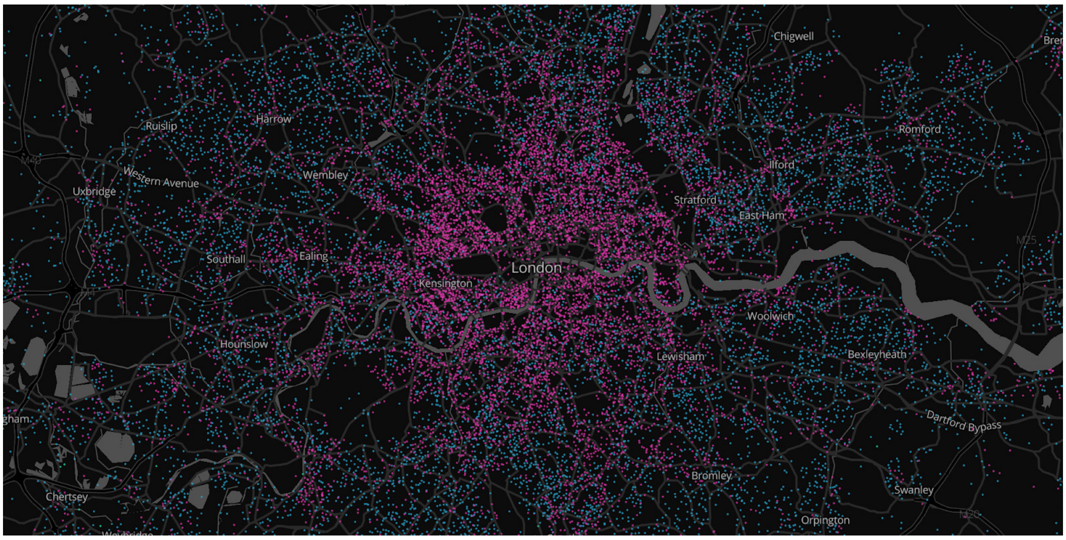
Reas et al. (2010) argue that when presented with data as an image, people understand and connect with it on a deeper level. They quote Card et al. (1999), who explain: “To understand something is called ‘seeing’ it. We try to make our ideas ‘clear’, to bring them into ‘focus’, to ‘arrange’ our thoughts (Card et al., 1999, p. 1)”. They explain how visual language can be used to construct meaning on a more intuitive level when compared to written communication. This is because the human brain has a fundamental understanding of visual communication from an early age, but years of education are required to develop the ability to read or write.

Reas et al. (2010) suggest that writing new software is one approach to move beyond common data representations and produce meaningful data visualisations. They argue that new techniques emerge as researchers and designers write software to fulfil their growing needs. Furthermore, Shaffer (2004) explains the theory of pedagogical praxis. This theory suggests that new technologies make it possible for people to participate in meaningful learning activities, acting as a bridge between professional practices and the needs of more novice learners (Shaffer, 2004). This has become increasingly apparent with the rapid increase in the use of Google Maps and similar data visualisation tools which can be used to explore the natural world. Barnett et al. (2011) further elaborate, explaining that the rapid increase of geographic information systems such as Google Earth and Google Maps have enabled specialists to overlay data and to evaluate the relationships between objects, locations, and other types of data, while also engaging the general public in performing simple geospatial analyses.

Obe (2013) explains that as the internet continues to develop, the increasing size of the average web page demonstrates the shift away from static, text-based content towards more interactive and immersive ways of visualising data. The Office for National Statistics in the United Kingdom has established a dedicated data visualisation centre, to explore methods of engaging people on the web through interactive, animated visualisations (Obe, 2013). In 2011, the centre developed a new approach to increase census return rates, adopting an animated approach for visualising census statistics over the last hundred years. The website was incredibly effective and generated over 100,000 more views than the plain text version, exemplifying the effectiveness of integrating more immersive data visualisation techniques to engage people.

Kirk (2016) argues that visual embellishments are an effective method to encourage viewers to engage with data visualisations through securing visual appeal and preserving communicated value. This is particularly useful when the aesthetic of the visual embellishments are compatible with the data being visualised, as it assists viewers in making visual connections between the visualisation and the data.

Refik Anadol is a media artist based in Los Angeles, who creates immersive installation artworks using a combination of machine learning and image data. He argues for the potential of media artists using data as a design tool to explore connections between art and science. American philosopher John Dewey's major writing on aesthetics *Art as Experience* explains that "science creates meanings; art expresses them" (Dewey, 1934). Anadol suggests designers take an interdisciplinary approach to explore connections between art and science through using data as a design tool, as this holds immense opportunity for future innovation and collaboration between disciplines (Anadol, 2020).



▲ **Figure 3.** *Dot density mapping of housing types in London based on census data.*

Note. From "Making census data open, accessible and useful for real people", by A. Barclay, 2020. (https://medium.com/nightingale/making-census-data-open-accessible-and-useful-for-real-people-f3ffb167913a). Image courtesy of Ahmad Barclay. Reprinted with permission.

THEME TWO :
SOUND VISUALISATION

Audio-visual representation techniques are becoming increasingly prevalent within society due to the growing popularity of films, television programmes, music videos, and computer games (Rogers, 2013). Many websites include a soundtrack, and the development of mobile phones and tablets allow users to engage with audio-visual activities quickly and easily. The increasing amount of audio-visual design has led to the development of a wide variety of methods and techniques which visualise sound, ranging from scientific interpretations to more artistic reflections.

Scientific Audio-Visual Methods

Sound is a form of energy that creates vibrations, impacting surrounding molecules and creating alternating bursts of high and low pressure (Carrier et al., 2012). Sound data is most commonly visualised using computer-generated images which create representations of what people hear, based upon technical parameters. There are a variety of visualisation techniques used to represent sound such as spectrograms, sonograms and other methods which attempt to capture these bursts of energy in a way that is targeted towards scientists. However, these visualisation strategies do not intuitively describe audio features in a way that could be understood by people who are not specialists in the field.

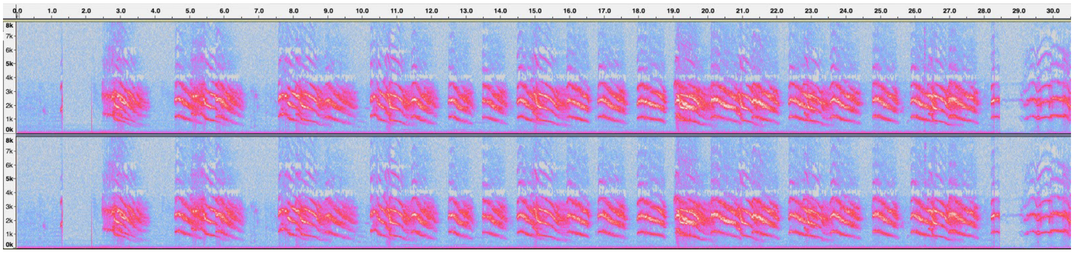
Sonogram/Spectrogram

Spectrograms represent sound through mapping amplitude and frequency on the y-axis, and time on the x-axis (Carrier et al., 2012). The lightness or darkness of the line represents the amplitude of a sound. A spectrogram is most commonly used to provide a visualisation of the audio spectrum which is perceptible to human ears, which is between 20 Hz and 20,000 Hz (Thibeault, 2011). However, this frequency range can be easily adjusted for the visualisation of more specific frequency ranges. Spectrograms can also be used to precisely detect the pitch and timbre of a sound through the analysis of the solid white lines, with lines closer to the bottom representing a lower pitch and vice versa.

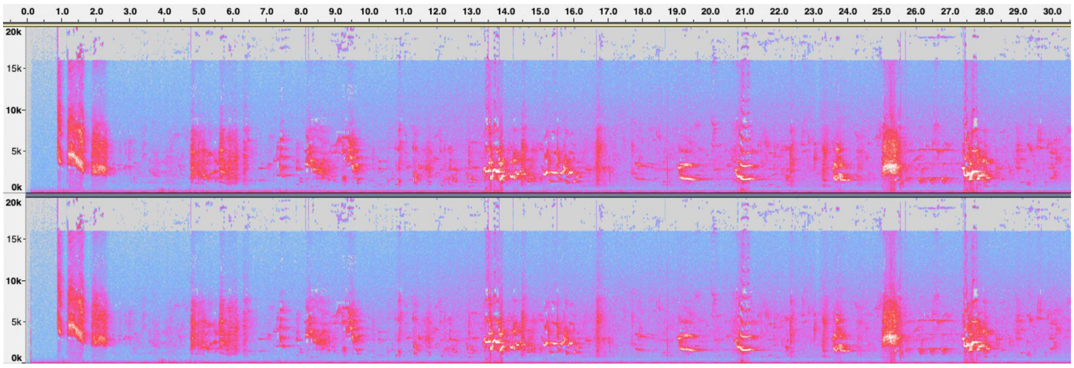
The ease of reading and writing a spectrogram, as well as the rise of sound as data in the age of new media has led to it becoming the most common method of sound visualisation, with software packages such as Audacity allowing people to generate spectrograms for free.

Fast-Fourier-Transform (FFT) Analyser

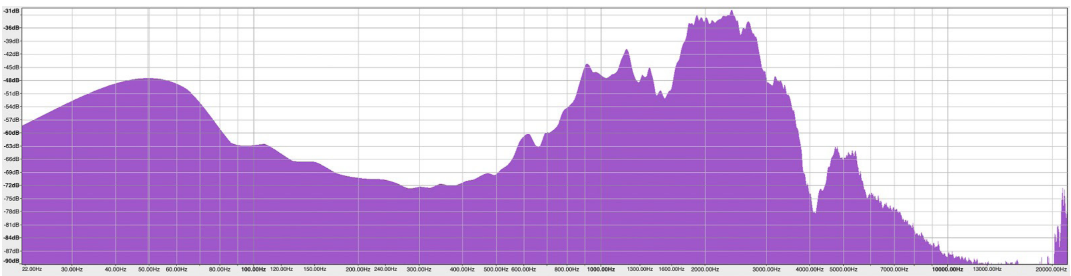
Another common technique for visualising audio is the use of a digital audio device known as a fast-fourier-transform (FFT) analyser (McCarthy, 2014). Audio is fed into the analyser where it is then separated by frequency and displayed through individual channels. The FFT analyser can visualise amplitude and phase over frequency in great detail, however, it has limited usability as it can be difficult to transcribe and understand for people who are not specialists in the field of sound analysis.



▲ **Figure 4.** *Kea song spectrogram displaying frequency range between 0 - 8,000 Hz .*



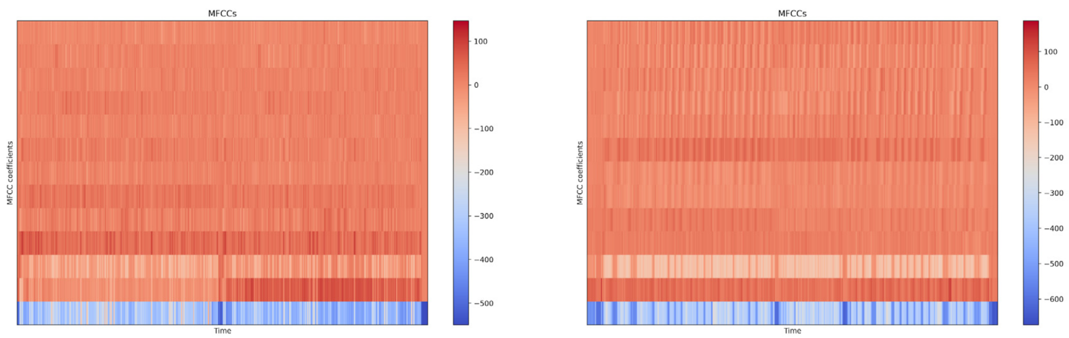
▲ **Figure 5.** *Kaka song spectrogram displaying frequency range between 0 - 20,000 Hz.*



▲ **Figure 6.** *FFT analysis of Kea song.*

Mel-Frequency Cepstral Coefficients (MFCC)

Pitch is one of the most identifiable elements of an audio signal and can be measured as the frequency of a signal. The mel-frequency cepstrum, also known as the mel scale (after the word melody), relates the perceived frequency of a tone to the actual measured frequency. Figure 7 provides an example of mel scale visualisations being applied to two different New Zealand bird songs. The mel scale captures the perceived differences between frequencies according to the human ear, which is much better at detecting small changes in pitch at low frequencies rather than high frequencies (Nair, 2018). For example, the perceived difference between 3,000 - 4,000 Hz would be much more significant in comparison to higher frequencies such as 7,000 - 8,000 Hz because of the way the human ear processes sound. The mel scale captures such differences by applying a logarithmic equation to frequencies measured in Hz. Although the mel scale measures sound to more closely resemble what is heard by the human ear, the visual representations generated by this system are very difficult to understand unless the user is a specialist in the field of soundscape analysis, and specifically trained in reading MFCC visualisations.



▲ **Figure 7.** *MFCC analysis of Kea and Kaka song.*

THEME THREE : AUDIO-VISUAL DESIGN

There have also been a large number of audio-visualisations that attempt to map sound as a form of artistic expression. Whereas scientific visualisation methods rely on computers to process sound and generate visual elements, artistic strategies have a greater reliance on humans to make audio-visual mapping decisions. In comparison to other human senses, the connection between auditory and visual perception is particularly pronounced. This emphasises the importance of designers to be able to optimise the correlation of auditory perception and visual elements within their work. To date, a systematic methodology for cross-sensory alignment between visual and auditory elements has not been developed. Within the fine arts, music, and literature, however, the search for connections has a long history (Rogers, 2013).

Synaesthesia/Phonaesthesia

The concept of sound having a shape, colour, or movement can be traced back to the perceptual phenomenon of synaesthesia. The term has been derived from the Greek words “syn” (together) and “aesthesis” (sensation). It refers to a specific manner of perception by which the stimulation of a certain sensory channel results in sensations being felt in one or more other sensory channels.

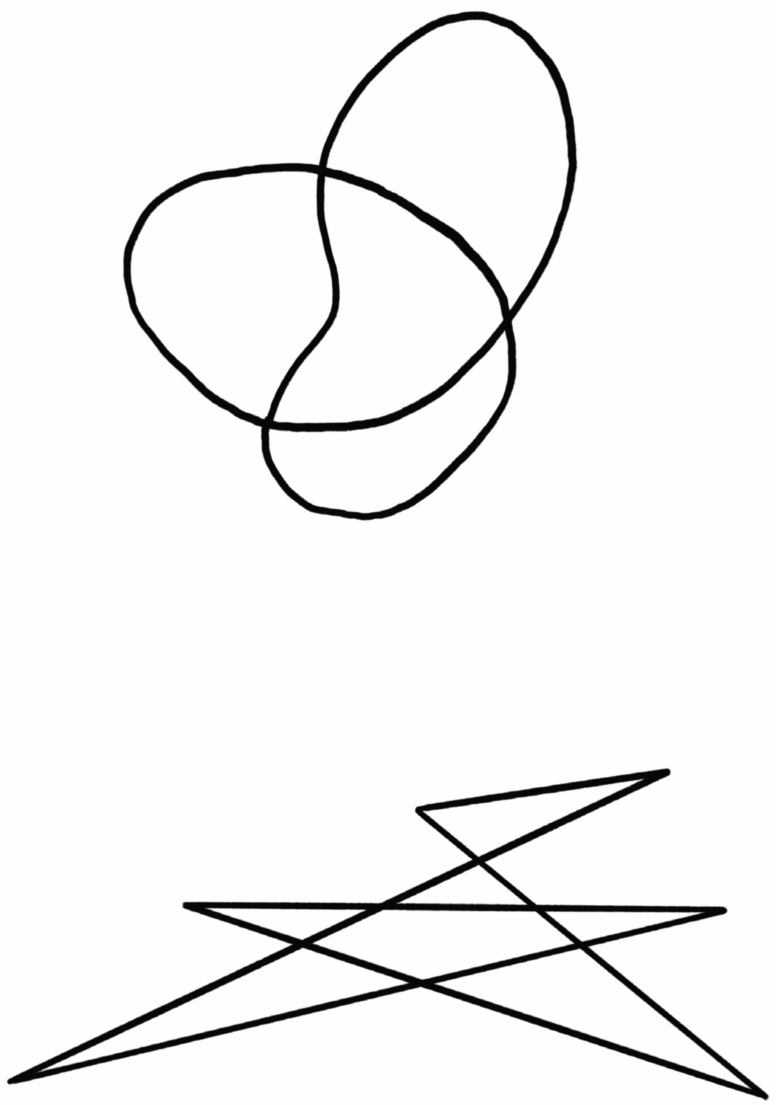
In this manner, the perception of sound can additionally stimulate a sensation of visual shape and color. Synaesthesia is a rare phenomenon found amongst few people, however, it is an intensely investigated matter within the field of synaesthetic research (Rogers, 2013). Those who experience synaesthesia are known as synaesthetes and can experience a variety of audio-visual connections such as the association of sound with perceptual elements such as taste, smell, and colour (Ward, 2013).

Although those with synaesthesia make up a small amount of the population, there is also a concept known as phonaesthesia which refers to the distinct audio-visual associations which exist within the human language and can be easily perceived by those without the condition. For example, the letter "p" can be seen to represent the sudden start of a sound through the words pop, puff, and plop. This phenomenon also can be applied to descriptions of objects which do not have a distinct sound, such as the rhyme "-ump" being applied to describe mass or swelling such as rump, stump, lump, and bump (Blake, 2016).

Maluma/Takete

Numerous studies of humans sharing systematic preferences for pairing audio with visuals have shown a consistent association of higher-pitched sounds with lighter colours, smaller sizes, spikier shapes as well as higher locations in space (Ludwig et al., 2011). These intuitive audio-visual associations can also be found within language, with the German phrase *dunkler Ton* (dark sound) referring to a low pitch sound, and the Japanese phrase *kiiroi koe* (yellow voice) referring to a high pitched sound.

In 1929, German psychologist Wolfgang Köhler asked their students to match a novel word form (maluma or takete) with an abstract line drawing (one angular, one curved) (Styles & Gawne, 2017). He reported that the majority of the students associated the angular drawing with the word takete, and the curved drawing with maluma. Since this initial study, various researchers studied this phenomenon in a variety of contexts. The most commonly tested pseudowords are the original maluma/takete combination, as well as bouba/kiki, developed by Ramachandran and Hubbard (2001).



▲ **Figure 8.** *Reproduction of the sketch used for the maluma/takete study.*

Note. Adapted from *Gestalt Psychology* by W. Köhler, 1929, New York, NY, United States: Liveright. In the public domain.

I

To understand why the maluma/takete effect occurs, Styles & Gawne (2017) explain that it is important to examine the prevalence rates for different consonants and vowels across different languages. In a study based upon the PHOIBLE Online database of 2,160 segments from 1,672 documented languages (Moran et al., 2014). The most prevalent sounds occurring across a variety of languages are the peripheral vowels of /a/, /i/, and /u/, followed by the mid-high vowels of /e/ and /o/. The most notable feature among these sounds is that they all occur within English and a majority of European languages, as well as being the sounds most commonly investigated within maluma/takete tests. A likely explanation for why such prevalent sounds are used in creating these test pseudowords is that there is a universal perception of sound symbolism within the human language. Sounds made from the front of the mouth form significantly different acoustic profiles versus sounds made from the back of the mouth (e.g. /k/ vs /m/ respectively). Almost all languages have this perceptual symbolism present within their phonemic inventories, making it a universal mapping point when it comes to visualising sound through size, brightness, and colour (Styles & Gawne, 2017). However, though these phonemic associations occur in almost all languages there is no specific sound that exists in all languages. For example, bouba, kiki and takete would all form legitimate words in Japanese but maluma would fail due to there being no Japanese equivalent of /l/.

Interestingly, intuitive audio-visual associations have also been found to be present within non-human species, with a study of chimpanzees indicating that they also spontaneously associate higher-pitched sounds with lighter colours, and lower-pitched sounds with darker colours (Ludwig et al., 2011). Researchers believe that this underlying audio-visual similarity between humans and non-humans can be attributed to a structural similarity between multi-sensory processing systems, either from evolutionary heritage or shared sensory experience (Styles & Gawne, 2017).

Artistic Audio-Visual Methods

In the 1920s, Russian painter Wassily Kandinsky began to experiment with using geometry within their artwork to describe relationships amongst colour, sound, shapes, and music. Kandinsky based their colour theory on the theosophical theory of electrical vibrations emitted from particular thoughts or feelings, vibrations that could be experienced through colour and shape. According to the artist, for instance, the colour yellow has an aggressive character and "may be paralleled in human nature with madness" (Rogers, 2013).

▲ **Figure 9.** *Succession - Vasily Kandinsky.*

Note. From "Succession" By V. Kandinsky, 1935. (https://www.wikiart.org/en/wassily-kandinsky/succession-1935). Image courtesy of Wikiart. In the public domain.

As audio-visual technology continued to develop, soundtracks became more prevalent in film and video work. This led to further development of the synaesthetic relationship between moving images and music. In the 1930's, German-American artist Oskar Fischinger attempted to create what he described as "visual music", matching visual and auditory elements from both a technical and synaesthetic standpoint (Arfini, 2013). The two-and-half-minute film *Studie Nr. 7* is structured to correspond with the musical piece *Ungarische Tänze no. 5*, composed by Brahms. This technique of audio-visual synchronisation was later introduced to the mainstream in 1940 when Fischinger worked with Bach's *Toccata and Fugue in D Minor*, for a sequence of visual music in Disney's *Fantasia* (Rogers, 2013).

In the statement *Sounding Ornaments*, Fischinger explained a method for encoding graphics onto film strips. Once encoded, the film strips were then able to broadcast tones when played through a machine. While completing these experiments, Fischinger observed the following:

> In reference to the general physical properties of drawn sounds, we can note that flat and shallow figures produce soft or distant-sounding tones, while moderate triangulation gives an ordinary volume, and sharply-pointed shapes with deep troughs create the loudest volume. Shades of grey can also play a significant role in drawn music-ornaments (Fischinger, 1932, p. 1).

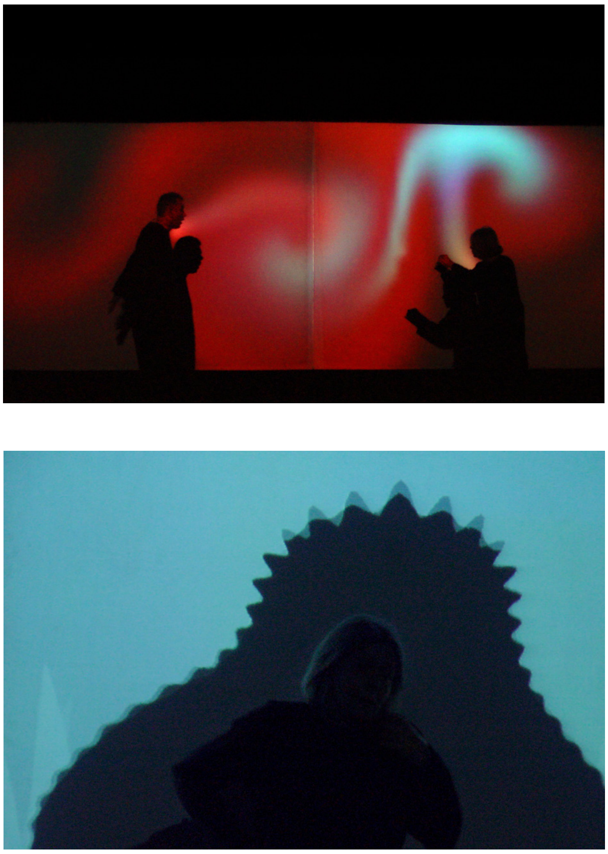
As opposed to the technique used by artists such as Kandinsky, these findings were physically determined rather than subjectively produced (Rogers, 2013).

Arfini (2013) explains that from a more formal standpoint, moving images can be incorporated as a kind of visual translation of musical thematic progress. One of the strong correlations that have arisen from attempts to create systematic audio-visual systems has been the association of pitch and colour. Pitch-to-colour connections have been important to many cultures, and early concepts of colour association can be found from Indian philosophy to Isaac Newton's *Opticks*, written in 1704 (Rogers, 2013).

Newton's theory was that the spectrum of sound is proportional, and could be tied to the principal degrees of colour - red, orange, yellow, green, blue, indigo, and violet (Silva, 2017). The addition of indigo to the sequence of hues now commonly known as ROYGBIV (red, orange, yellow, green, blue, indigo, and violet) is the result of Newton adding a colour to their prism experiments so there would be seven colours to match the seven notes of the musical scale. Although Newton's theory was not largely accepted by philosophers of the time, it has continued to inspire a huge amount of audio-visual work exploring the perceptual and aesthetic possibilities of pitch-to-colour associations.

▲ **Figure 10.** *Isaac Newton's colour wheel.*

Media artist Golan Levin has been a key part of the recent vanguard of artists and designers exploring audio-visualisations through digital media. Levin's work explores the concept of phonaesthesia, exploring the mapping of audio to visual features such as hardness, sharpness, brightness, and darkness (Levin, 2009). The project *Messa di Voce* by Golan Levin and Zachary Lieberman features two professional vocalists, Jaap Blonk and Joan La Barbara, whose vocal performance is transformed into abstract forms and textures which express the sound on a phonaesthetic level (Perron, 2004). No actual words are present within the performance, with it focusing on visualising different aspects of sonic expression such as screams, yells, and cries. Projects such as *Messa di Voce* exemplify how the concepts found within phonaesthesia can be used to inform audio-visual mapping decisions.



▲ **Figure 11.** *Messa Di Voce performance at Ultrasound 2003.*

Note. From "Messa Di Voce" by G. Levin, Z. Lieberman, J. Blonk and J. L. Barbara, 2003. (https://www.flickr.com/photos/tmema/sets/72157594393980031/). Image courtesy of Golan Levin, Zachary Lieberman, Jaap Blonk and Joan La Barbara. Reprinted with permission.

▼ **Figure 12.** *Inmi Lee - Mother.*

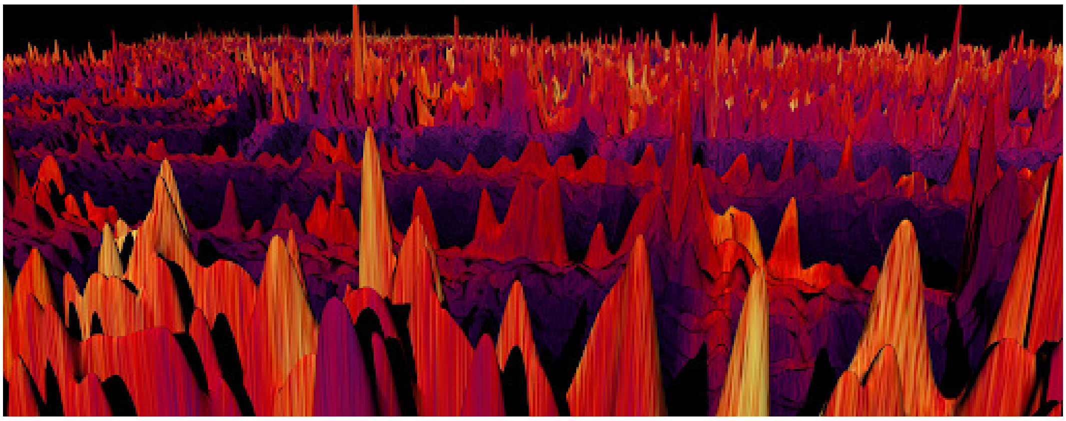
Note. From "Mother", by I. Lee and K. McDonald, 2013. *Leonardo, 47*(4), 410-411. (https://www.muse.jhu.edu/article/552313). Copyright 2013 by Inmi Lee and Kyle McDonald. Reprinted with Permission.

More recently, in 2013, artist Inmi Lee created a series of generative forms for the project *Mother*, which explored the synaesthetic connections between language and form (Lee & Mcdonald, 2014). To create the forms, participants in the project were asked to make hand movements to describe a range of different sounds. The hand movements have been captured using Xbox Kinect, and then 3D printed into physical objects. This project demonstrates how audio can be translated into tangible 3D forms which visualise and express identifiable characteristics of the sound.

Alican Inal, an architect, urban designer and data artist based in London, argues that to fully understand the benefits of sound as a design tool, we need to be able to visualise it. Using spectrogram data as a visual mapping tool, their work focuses on creating complex, site-specific 3D models of sonic environments (Inal, 2020).

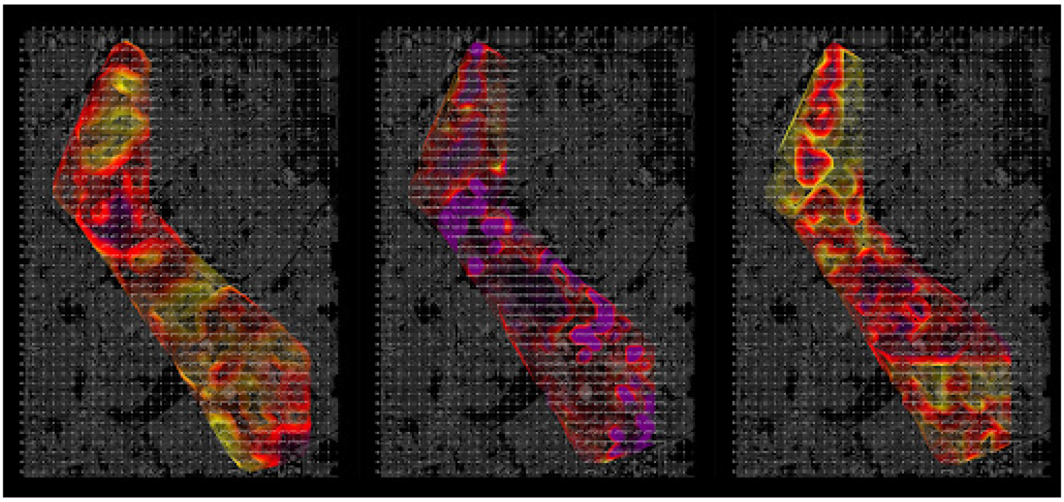
Inal's work also explores applying this technique to architectural space, showing how environmental noise hotspots can be identified through the study of visualised spectrogram data. Through identifiable visual properties such as sharpness and colour intensity, parts of the building which may be unpleasant for the human ear can easily be identified and improvements can be made to the building's spatial design. This exemplifies how the study of sound data as a design tool can help to improve the way things are created (Inal, 2020).

I



▲ **Figure 13.** *Alican Inal - 3D spectrograms.*

Note. From "Alican Inal - Work" by A. Inal, 2021. (https://alicaninal.com). Image courtesy of Alican Inal. Reprinted with permission.



▲ **Figure 14.** *Alican Inal - soundscape studies of the Queen Elizabeth Olympic Park, London.*

Note. From "Sonic Architecture: Exploring the Relationship Between Spaces and Sound" by A. Inal, 2020. (https://populous.com/sonic-architecture-exploring-the-relationship-between-spaces-and-sound). Image courtesy of Alican Inal. Reprinted with permission.

THEME FOUR :
SOUNDSCAPE ECOLOGY

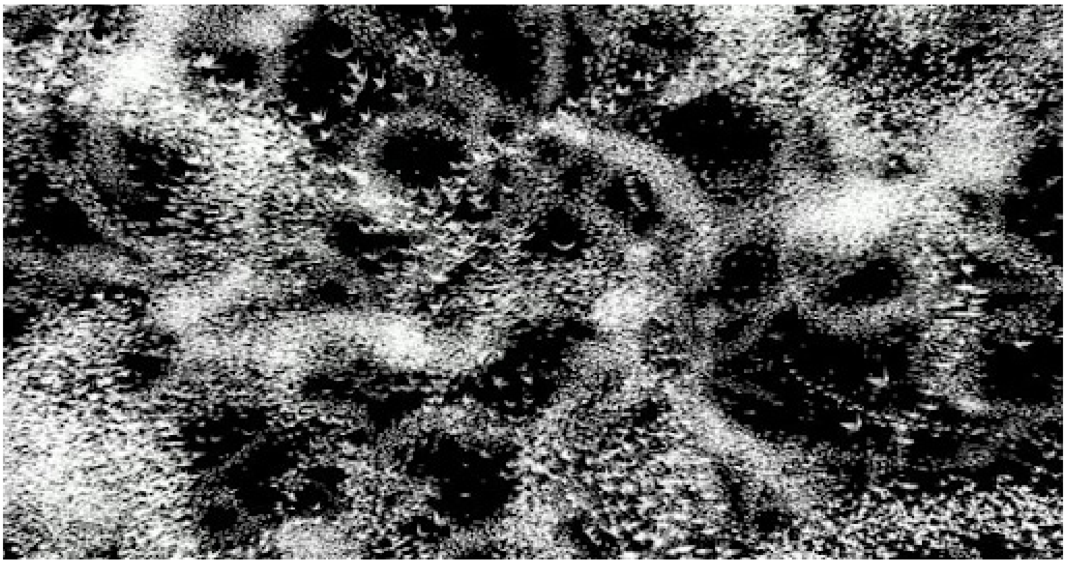
Milone and Carmada (2017) argue that the sounds of nature are closely connected to the quality of the environment around them, and are dramatically influenced by human activity. They explain that recent technological innovations such as automated recording devices, low cost storage, developments in the processing of acoustic data, as well as new methodologies in the field of ecology (a branch of biology that studies the relationship between organisms and their environment) has given a significant drive to research within this field. The natural environment analysis methodology is based on the principles of soundscape ecology, which uses specific interfaces known as sonometers which allow for the qualitative analysis of photometric and spectrographic recordings (Milone & Carmada, 2017).

Maher (2011) completed a research project investigating long-term soundscape audio recordings. The project explored spectrograms, audio snapshots, as well as time-lapses of sound information. As they explain in their study, their goal was to define and capture environmental sound texture and present it to the public. One of the main challenges which arose was: how can the element of sound be presented to the viewer in an engaging way? As acknowledged by Maher (2011), visitors to a website or visitor centre spend an average of only a few minutes there, which highlights the need for engaging visual design within this field.

I

Designers around the world have been exploring ways of visualising environmental data so that viewers reflect on the relationship between natural and human-made systems. Victoria Vesna, an artist based in the United States, has been motivated by artist-scientist collaborations that explore areas of environmental science while raising awareness of that research to broader audiences through visual design (Vesna et al., 2019). In the project *Noise Aquarium*, Vesna projects 3D scans of plankton into an interactive space using scientific imaging techniques, enlarging the micro-creatures so they appear like they are the size of whales. The presence of the audience activates harsh visual and audio effects, resembling noise pollution within the ocean which is caused by humans using tracking technologies such as sonar. The purpose of this project, Vesna explains, is to demonstrate the implications of human impact on the environment (Vesna et al., 2019).

In another of Vesna's projects, *Bird Song Diamond*, audiences are encouraged to mimic different bird songs in groups, experimenting with male and female vocalisations. The sounds and movement of the audience influence the sound and images being projected around the room in real-time (Ars Electronica, 2018). Vesna's work provides an example of how environmental data can be integrated into design projects to elicit an emotional response from people.



▲ **Figure 15.** *Victoria Vesna - Bird Song Diamond.*

Note. From "Bird Song Diamond" by V. Vesna, 2018. (https://birdsongdiamond.com). Image courtesy of Victoria Vesna. Reprinted with permission.

THEME FIVE :
NEW ZEALAND BIRD SONG

The structure of bird song often consists of a series of monotonous patterns containing only a few simple phrases. However, some birds have extremely complex songs with a large phrase repertoire which are organised with a set of perceptible patterns (Sasahara et al., 2012). Bird song structure also usually contains rapid temporal modulations, which is useful for identifying species-specific characteristics (Stowell & Plumbley, 2014).

Morton (1975) studied the structure of bird song in relation to their environment and found that forest birds use lower frequency notes than birds living in more open habitats. This was confirmed by subsequent studies that compared the songs of birds living in open and closed habitats (Badyaev & Leaf 1997; Bertelli & Tubaro 2002; Tubaro & Lijtmaer 2006). Ambient noise within the environment can also be a powerful selective force, causing changes within bird song frequencies among species, as well as populations (Slabbekoorn & Peet 2003; Brumm & Slater 2006).

Though there have been numerous studies of bird song that have measured repertoire size and approached song structure using simple statistics and models, the structural properties of complex songs remain largely unexplored (Sasahara et al., 2012). Azar and Bell (2016) found that there has been only a small amount of research into the vocalisations of bird communities in the temperate broadleaved rainforests of the southern hemisphere. New Zealand's distinctive bird community and temperate broad-leaved rainforests (Ovington & Pryor, 1983) allow for the study of different structures of bird song within a mixed community of native and exotic birds (Azar & Bell, 2016).

I

There is also further scope for examining changes of vocalisations of introduced species since their arrival in New Zealand, and for determining why their vocalisations have stayed relatively the same compared to native species (Azar & Bell, 2016). One emerging property from this field is that certain phrases are often constructed from sequences that are neither uniform nor random, but possess some discernible pattern (Sasahara et al., 2012).

A bird song is defined as a noise that serves a specific function, which is almost always related to mating. In contrast, a bird call can serve a wide variety of functions such as “follow me” or “danger is near” (Truslow, 2014). While bird songs tend to follow a definite rhythm and structure, calls are usually short notes or phrases used to convey everyday concepts. Many bird sounds which may not sound musical to the human ear are technically considered bird songs by this definition.

Within the more specific area of bird song, there has been some exploration of how sound analysis techniques can be applied as an effective research method. Farina et al. (2011) completed a study into the soundscape methodology for long-term bird monitoring. They explain that the complexity and increasing fragility of the link between natural and human-made systems will require new types of investigation to prepare for potential environmental surprises in the future.

Farina et al. (2011) further expand on this concept, explaining that certain aspects of the environment such as the disruption of natural communication systems and the constant exposure to human noise activity are still poorly understood within modern society, and alternative research techniques into these fields have the potential to uncover new knowledge. Their study was within the scientific field, but exploring these fields in an innovative, more artistic way could help encourage the public to engage with these topics. Sasahara et al. (2012) explain that further research into the area of bird song and the development of a rich set of bird song patterns has the potential to further human understanding of the evolution of vocal communication in birds, of animal communication systems, as well as the broader areas of syntax and semantics.

Wang et al. (2017) argue for the advantages of representing bird song within a 3D space, as opposed to the 2D representations commonly found among bird song visualisations today. They explain that 3D spectrograms not only improve the readability of bird song, but also allow for a more accurate representation of the energy change curve which represents bird tweet signals, as well as clearer monitoring of any changes in speech signal energy over time (Wang et al., 2017).

SUMMARY

As identified by both Azar and Bell (2016) and Sasahara et al. (2012) the topic area of bird song analysis within a country like New Zealand provides large scope for research and the development of new knowledge. There has only been a modest amount of research into the structural properties of complex bird song and the vocalisations of bird communities within temperate environments such as New Zealand. Soundscape ecology, as explained by Milone and Carmada (2017), provides a methodology for analysing environmental noise information such as the structure of bird song. There is, however, a lack of available audio-visual techniques able to emphasise research findings and provide non-specialists with a meaningful way to engage with this information, as acknowledged by Maher (2011) in their research.

Anadol (2020) argues for the potential of media arts to bridge the gap between science and art, and that exploration of these interdisciplinary connections will encourage innovations, perspectives, and greater collaboration between disciplines. Although current methods of visualising environmental sounds such as bird song are suited for the optimised reading of information by sound analysts and computers, the potential for generating visuals from this sound information which capture the vivid, dynamic elements of bird song remains largely unexplored. Consequently, there is an opportunity for creating an audio-visual system that uses the sound data of New Zealand bird song as a design tool for generating audio-visualisations. This prompted the research question:

How can sound data be used as a design tool to create an audio-visual design system capable of comparing New Zealand bird songs through an artistic format?

II

CHAPTER TWO

METHODOLOGY

OVERVIEW

To answer the research question, it was important to identify an appropriate approach. Frankel and Racine (2010) provide a framework for understanding approaches to design research, building upon previous research in the field. Research through design is an approach to conducting research where design methods and processes are used to generate new knowledge (Zimmerman & Forlizzi, 2014). It employs reflective design strategies to find innovative solutions to problems throughout the making and critiquing process. Research through design suggests researchers take a holistic approach to problem solving, and integrate knowledge and theories from a wide variety of disciplines (Zimmerman & Forlizzi, 2014). It is also known as practice based research, and topics can range from an idea or concept to a new material or process (Milton & Rodgers, 2013). Following the recommendation of Ramírez (2017), a criteria-based approach was used to analyse various stages of the process and measure the effectiveness of the final design.

Aim 1 :

- Investigate how sound data can be visually represented using audio-visual design methods.

Objectives :

- Research current methods of visualising sound data through digital design.
- Trial various digital design methods for the visualisation of New Zealand bird song.

AIM 1 METHODS :

Literature Review

To contextualise the research project, a literature review was conducted to explore the areas of data visualisation, sound visualisation, audio-visual design, soundscape ecology and bird song analysis. A literature review involves the analysis of the current field of knowledge relevant to the research project, to inform the researcher's perspective on the subject and identify where gaps in knowledge may exist (Miltons & Rodgers, 2013).

Design Experimentation/Iteration

After situating the research through conducting a literature review, the next phase of the research consisted of design experimentation using a variety of software. Ramírez (2017) explains that the design experimentation phase can involve a variety of mediums, and be based upon a loosely defined idea. This process informed which software would be the most effective for visualising audio data, with a specific focus on using New Zealand bird song as an input. Following the recommendation of Ramírez (2017), these experiments also helped inform the development of the research criteria.

Aim 2 :

- Develop an audio-visual design system for New Zealand bird song.

Objectives :

- Create an audio-visual design system that is able to generate artistic and expressive audio-visualisations of New Zealand bird song through the perceptual mapping of sound data.

- Develop a series of videos and images which showcase the audio-visual design system being applied to a range of New Zealand bird song, expressing unique sonic elements of each sound.

AIM 2 METHODS :

Generative Design :

Generative design is a design method where the generation of form is based on mathematical rules or algorithms (Gross et al., 2018). In recent years, the development of computer software has introduced generative design to the mainstream, offering an innovative alternative to form making. Generative design is often instead described as form finding, as the process consists of modifying parameters and systems which instruct the computer to generate visual elements (Agkathidis, 2015). It is based on the iterative development of design systems, where the most visually compelling results are selected and refined. The generative design process puts the designer in the role of the conductor, orchestrating the decision making process of the computer rather than making direct aesthetic decisions.

Generative design was a key design method used for this research as it allowed for the iteration and refinement of the audio-visual design system through incremental adjustments to mapping parameters.

Documentation :

Through consistent design documentation, the design process could be analysed and different design software and strategies could be compared. Screenshots, annotations, renders, and short animations were recorded and then used to reflect upon various stages of the design process. Hanington and Martin (2012) explain that visual documentation is an essential part of the research process in order to appropriately analyse the positive and negative qualities of the design at various stages.

Critical Reflection :

Critical reflection describes the process of stepping back from the design process to focus on what may or may not be successful about the design, analyse how to improve it, and figuring out how to prevent similar mistakes from occurring in future design iterations (Bassot, 2015). Kolb's Learning Cycle was developed in 1984, and it has since become seminal within this area of research. It breaks down the process of critical reflection into four key stages: concrete experience, reflective observation, abstract conceptualisation and active presentation. Kolb (1984) argues that the process is not necessarily linear and can begin or end at any point.

Critical reflection was used to iteratively develop the final audio-visual design system and encouraged the iteration of generative mapping decisions throughout the research process.

Criteria Based Approach :

The criteria-based approach served as a guide throughout the design process and was used to analyse the success of each phase of the design development. This ensured the research process finished with a systematic assessment of the final design, and the criteria could then be used to describe the design and its contribution to the research field (Ramírez, 2017). Ramírez (2017) argues that the criteria used to measure the success of design solutions should continue to evolve throughout the design process. This recommendation was followed, with the criteria being developed alongside the research project.

CRITERIA

The following criteria was developed throughout the various phases of the research project. It was then identified as being the foundation for measuring the success of the final audio-visual design system.

- The system should generate 3D visualisations so that the resulting form can be examined from multiple angles.

- The system should be able to contrast and compare different New Zealand bird songs and generate identifiable visual differences or similarities, which have a direct relation to the difference in sound.

- Sound data of New Zealand bird song should be able to be used as an input to generate comparable forms using the final system.

- The audio-visual mapping choices should consider human's phonaesthetic understanding of audio-visual connections, such as the maluma/takete phenomenon.

- The system should be able to identify changes in pitch and have enough visual variance to represent the frequency range of New Zealand bird song.

- The movement of the visual elements should have a direct relation to the sound.

- Visual embellishments of the data should consider the appearance of New Zealand birds to allow people to draw distinct associations to the data being visualised.

- The audio-visual mapping decisions should integrate the full colour spectrum to create as much variance between different frequencies as possible.

III

CHAPTER THREE

DESIGN EXPERIMENTATION

OVERVIEW

The first stage of development consisted of design experimentation to determine what would be the most suitable approach for generating audio-visualisations of New Zealand bird song.

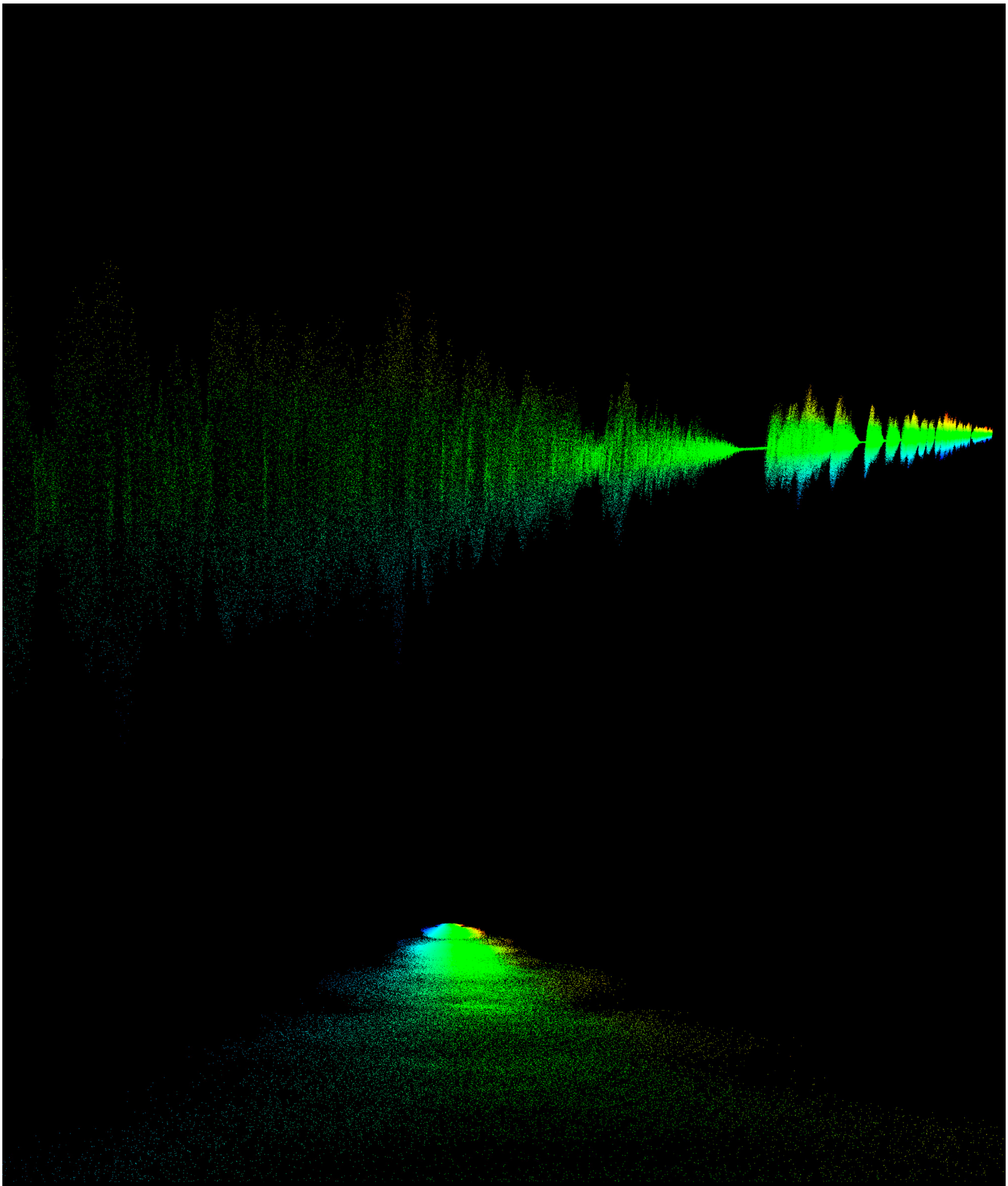
- ***Point Clouds***
- ***Touch Designer***
- ***Houdini***
- ***Processing***

EXAMINING SOUNDWAVES AS POINT CLOUDS

One of the earliest experiments within the design process was to explore what sound data looked like when converted into a point cloud. FME, developed by Safe Software, is a data integration platform that allows for the simple transformation of data from numerical data into visual elements (Safe Software, n.d.). Audio files consist of binary information which can be visualised as points, with the x-value representing time and the y-value representing amplitude.

The first stage of converting a sound into a point cloud to convert the audio file into numerical information. Audio files were converted into .DAT files that list the time and amplitude values in a chart. The .DAT files could then be opened using FME, where they could be converted into point cloud visualisations.

Examining the data in a 3D space allowed for specific areas of the sound to be closely inspected from multiple angles. After further research, it was found that a more suitable approach for generating 3D visualisations would be to extract numerical data from the audio using generative 3D software such as Touch Designer and Houdini. The data could then be used to directly manipulate visual elements through parametric mapping.



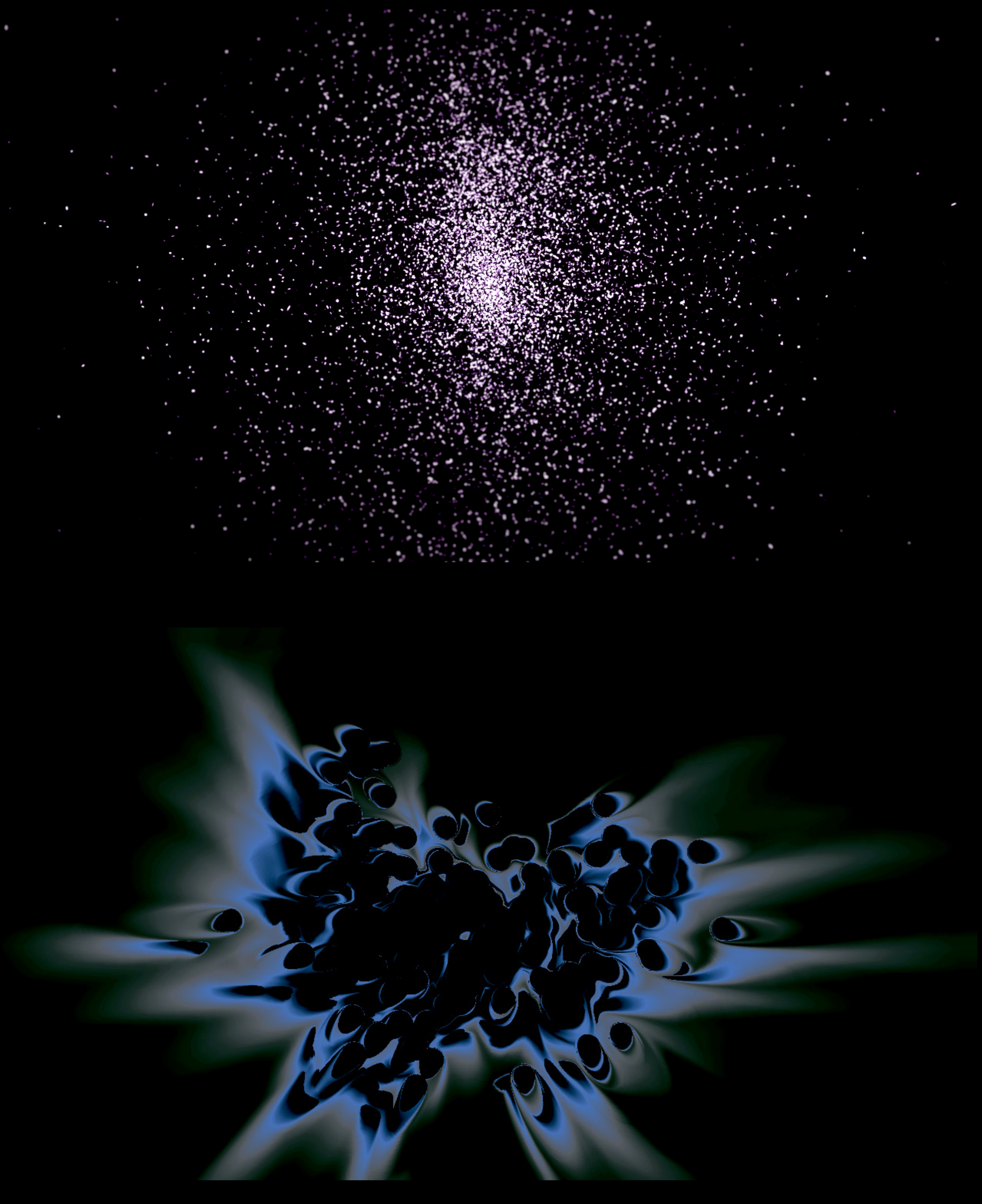
▲ **Figure 16.** *Audio file after being converted into a point cloud.*

INITIAL EXPERIMENTS USING TOUCH DESIGNER

Following these experiments, the next phase of the design process was to explore the possibilities of using Touch Designer. Touch Designer is a visual programming language built for creating real-time interactive multimedia content (Derivative, n.d.). It was developed by the Toronto-based company Derivative and is used by artists, programmers, coders, and performers to create real-time, interactive multimedia content.

Initial Touch Designer experiments explored the potential of mapping audio parameters such as volume and frequency to animate visual elements. While some interesting visual results were generated from these experiments, more development was needed to create a clear correlation between the visual and auditory elements. However, from these experiments, it became clear that 2D visualisation methods are considerably less effective for the expression of sound as they have considerably fewer parameters to manipulate due to the movement and form of the visual elements being restricted to a 2D axis. Earlier background research into concepts such as maluma/takete established that the manipulation of form was one of the most effective ways to express sound due to the intuitive perceptions that exist in society surrounding sound and shape.

Touch Designer was also used to explore the manipulation of particles within a 3D space, mapping audio parameters to drive the movement of the particles. These experiments lacked the expressiveness and legibility required to create a successful audio-visual design system using New Zealand bird song. The particles would move in relation to the sound, but it was difficult to control their movement in a way that would reflect sound based on human's phonaesthetic understanding of sound.



▲ **Figure 17.** *Touch Designer particle experiments.*

INITIAL EXPERIMENTS USING HOUDINI

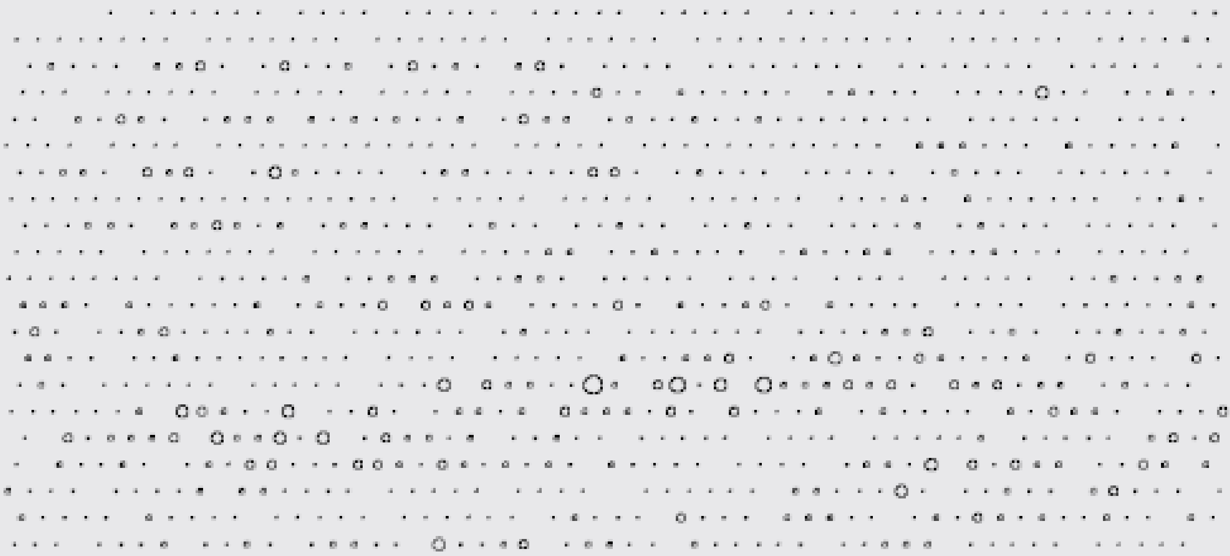
Houdini, a 3D procedural software developed by SideFX (SideFX, n.d.), was also used to carry out initial audio-visual design experiments. To explore the possibilities of 3D audio-visualisations, experiments consisted of generating a 3D mesh and manipulating the growth of the form through audio properties. Within Houdini, this can be achieved through the use of a channel operator, also known as a CHOP. Using bird song sound data as an input, numerical values extracted from the sound file were re-mapped and manipulated in order to generate a 3D mesh. However, more audio-visual mapping development was required in order to be able to control visual elements to reflect the sound.

It was discovered that through linking the audio to a noise manipulator, the form could be controlled so that it became increasingly sharp and jagged. This was useful to note for future experiments as this audio-visual mapping technique had a distinct link to the maluma/takete phenomenon studied by Köhler in 1929, but applied it to a digital 3D space.

▲ **Figure 18.** *Generating a 3D mesh from an audio file using Houdini.*

INITIAL EXPERIMENTS USING PROCESSING

The software and programming language Processing was used to explore another potential way of generating forms based upon parameters extracted from New Zealand bird song. The minim library is a java audio library that can be used to extract features from audio files using Processing (Processing, n.d.). Using the minim library, audio files could be input into Processing for both manipulating and generating visuals. Some simple experiments were carried out using Processing, such as creating a grid of circles that increase in size based upon the amplitude of the audio signal. Processing was also used to generate MFCC, FFT, and spectrogram visualisations from different New Zealand bird songs to examine if there were any discernible differences. However, because of the limitations of Processing's 3D visual generation when compared to programs such as Touch Designer or Houdini, it was decided that it would not be a viable option for creating an audio-visual design system.



```
import ddf.minim.*;

Minim minim;
AudioPlayer song;

int spacing = 16; // space between lines in pixels
int border = spacing*2; // top, left, right, bottom border
int amplification = 3; // frequency amplification factor
int y = spacing;
float ySteps; // number of lines in y direction

void setup() {
  size(800, 800);
  background(255);
  strokeWeight(1);
  stroke(0);
  noFill();
  minim = new Minim(this);
  song = minim.loadFile("kea-song.mp3");
  song.play();
}

void draw() {
  int screenSize = int((width-2*border)*(height-1.5*border)/spacing);
  int x = int(map(song.position(), 0, song.length(), 0, screenSize));
  ySteps = x/(width-2*border); // calculate amount of lines
  x -= (width-2*border)*ySteps; // set new x position for each line
  float frequency = song.mix.get(int(x))*spacing*amplification;
  ellipse(x+border, y*ySteps+border, frequency, frequency);
}

void stop() {
  song.close();
  minim.stop();
  super.stop();
}
```

▲ **Figure 19.** *Processing script which reads audio and converts it into ellipses based on the amplitude of the audio signal.*

EVALUATION

Following these experiments, it was decided that Touch Designer would be used for the next phase of the design development due to its real-time visual generation which allows for a fast workflow when drafting design ideas. Touch Designer also has a complex range of audio manipulation tools due to it being created for real-time audio-visual performances and installations. Using these experiments as a foundation, preliminary criteria were developed for the next phase of the design development.

- The system should generate 3D visualisations so that the resulting form can be examined from multiple angles.

- The system should be able to contrast and compare different New Zealand bird songs and generate identifiable visual differences or similarities, which have a direct relation to the difference in sound.

- Sound data of New Zealand bird song should be able to be used as an input to generate comparable forms using the final system.

- The audio-visual mapping choices should consider human's phonaesthetic understanding of audio-visual connections, such as the maluma/takete phenomenon.

IV

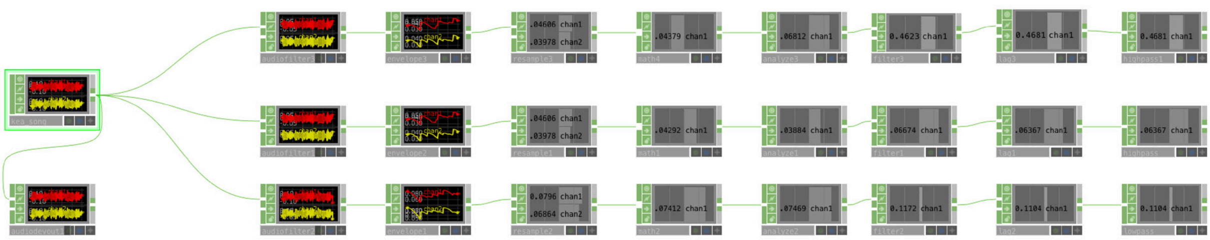
CHAPTER FOUR

TOUCH DESIGNER

AUDIO PROCESSING

To begin developing the audio-visual design system using Touch Designer, a range of bird songs were sourced from both The Department of Conservation (n.d.) and McPherson (n.d.). After gathering a collection of bird songs, the auditory data could then be used to test whether the system could be used to generate visual comparisons of New Zealand bird song, as outlined in the criteria.

Touch Designer can be used to separate a defined frequency range from an audio file. This was especially useful for mapping the audio parameters of New Zealand bird song as most species of birds have a condensed frequency range of between 1,000 Hz - 8,000 Hz (Sasahara et al. 2012). Audio files were divided into numerical values to represent the low, medium, and high frequencies of the bird song, as seen in Figure 20. Numerical values extracted from the audio files were then connected to separate channels which could manipulate these values for the generation of visual elements.



▲ **Figure 20.** *Touch Designer audio processing system.*

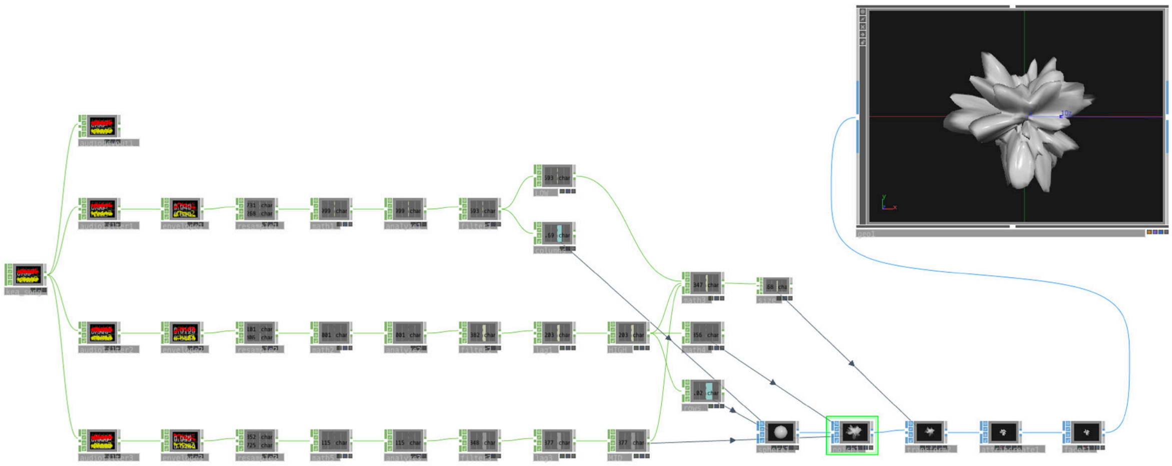
AUDIO-VISUAL MAPPING
DEVELOPMENT

Form

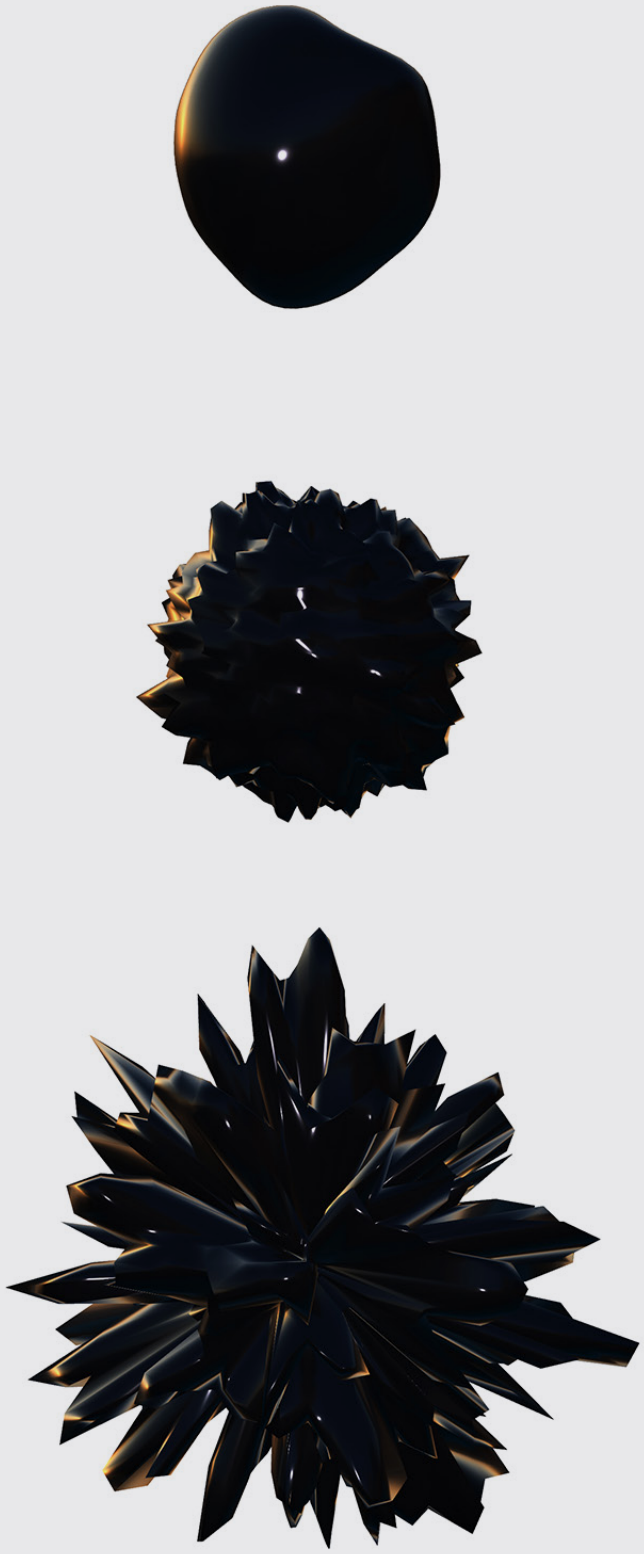
After separating audio files into low, medium, and high frequencies, the next stage of the design process was to explore how form could be manipulated to represent specific elements of New Zealand bird song.

In order to manipulate the form of the mesh using audio, a noise operator was applied, which allows for the creation of generative shapes through the mapping of specific properties, such as amplitude and roughness. This allowed for the manipulation of a 3D shape so that it became spikier when a higher frequency was occurring within the bird song, and smoother when a lower frequency was occurring. This created visual variance when comparing different bird songs, and was also reflective of the sound through referencing the maluma/takete phenomenon explored within the background research.

This prototype allowed for definitive mapping decisions based upon a selected range of the audio spectrum, as well as for the definition of how smooth or spiky the shape became based upon the specified frequency range.



▲ **Figure 21.** *Connecting noise manipulation to a form.*



▲ **Figure 22.** *Touch Designer audio forms - low frequency (round shape) to high frequency (spiky shape).*

▲ **Figure 23.** *Adding a feedback trail to an audio-driven form.*

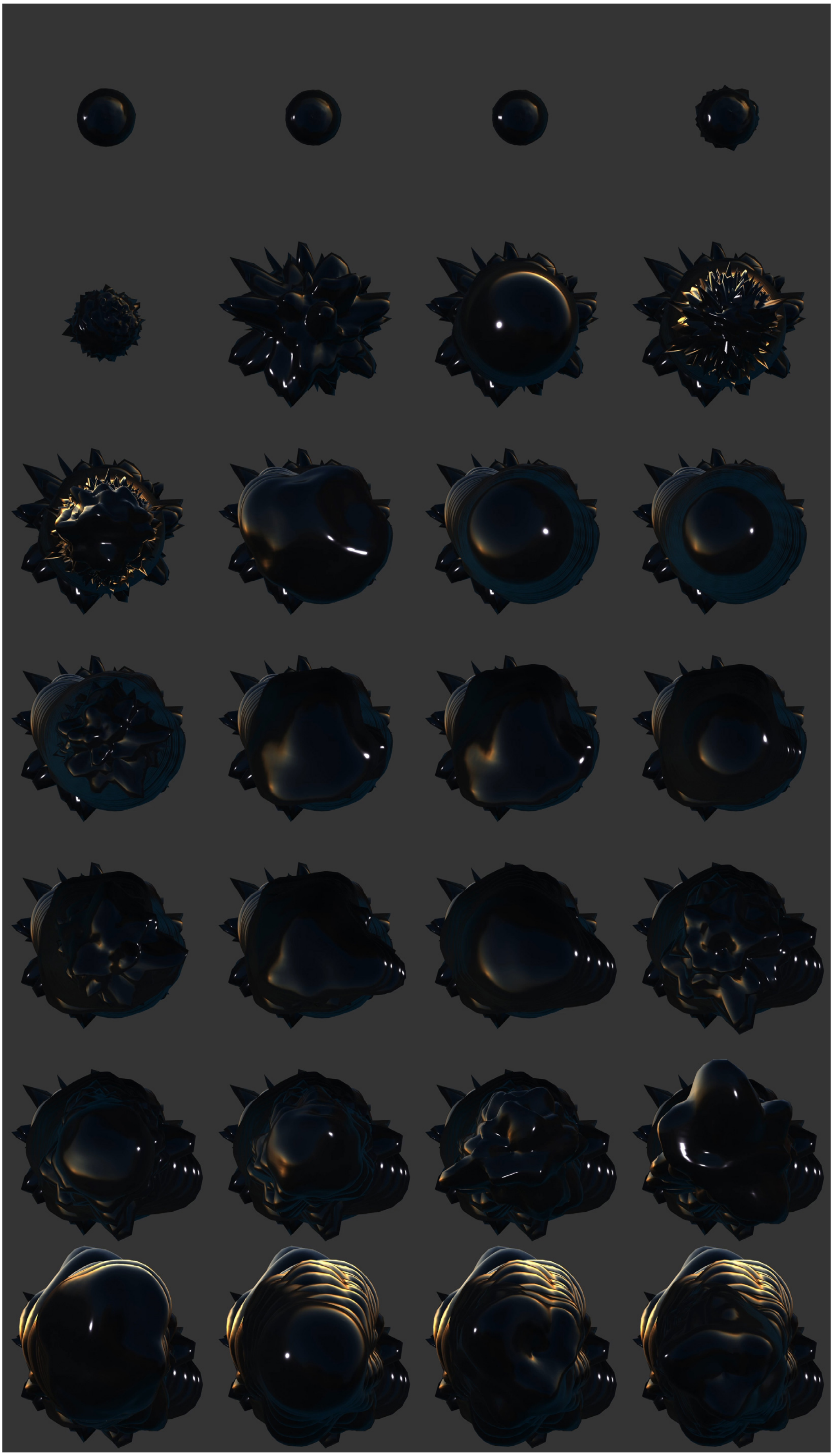
Movement

At this stage of the design development, methods for visually representing the timing and pattern of sounds were explored. Timing and pattern are some of the most definable properties of sound, and bird song, in particular, has a strong emphasis on the timing and pattern of frequencies. It was decided that this could be represented through the use of a visual feedback loop, which would create a trail of the shapes movement as the bird song evolved. This allowed viewers to see the past movements and form of the shape over time, as seen in Figures 24 and 25.

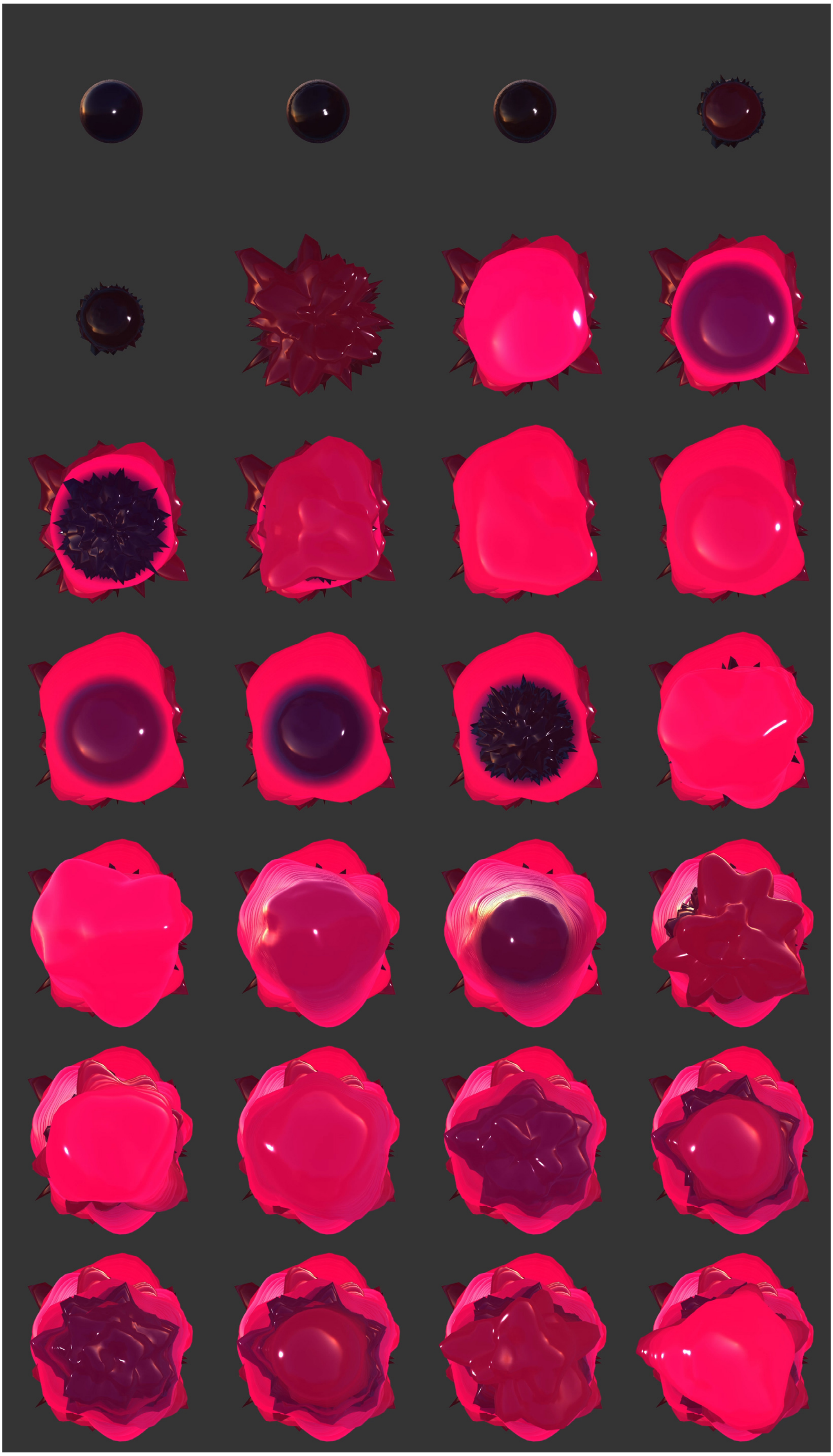
Integrating this feedback loop using a static shape was relatively successful. To further manipulate the mesh, the shape was manipulated to become larger when there was an increase in volume. However, once the shape increased in size, previous forms were no longer visible. To address this issue, the shape was animated to move across the screen from left to right. This decision was made to match how information is read and understood in other formats. This allowed for the trail of the shape to remain visible and allowed for the visual identification of patterns occurring within the bird song.

After further experimentation, it was found that moving the shape upwards as the volume increased created more abstract and varied shapes as frequencies occurred at different volume levels. This created further visual variance between bird songs. For example, a bird with a more complex song such as a Tui would be reflected through a more intricate form, while a simpler bird song such as a Fantail would be visualised as a simpler form, as can be observed in Figure 26. The larger variety of audio frequencies present in the bird song of the Tui created a greater variation in the form, size, and colour of the shape as it moved and changed over time when compared to the song of a Fantail which has a more condensed frequency range.

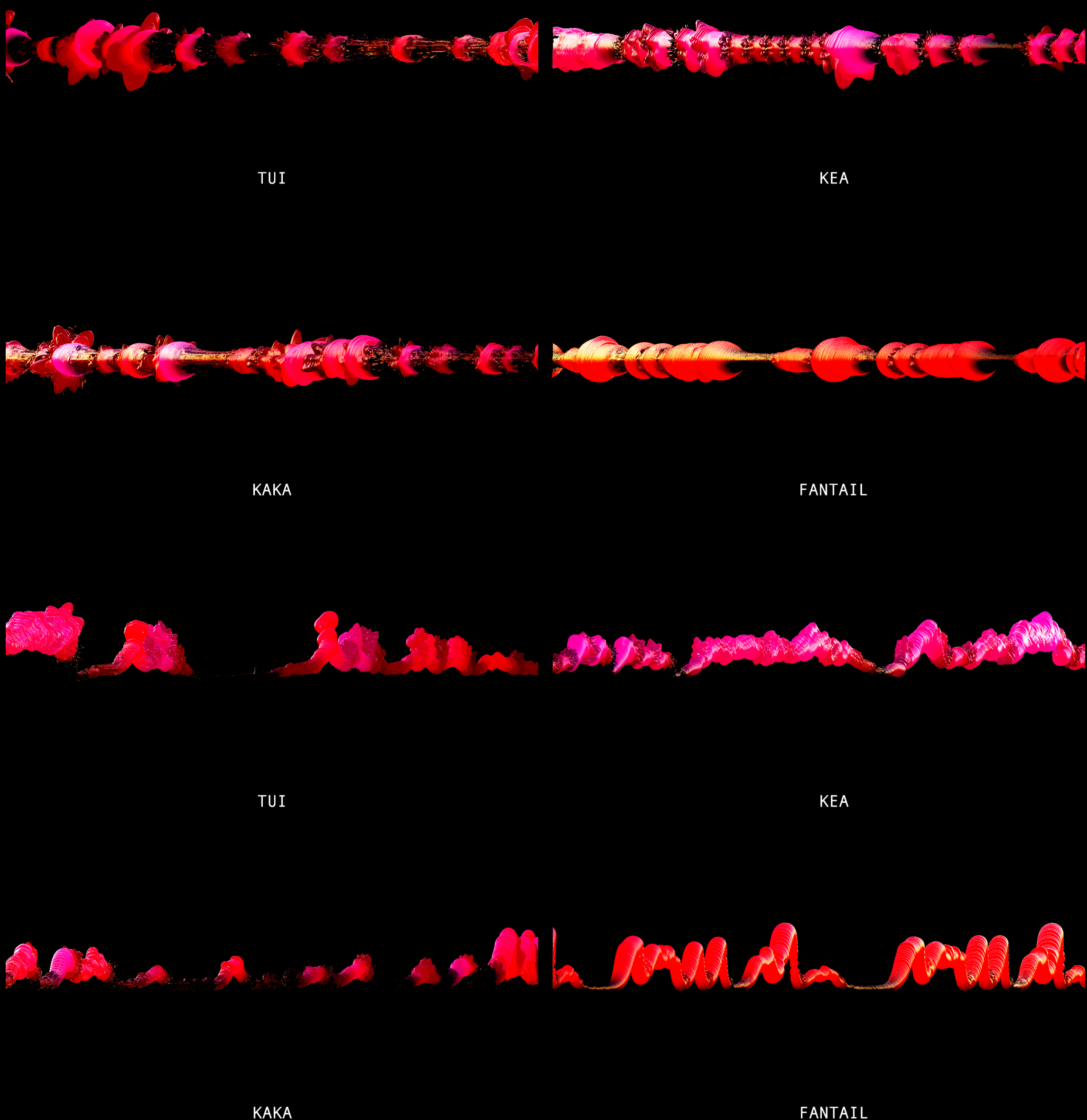
Other experiments also consisted of moving the shape in vertical space once it reached the end of the screen, and having the trail of the form fade away over time. However, these experiments did not provide any intuitive benefit to the system. At this stage of the design process the decision was made to have the form generate from left to right, as the resulting visualisations could be read intuitively and were able to represent the timing and pattern of bird song.



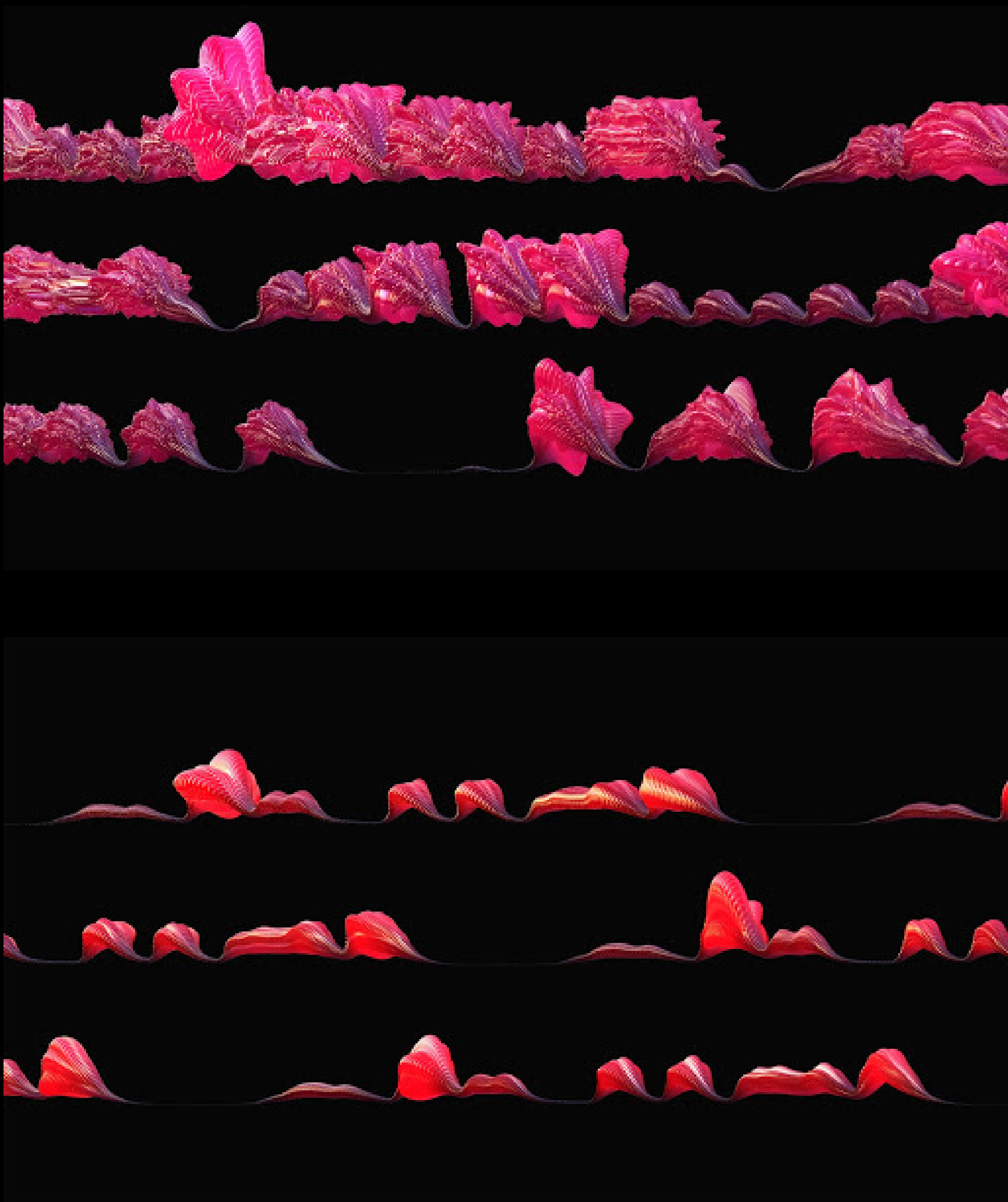
▲ **Figure 24.** *Audio-driven feedback trail timelapse.*



▲ **Figure 25.** *Audio-driven feedback trail timelapse - coloured.*



▲ **Figure 26.** *Audio-visual experiments created using Touch Designer comparing different New Zealand bird songs.*

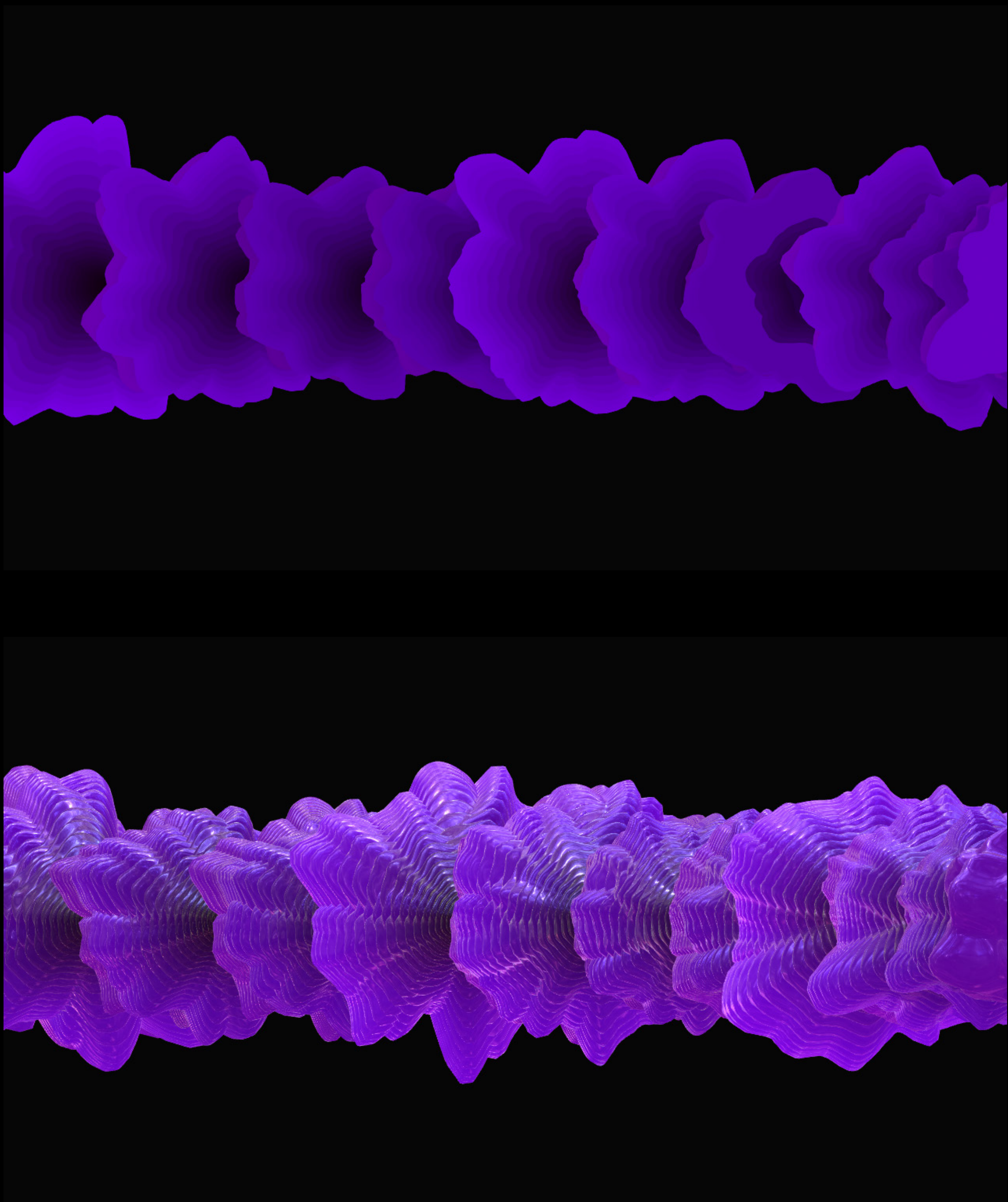


▲ **Figure 27.** *Kea song (top) and Fantail song (bottom) being visualised with a vertical time system.*

Colour

Colour is one of the most distinct visual elements for representing sound, so the consideration of how it would be used to represent bird song was a crucial part of the design process. For the initial experimentation, higher frequencies were mapped to light red colour values, while lower frequencies were mapped to dark blue colour values. This resulted in lighter shades of pink and red when a higher frequency was being visualised, and darker shades of blue and black when a lower frequency was being visualised. It also created variation in colour over time as the hue would shift to a purple shade when there were both low and high frequencies occurring within the bird song. This mapping decision was based upon the research of Ludwig et al. (2011) who explain that humans have an intuitive instinct to link high frequencies with lighter colours and lower frequencies with darker colours. A colour system still needed to be developed which could integrate a greater range of colours. At this stage, the colour mapping was limited to individual channels from either the red, green, or blue colour channels.

Adding texture to the form greatly improved its aesthetic appearance, as well as adding further definition to the shape through the increased reflection of the light as can be seen in Figure 28.



▲ **Figure 28.** *Touch Designer renders without environment light (top) and with environment light (bottom).*

EVALUATION

As the Touch Designer programming language is optimised to generate real-time visuals, it allowed for simple and efficient adjustments to be made to both the audio manipulation and the generated visual elements. At this stage of the design process, audio-visualisations which varied in form, size and colour could be generated using New Zealand bird song as an input. However, visualisations of different bird songs lacked the range of expression required to reflect unique sonic elements, or enable viewers to visually identify the differences between different New Zealand bird songs.

This highlighted the need for the audio-visual system to be able to identify a range of defined frequencies, and for the system to have enough visual variance to show a definable difference between different bird songs in order to meet the criteria.

The trail concept allowed the viewer to identify patterns within bird song recordings. However, it was 2D which meant that it could not be viewed from multiple angles, and limited the amount of detail able to be observed by the viewer. This limited the aesthetic quality of the visualisation, and needed more development in order to meet the criteria.

At this stage of the design process, it was decided that more work was needed within the audio processing segment of the system. There needed to be a more defined, accurate measurement of sound to create greater visual variance between each bird song. Other visual possibilities also needed to be explored to further integrate the maluma/takete phenomenon and other phonaesthetic audio-visual connections explored within the background research.

As Houdini is not a real-time programming language, it does not have a focus on the optimisation of visual elements and instead offers a much larger variety of audio processing and visual manipulation capabilities. Therefore, it was decided that Houdini would be used for the next phase of the design process.

Following the evaluation of this phase of the research, the criteria were expanded upon:

- The system should generate 3D visualisations so that the resulting form can be examined from multiple angles.

- The system should be able to contrast and compare different New Zealand bird songs and generate identifiable visual differences or similarities, which have a direct relation to the difference in sound.

- Sound data of New Zealand bird song should be able to be used as an input to generate comparable forms using the final system.

- The audio-visual mapping choices should consider human's phonaesthetic understanding of audio-visual connections, such as the maluma/takete phenomenon.

- The system should be able to identify changes in pitch and have enough visual variance to represent the frequency range of New Zealand bird song.

- The movement of the visual elements should have a direct relation to the sound.

V

CHAPTER FIVE

HOUDINI

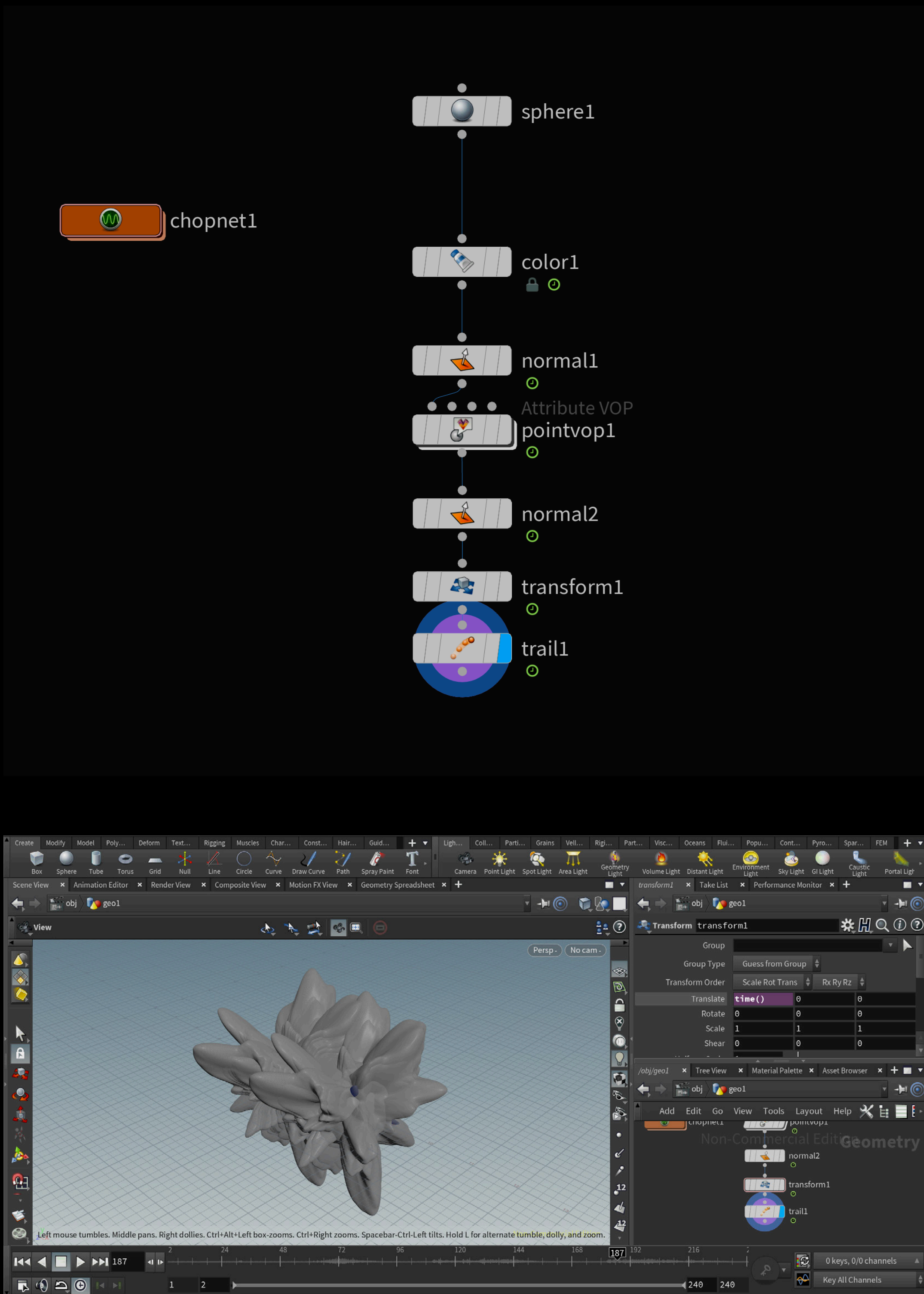
REBUILDING THE SYSTEM

The audio-visual system developed with Touch Designer was rebuilt using Houdini where it could then be further iterated upon. A CHOP network was used to manipulate numerical data from the audio files. This data was then remapped into specified values which could be used to reflect audio attributes through visual elements.

By rebuilding the system within Houdini, the established audio-visual design system was able to be built and iterated upon without the limitations of real-time optimisation. As Houdini is not a real-time engine like Touch Designer, it had a larger range of visual possibilities, but there had to be more consideration in regards to optimisation. Due to the slower iteration process, it was beneficial to begin programming the system with a more informed workflow. However, as identified in the criteria, greater visual variance was required to represent the frequency range of New Zealand bird song within the design system, which made the switch to Houdini necessary.

To recreate the maluma/takete effect, noise was applied to manipulate a form. This allowed for the distortion and transformation of the shape by manipulating the noise levels through CHOPS. The form became either smooth or spiky depending on the amplitude of the noise.

The trail effect was one of the most successful aspects of the audio-visual system developed using Touch Designer. Therefore, it was decided that this concept would be developed further using Houdini. While the Touch Designer trail allowed the viewer to identify patterns within bird song recordings, it was 2D which limited the amount of detail able to be observed.

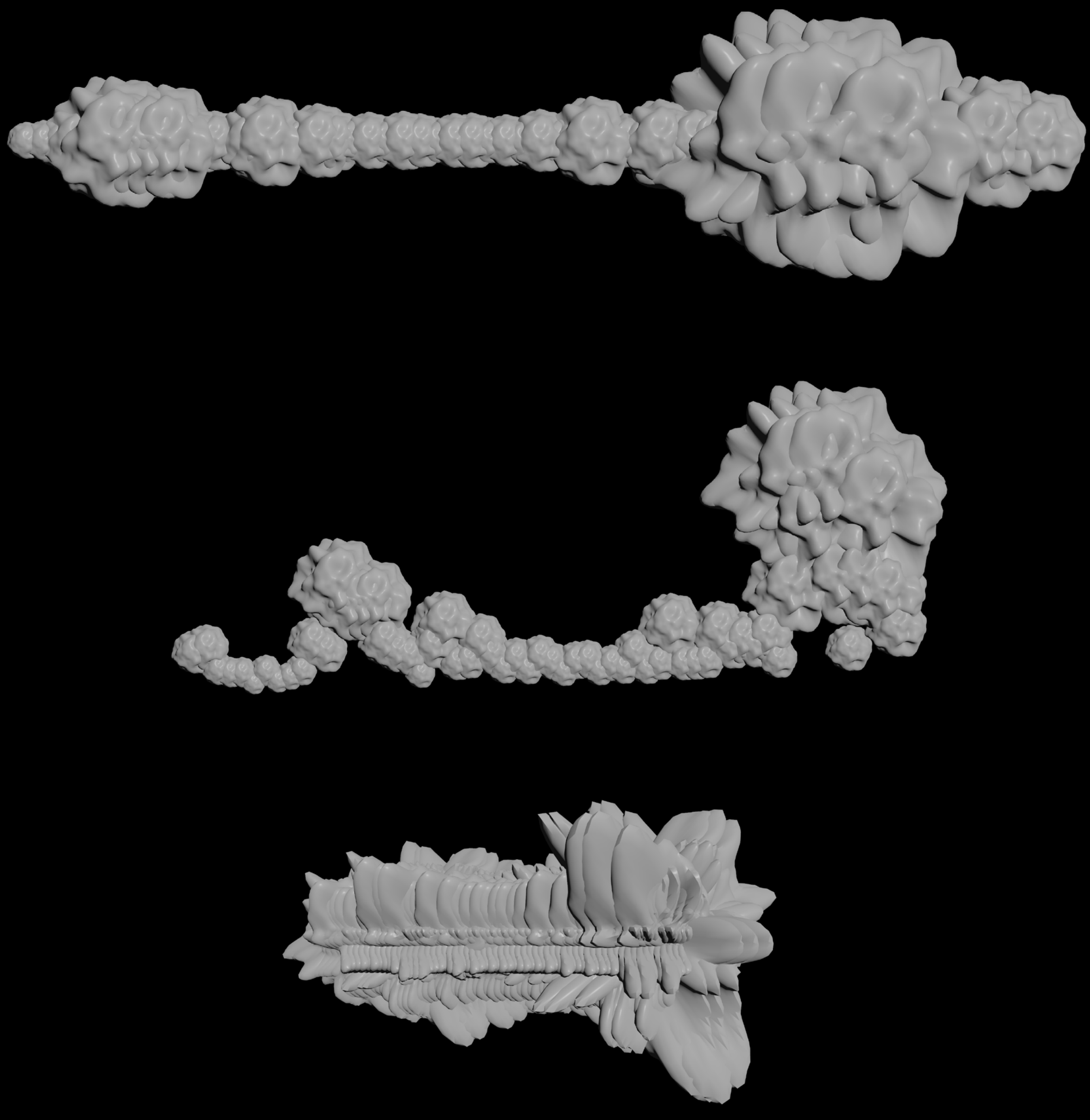


▲ **Figure 29.** *Rebuilding the audio-visual system using Houdini.*

To achieve a trail effect in Houdini, the form was connected to a transform node which moved the shape along a vertical axis. A trail node was then applied to the form, which allowed it to operate within a 3D space and to be orientated by the camera and examined in close detail, in accordance with the criteria.

Several techniques were explored to develop the audio processing setup within Houdini. The initial audio setup was similar to Touch Designer, with the volume levels of the bird song being used to control the size and form of the shape. At this stage, high and low frequencies had been separated at a very basic level, with any frequency over 2,000 Hz making the form spikier. Visual experiments were carried out which explored manipulating the form in various ways.

A basic colour mapping system was established, which linked higher frequencies to the colour red and lower frequencies to the colour blue. However, once again the system lacked both legibility and variance between bird songs. Most bird songs would end up being roughly the same colour due to only the red and blue colour values being mapped. Because of this limitation, it was decided that further possibilities for mapping audio frequency to colour would be explored in order to meet the criteria and create identifiable variations between bird songs. The next stage of the design process was to explore the audio manipulation capabilities of Houdini and to improve the audio-visual mapping of the design system.



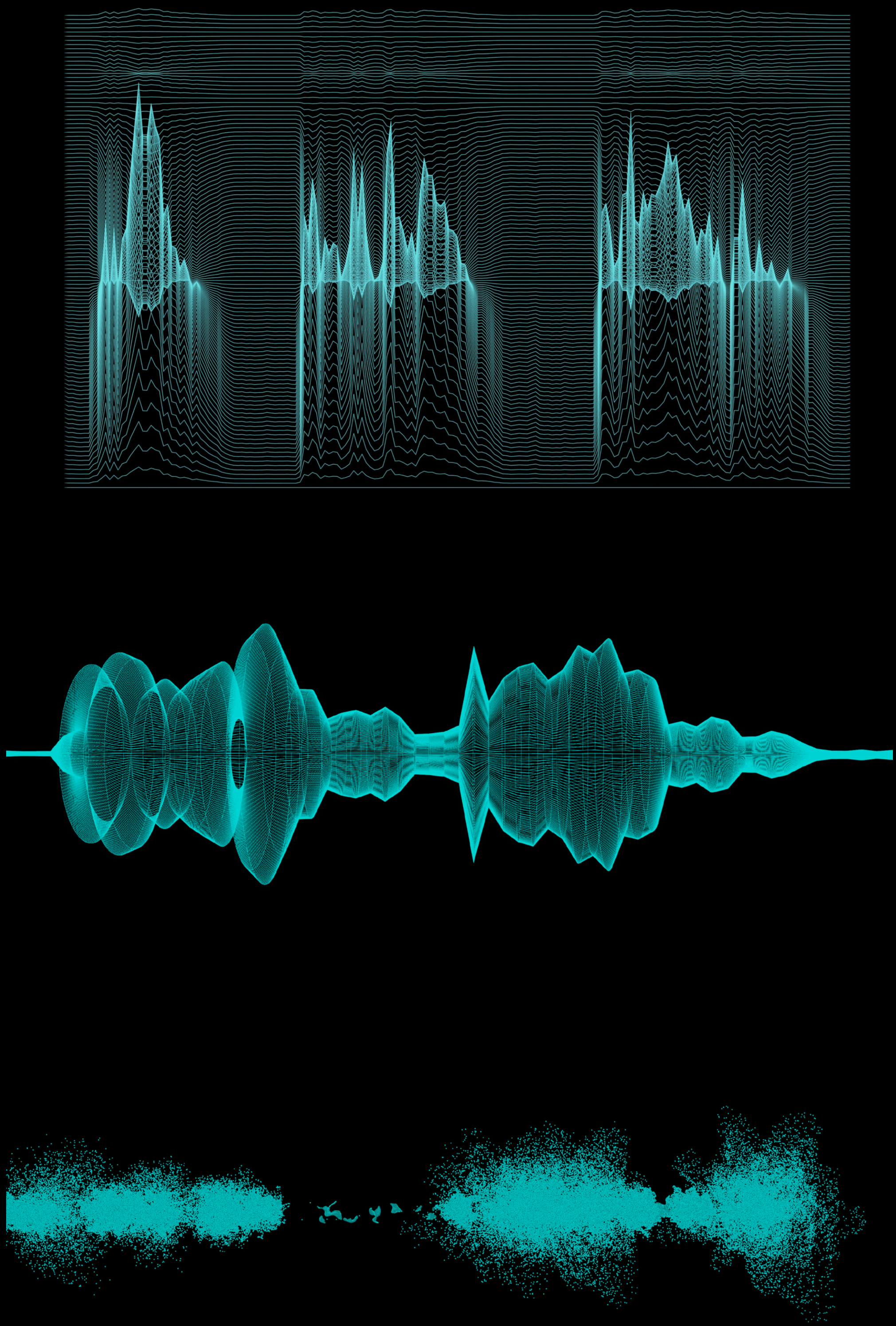
▲ **Figure 30.** *Experimenting with form manipulation using Houdini.*

AUDIO VISUAL MAPPING
REFINEMENT

In order to improve the clarity of the mapping parameters, a variety of visual experiments were carried out, experimenting with how form, movement, and colour could be mapped using sound. The shapes currently being generated were not defined well enough to intuitively reflect sounds, and did not have enough variance to reflect the definitive elements of bird song. As this was identified as being a key feature of the design system within the criteria, the next phase was to improve upon how sound properties were being visually represented.

FORM

With the audio processing segment of the audio-visual system functional, aesthetic options were then explored to determine what would be the most suitable for the audio-visualisation of New Zealand bird song. As outlined by Kirk (2016), aesthetics are an important design element to consider for the final design, as they have the potential to draw interest towards the project. Ideally, the aesthetic style of the audio-visual system would have some visual resemblance to New Zealand birds to help viewers associate visual elements of the design with what they are hearing.



▲ **Figure 31.** *Experiments with procedural form generation.*

Using a line as the visual basis of the system resulted in a visual effect which was similar to a typical soundwave. However, this method was lacking in visual variance and therefore was not effective for comparing and contrasting different bird songs. Furthermore, this concept also lacked innovation as this is already a common method in the field of sound visualisation.

Design experimentation at this phase also explored the possibility of generating a shape that connected as a smooth mesh. However, this created issues surrounding conflicting faces and broken geometry. This was due to the unpredictability of procedural shape generation when using sound as a design tool for generating visual elements.

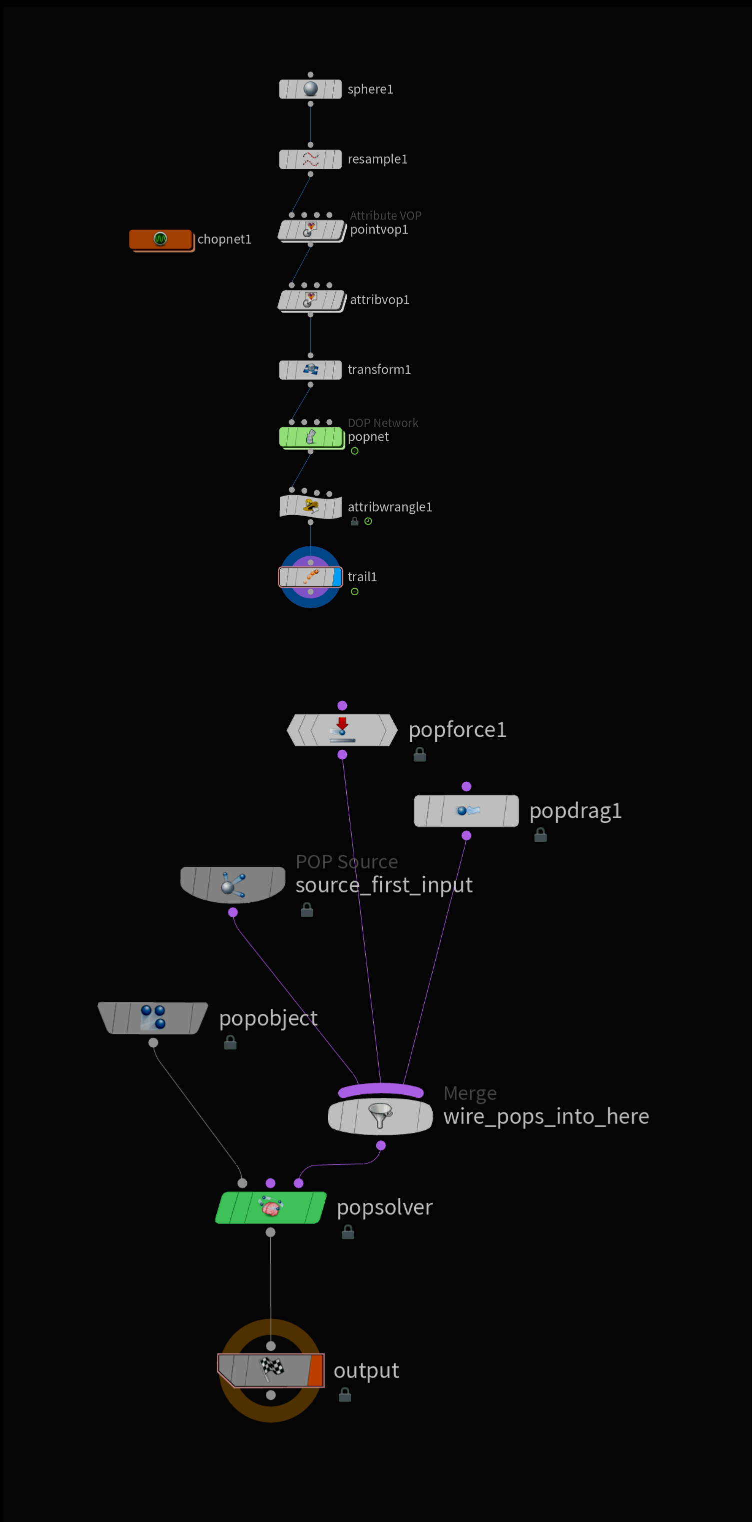
It was found that the most successful result was achieved through the use of a sphere. Transforming the sphere into scattered points created an aesthetically engaging visualisation which allowed the viewer to examine the interior and exterior of the shape from multiple angles. The generated form could be manipulated to become either spiky or smooth depending on the frequency. This experiment also contained the largest amount of visual variance compared to previous options.

At this phase of the design process, the optimisation of geometry had to be considered due to the large amount of computing power required to simultaneously read sound data and generate procedural geometry. The generated geometry was converted into a particle system, as seen in Figure 33, which reduced loading times and made the iteration process more manageable. Through converting the form into a particle system, the movement and generation of individual particles could then be manipulated.

A Houdini particle network allows for the designer to specify the number of seconds the particles exist. By modifying this parameter, the visible length of the audio trail could be manipulated. The amount of particles being generated could also be specified within the particle network. Both these parameters were reduced to a lower value to optimise the design system for the development phase.

The particle system was developed further, integrating the trail effect which had been explored in previous design experiments. The addition of a trail operator to the particle network meant that the individual movement of each particle could be manipulated to create more expressive forms. These forms also had somewhat of a resemblance to New Zealand birds, with the generated forms looking like fur or feathers. Therefore, it was decided that this iteration of the design would be developed further.

Further audio-visual manipulation of the particles was then explored. It was discovered that by controlling the size of the shape being used to generate particles, the particle trails could be manipulated to vary in density based upon the frequency of the occurring pitch. Denser forms appeared smoother, while more sparse forms appeared spikier. Therefore, the decision was made to generate denser forms from lower pitches to link to the maluma/takete phenomenon. To improve the readability of this mapping decision, the frequency value was also mapped to control the vertical position of the form. The higher the frequency value, the higher the form would appear in vertical space, which created greater variance for the comparison of different bird songs.



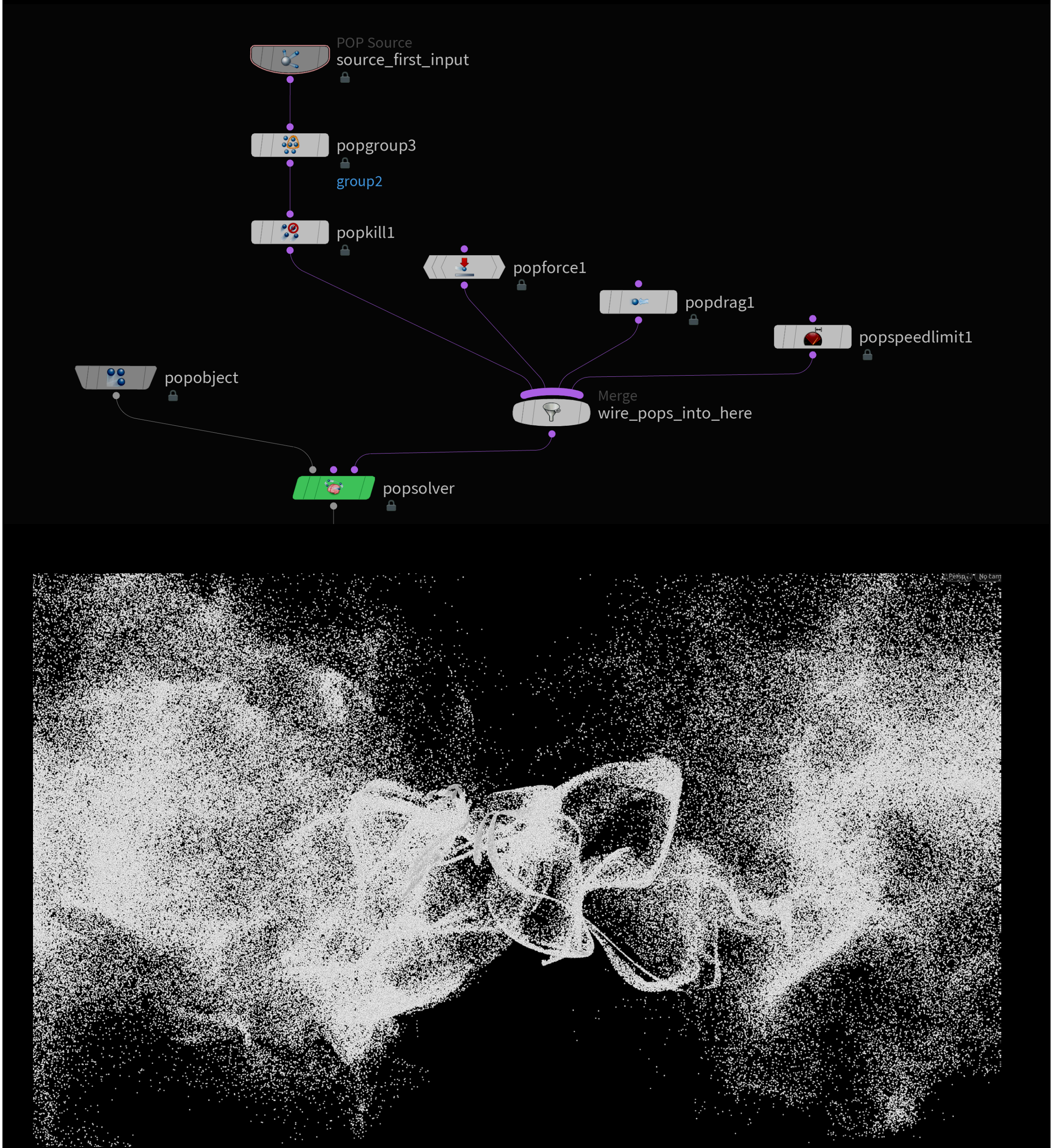
▲ **Figure 32.** *Houdini particle trail system.*

▲ **Figure 33.** *Particle system with no trails (top), and with trails (bottom).*

MOVEMENT

The next phase of the design was to refine the movement of the particles so that they were more representative of the sound being visualised, as defined in the criteria. The main design problem which needed to be addressed was that the movement of the particles was too fast, and the forms being generated would lose their shape due to the rapid dissipation of the particles. Particles would also continue to move at a rapid speed after the sound was over. As a result, bird songs that were being visualised would appear to have very long echoes. Since most bird songs consisted of a series of short chirps, the decision was made to develop this aspect of the design further to have a direct relation to the sound, as specified in the criteria.

To address this issue, the movement of the particles was adapted so shortly after a particle was generated, it would reduce in speed. It was found to be important that the particles did not become completely still, as this made the animation look very stagnant aesthetically and limited the expressiveness of the generated forms. Therefore, reducing the speed to a very low value after a short period of time signified that the frequency had already occurred without compromising the aesthetic of the animation or form. Further experimentation led to the discovery that the particle movement could be manipulated so that the generated forms would have some further visual resemblance to New Zealand birds. Various shapes within each generated form appeared to be similar in form to bird feathers, wings, and beaks. Interestingly, some shapes also resembled birds in flight.



▲ **Figure 34.** *Particle system design iteration.*

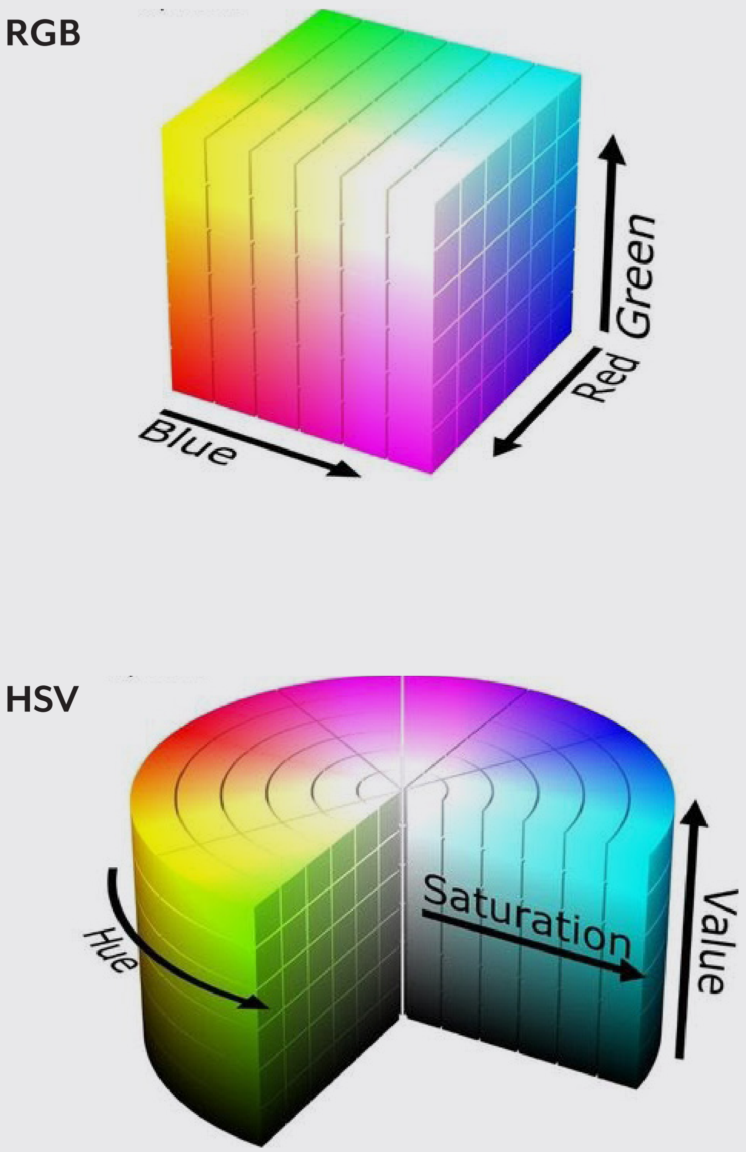
COLOUR

Pass Filter

The audio-visual manipulation of colour in prior experiments had consisted of only red and blue values which limited the colour of the generated forms to red, blue, or a shade of purple. A much more comprehensive colour mapping system was required to take full advantage of the colour spectrum. By using the full colour spectrum, there would be greater visual variance to differentiate audio frequencies with, which would allow for more effective audio-visual comparison of New Zealand bird song.

Houdini colour values are in RGB (red, blue, green) by default, which meant that by default only red, green, or blue values could be controlled using CHOPS. The colour system needed to map the audio levels in an HSV (hue, saturation, value) colour space rather than RGB. This would allow audio parameters to be used to control the full colour spectrum through manipulation of the hue.

After converting the colour space from RGB to HSV, a strategy for mapping the bird song to separate colour values was required. Reflecting upon prior background research, the mel scale provides a method for measuring the perceptual scale of pitch (Nair, 2018). Using a logarithmic conversion algorithm, the mel scale creates a more accurate representation of how humans perceive pitch based on audio frequencies, as seen in Figure 36.

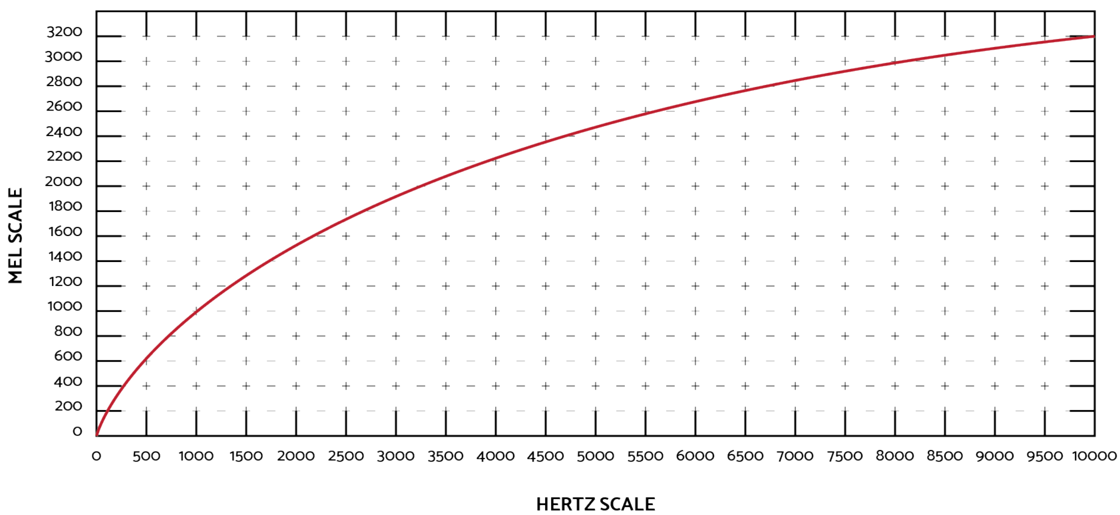


▲ **Figure 35.** *RGB and HSV colour space comparison.*

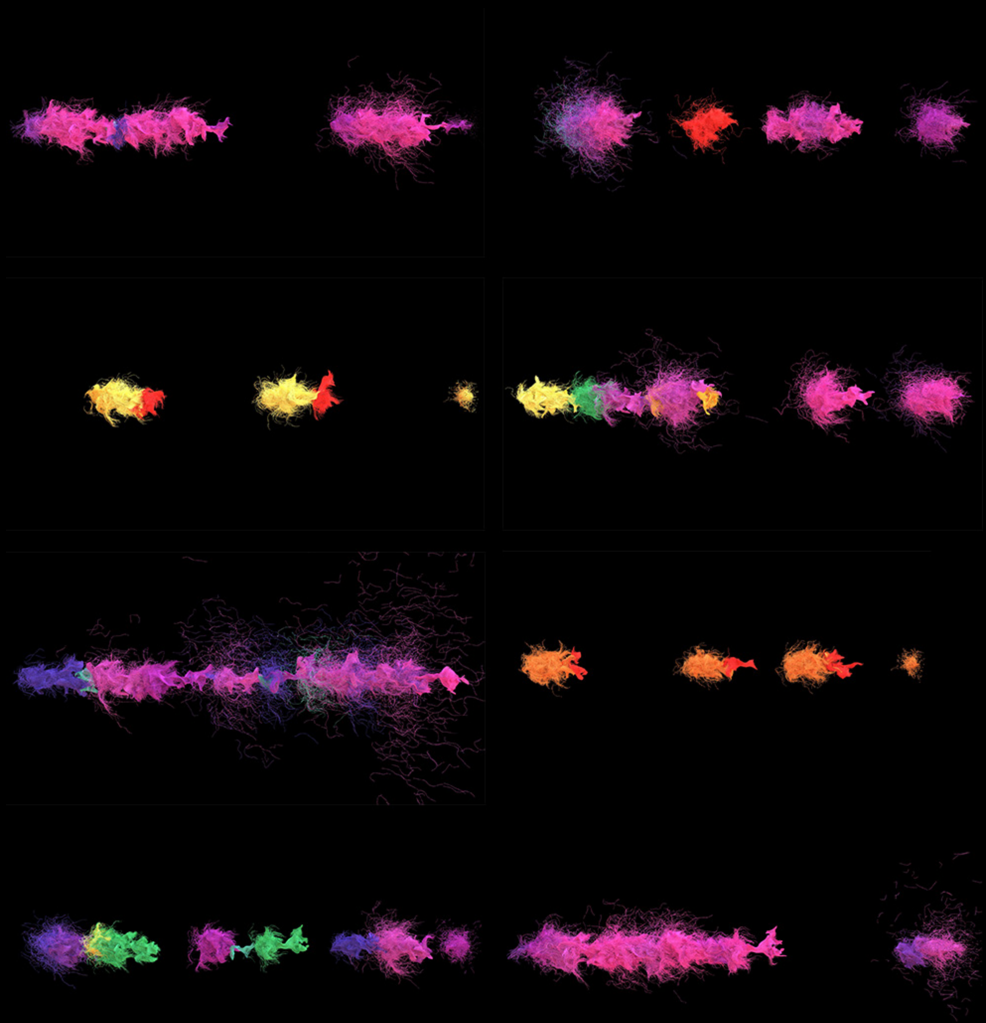
Note. From "Wikimedia Commons" by SharkD, 2008. (https://commons.wikimedia.org/wiki/File:Color_solid_comparison_hsl_hsv_rgb_cone_sphere_cube_cylinder.png). In the public domain.

The mel scale was used to separate audio files into frequency divisions which associated the frequency value to its estimated pitch using a pass filter CHOP. To trial this method, the system focused on mapping the twelve notes of the chromatic scale across one octave. Each frequency value was mapped to a specified colour value, where the highest occurring colour value would then be calculated and mapped to the form.

While this allowed for the full spectrum of colours to be used for audio-visual mapping, the system would always prioritise the highest colour value, rather than the most prominent pitch within the bird song. This made the audio-visual mapping inaccurate, as the colours most commonly mapped were those with higher numerical colour values according to the HSV colour space. A new colour mapping system was required in order to determine the highest occurring frequency. Following an analysis of the values being output from the audio files, it was discovered that the pass filter CHOP was not functioning accurately. It was relatively successful at separating high and low frequencies, but lost accuracy in isolating a more specific range of frequencies which was required for the visualisation of bird song. Therefore, it was decided that other pitch-to-colour mapping options would be explored.



▲ **Figure 36.** *Diagram illustrating the mel scale.*



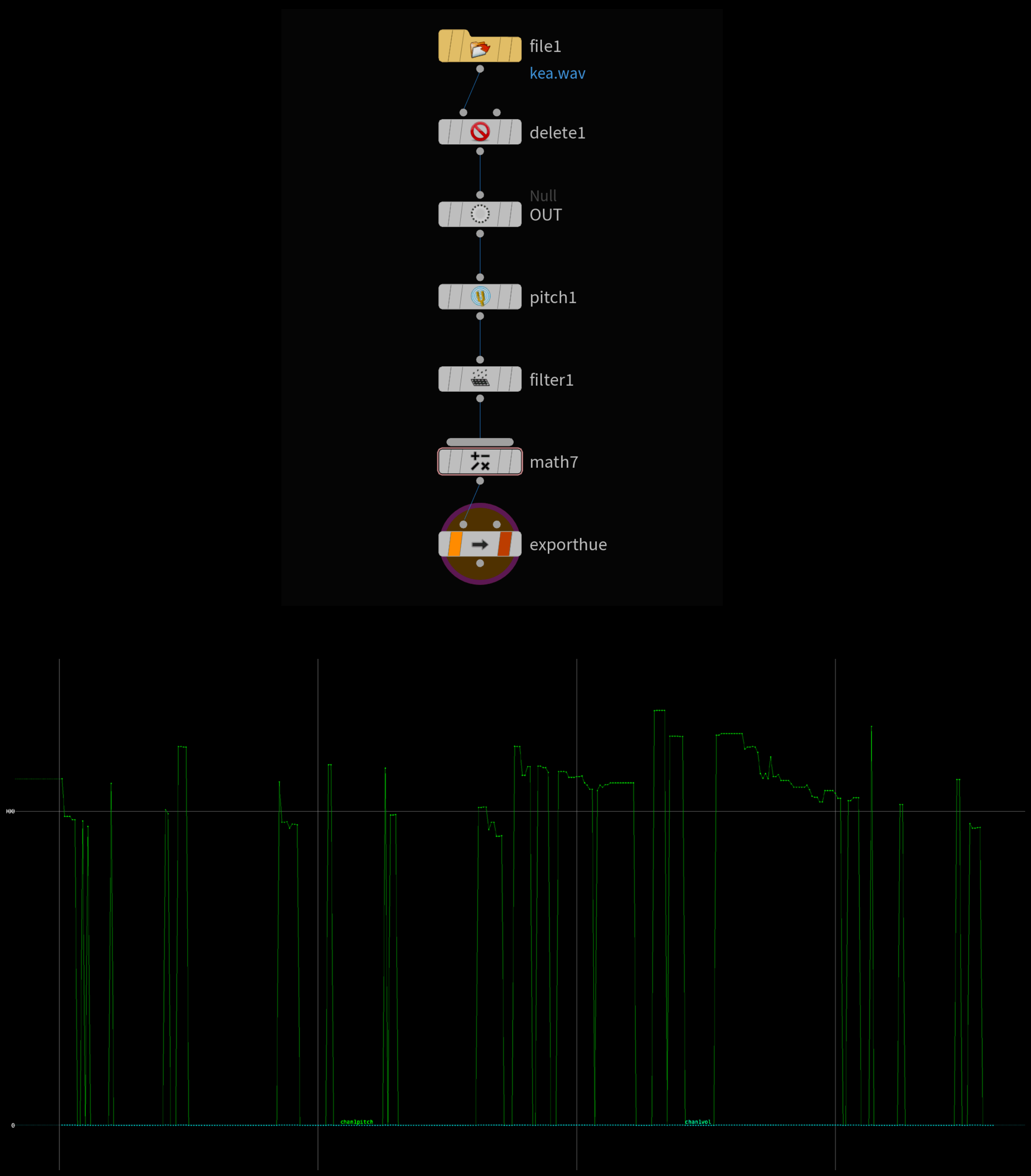
▲ **Figure 37.** *Design iteration using the initial colour mapping system.*

Experimenting with Pitch Detection

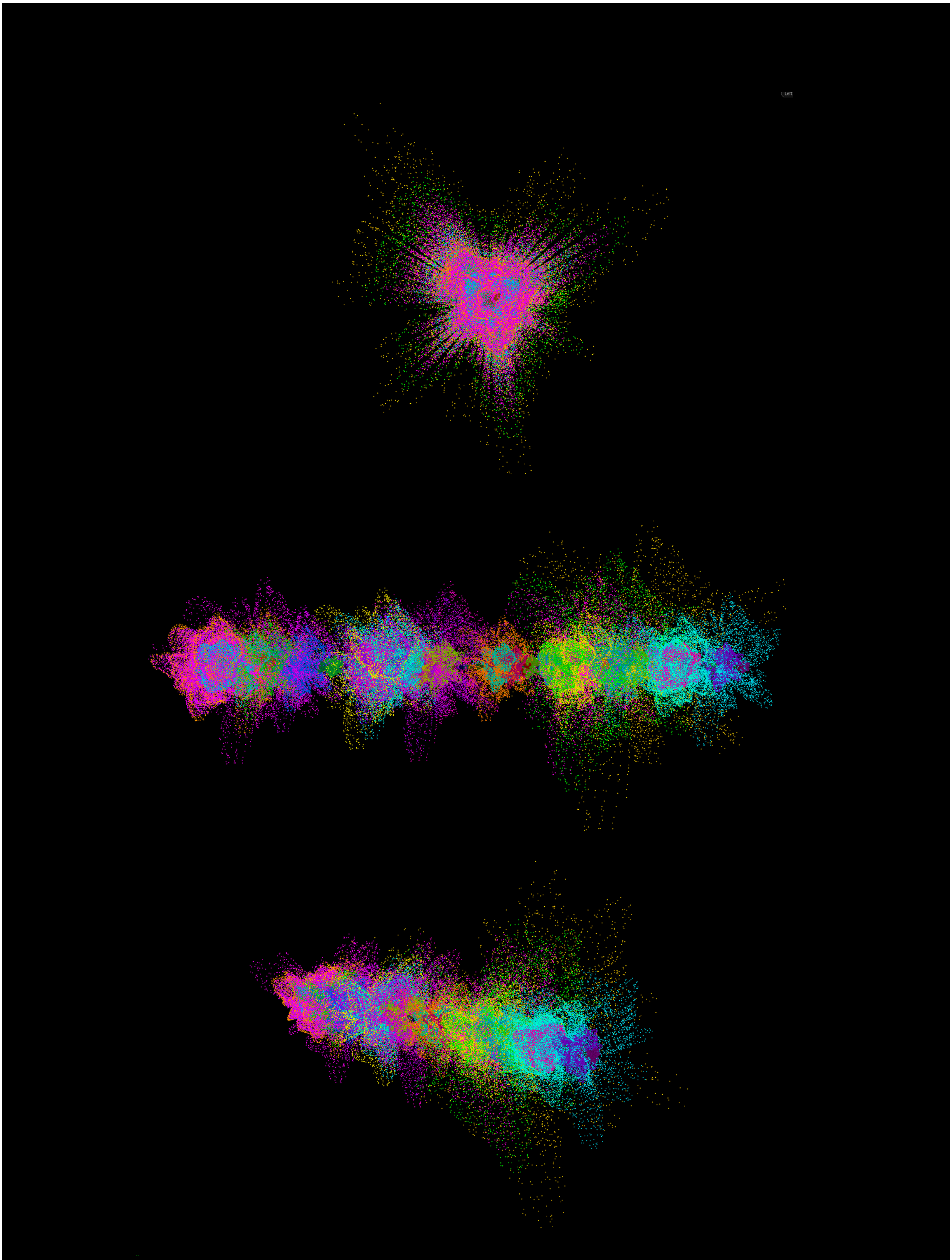
Since the pass filter CHOP was not functioning as accurately as was required, the next stage of the design process explored using Houdini's inbuilt pitch detector CHOP.

Houdini's inbuilt pitch detector determines the most dominant frequencies within an audio file over time and converts the information into numerical values. Depending on the bird song, these values ranged from 1,000 to 8,000 Hz. HSV colour values range from 0 - 360, so these values were remapped accordingly. This meant that a high-frequency would generate a colour higher on the spectrum according to the HSV colour space.

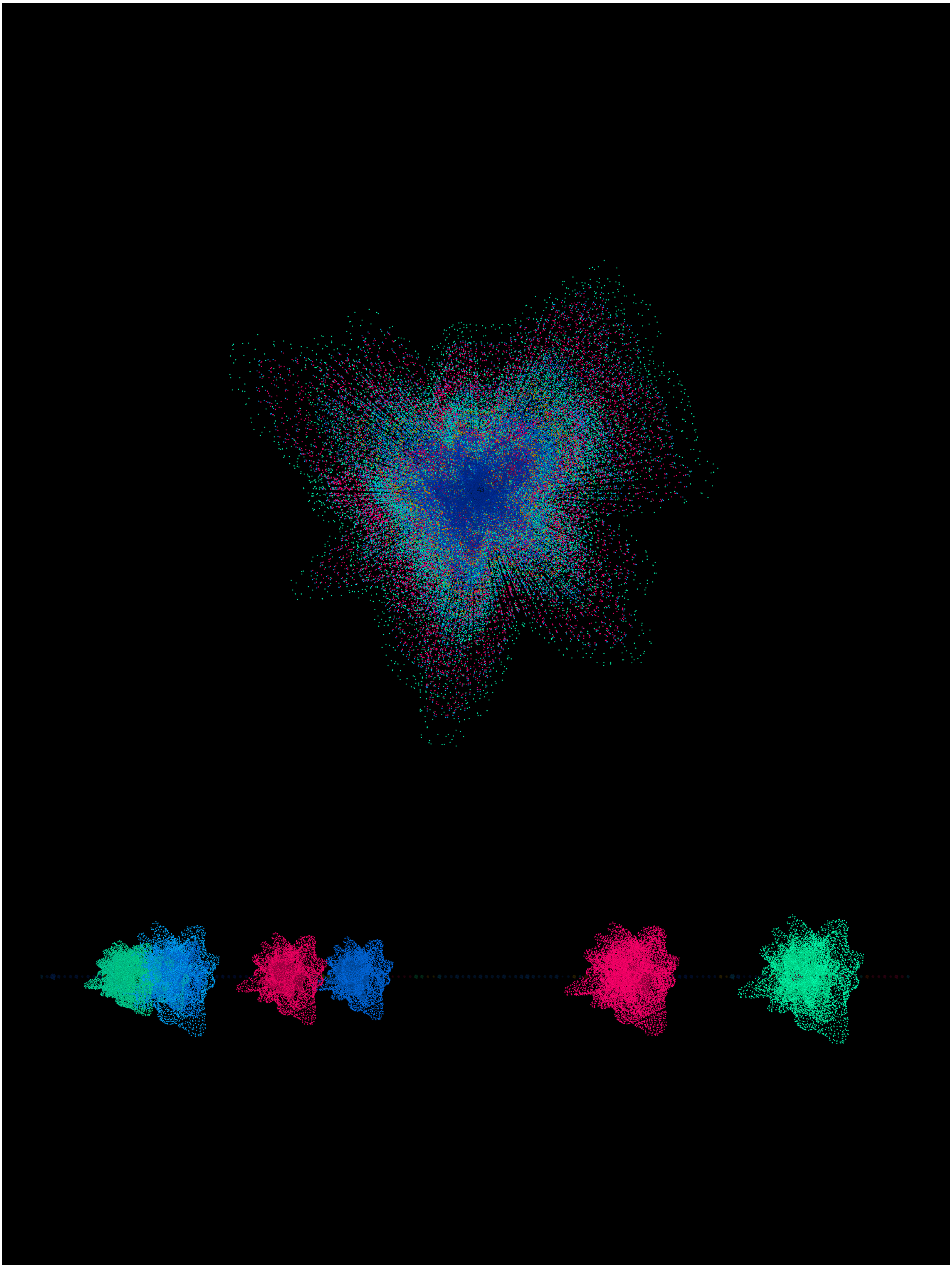
Although this system was more successful than the previous, it was once again not accurate enough to reflect changes in pitch through colour. Frequency values would often accelerate rapidly from zero to a specified colour value due to the high frequency range of bird song. This meant that the assigned colour value would cycle through the spectrum in accord with the increase in frequency. This would result in the visualisations containing every colour on the spectrum, resulting in a rainbow effect. Therefore, it was decided that Houdini's inbuilt pitch detector did not provide enough control of the frequency and colour values to create an accurate mapping system.



▲ **Figure 38.** *Pitch detector CHOP mapping frequency values.*



▲ **Figure 39.** *Houdini visualisation of a Bellbird song with the system mapping the entire colour spectrum to the generated form.*



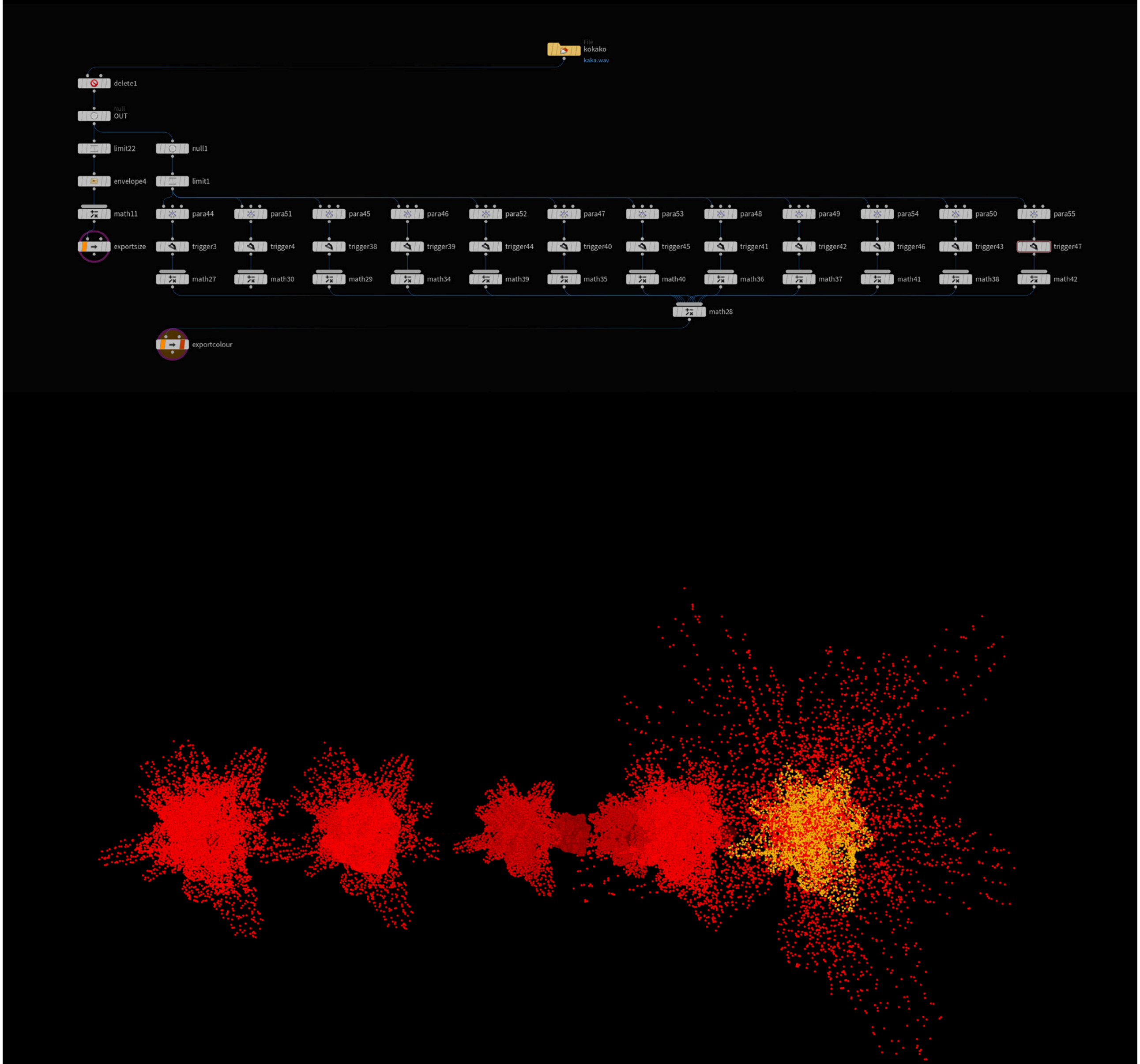
▲ **Figure 40.** *Houdini visualisation after remapping the colour values of the pitch detector.*

Building a Trigger System

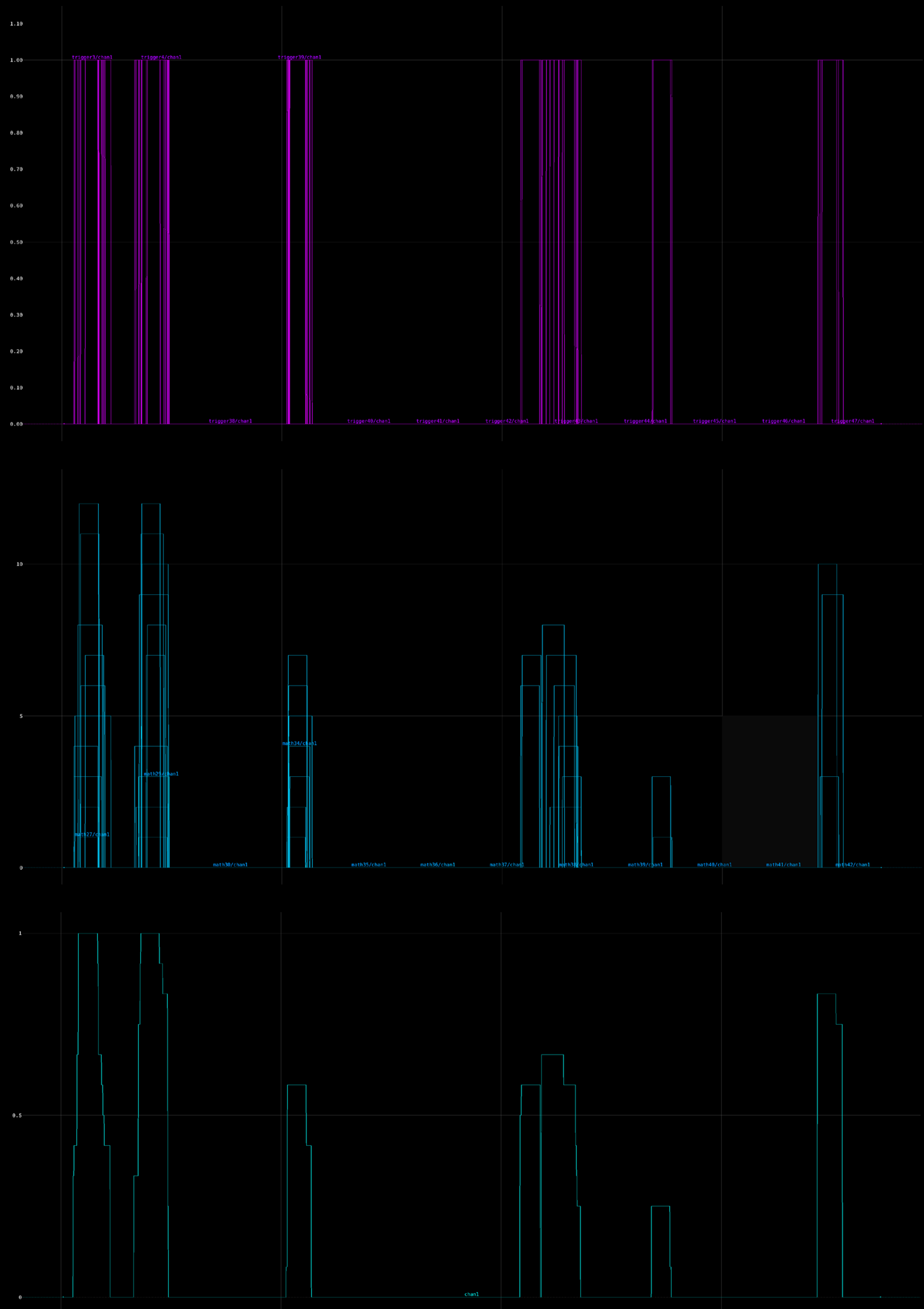
To address this issue, the next phase of the design process was to explore the potential of using Houdini's trigger CHOP to activate colour values. This meant that colours would only be mapped to the form when a frequency range reached a defined volume level.

Since both the bandpass and the pitch detector had lacked the accuracy required to map frequency to colour, for this experiment Houdini's parametric equaliser CHOP was used to divide frequency values. The parametric equaliser allowed a more defined selection of frequencies. The octave of frequencies being mapped was divided into twelve, to correspond with the twelve notes of the chromatic scale. The mel scale was once again used to define at what value frequencies should be separated.

This method for mapping frequency to colour was much more accurate. Rather than the colour value being directly mapped from the frequency value. The trigger system activated the colour value once the corresponding frequency met a defined volume threshold. This meant that the frequency range, colour value, and volume threshold could all be manipulated and manually defined, providing more control of the individual mapping parameters. This method also prevented the rainbow effect from occurring as the colour value would no longer cycle up with frequency, as seen in Figure 41.



▲ **Figure 41.** *Testing a new frequency-to-colour system using triggers.*



▲ **Figure 42.** *Trigger mapping frequency-to-colour values.*

Remapping Colour Values

While this system could accurately detect changes in pitch, there were still some issues regarding the selection of colours being triggered by frequency ranges. Currently, when there was no sound occurring, the numerical value mapped to colour was zero because there was no frequency range reaching the defined volume threshold, meaning no numerical values were being output from the CHOP network.

In in the HSV colour space, a numerical value of zero results in a vibrant shade of red. This made it hard for the viewer to visually identify patterns occurring within bird song, as there were no empty spaces in the form to visualise the spacing and timing between sounds.

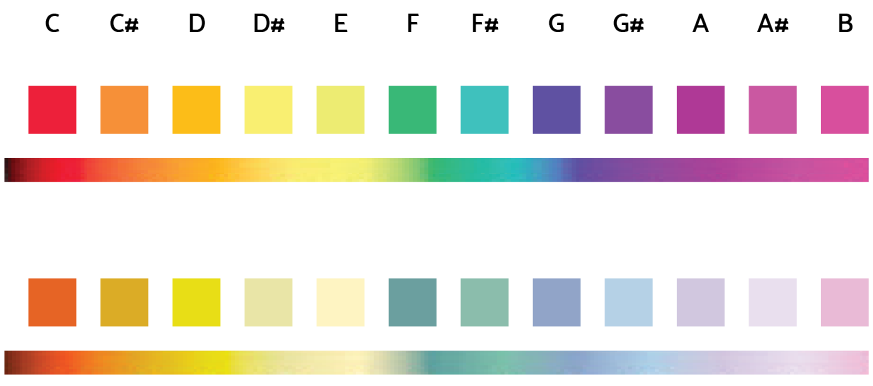
To address this issue, the HSV colour mapping was removed from the system. Instead, the colours were mapped using a ramp which was manually defined by a channel wrangle operator. This allowed for the manual arrangement of colour values. A numerical value of zero could be remapped so that it no longer had a colour association and other frequency ranges could also be remapped to a manually defined colour.

As a numerical value of zero now no longer had a colour association, black particles were now being generated when there was no sound. This was making the audio-visual mapping unclear, as there were no gaps in the form to identify the timing and pattern of bird song. To address this issue, the particles were sorted into groups which allowed for them to be defined by their associated colour value. This meant that when there were no frequencies being triggered, no particles would be generated. This made the audio-visual mapping significantly clearer, as particles would only appear when the corresponding frequency values reached the defined volume threshold of the trigger system.

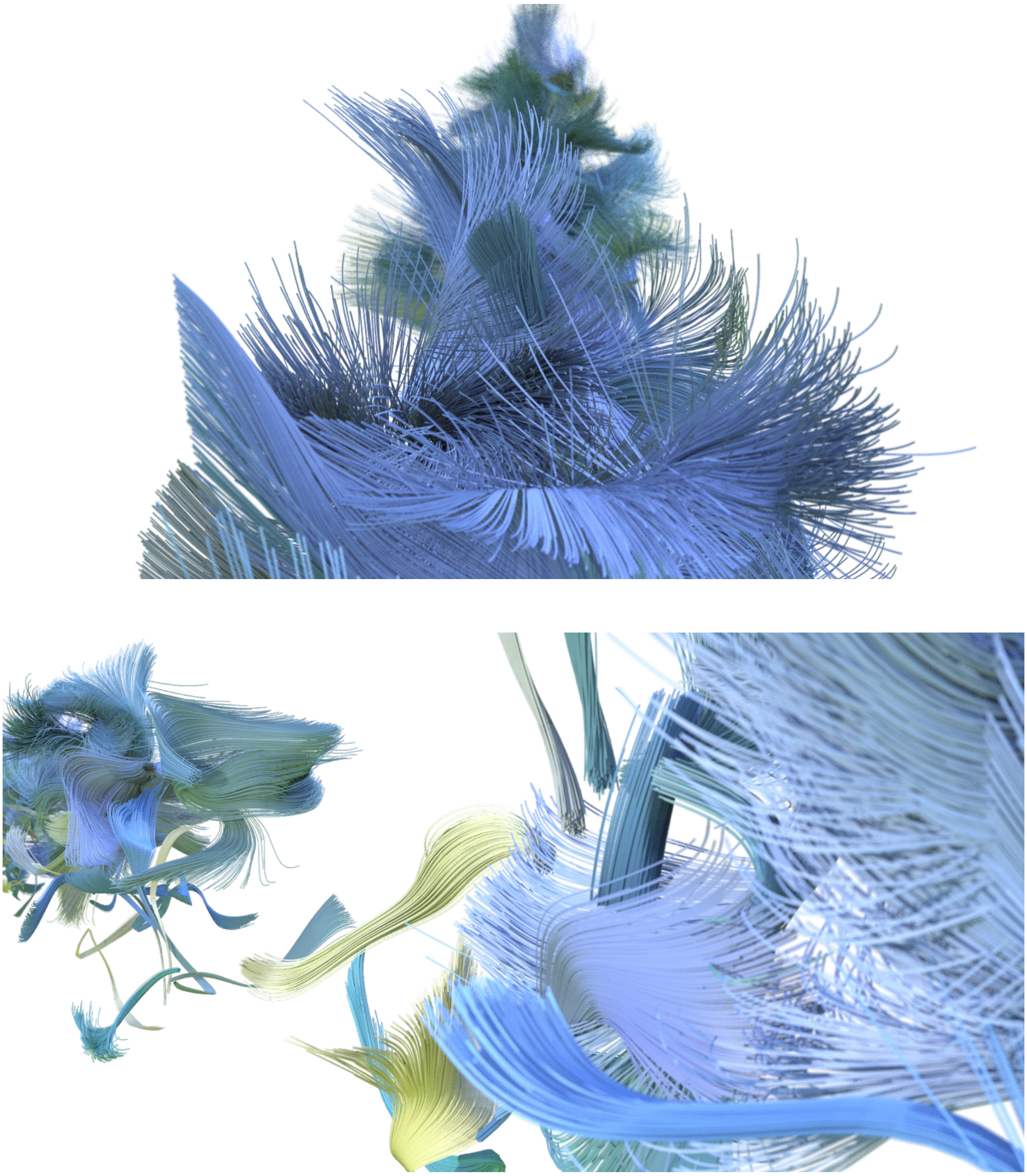
Colour Choices

To develop the visual aesthetic of the design, further consideration was placed into which colours would be used to represent various pitches. The initial colour scheme was based on Isaac Newton's colour theory, which provided a theoretical framework for the perceptual linking of sound to colour. This meant that the colours followed a ROYGBIV pattern to represent the chromatic scale. However, after generating various audio-visualisations of New Zealand bird song, it became clear that they were too synthetic in appearance, and lacked the more organic aesthetic required to create a visual association to New Zealand birds.

Using Newton's pitch-to-colour theory as a starting point, the colours being used for mapping were modified to create a resemblance to the distinct colour characteristics of New Zealand birds. The order of the colours to represent each frequency remained the same, but each hue was adapted to a colour inspired by the feathers of New Zealand birds. This visual embellishment created an aesthetic connection to the data which was being visualised.



▲ **Figure 43.** *Adapted colour mapping system.*

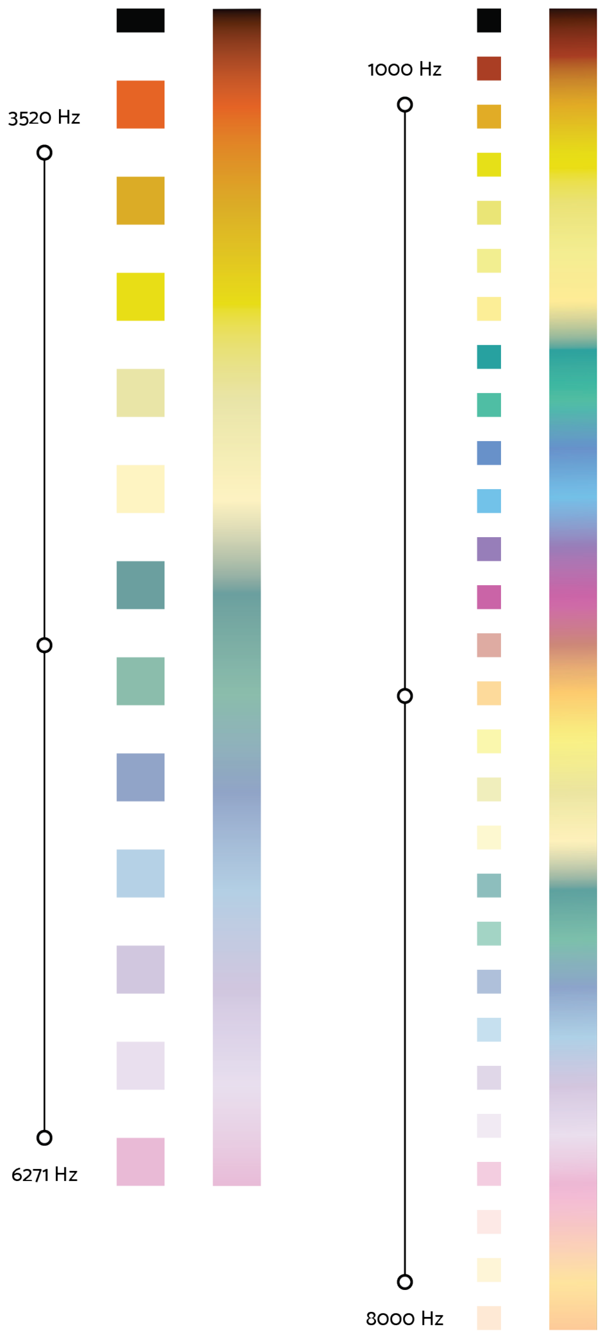


▲ **Figure 44.** *Visualisations of Kea song after changing colour scheme.*

The initial colour-to-pitch mapping system only covered one octave of the chromatic scale, ranging from 3520 Hz - 6271 Hz. The frequencies present within New Zealand bird song had a larger range than what was currently being visualised, so it was important to consider how this information could be represented through colour.

The frequency range of the bird songs being visualised was between 1,000 - 8,000 Hz, so a method for mapping the remaining frequencies at higher and lower ranges had to be considered. As the colour was being mapped to represent what pitch was currently occurring within the bird song, the decision was made to use different shades of the same colour to represent the same pitch value occurring at higher and lower octaves. Background research into perceptual colour association indicated that darker colours are seen to represent lower frequencies whereas lighter colours are seen to represent higher frequencies (Ludwig, Adachi & Matsuzawa, 2011).

To map the lower range of frequencies down to 1,000 Hz, another set of twelve colours was created. These colours were arranged in the same order as ones being used to map higher frequencies, as seen in Figure 45. However, each colour was set to be a shade darker to represent the same pitch occurring at a lower frequency. This meant when the same pitch occurred at different octaves, the hue of the colour would be the same (e.g. the pitch of C would always be visualised using a red hue), but the shade would be brighter or darker depending on the frequency value. A further three colours were added to the higher frequency ranges up to 8,000 Hz, to cover the full range of the bird song being mapped,. At this point, any further addition of frequencies to the colour-to-pitch system would have been for a frequency range outside what was necessary to visualise the data being used.



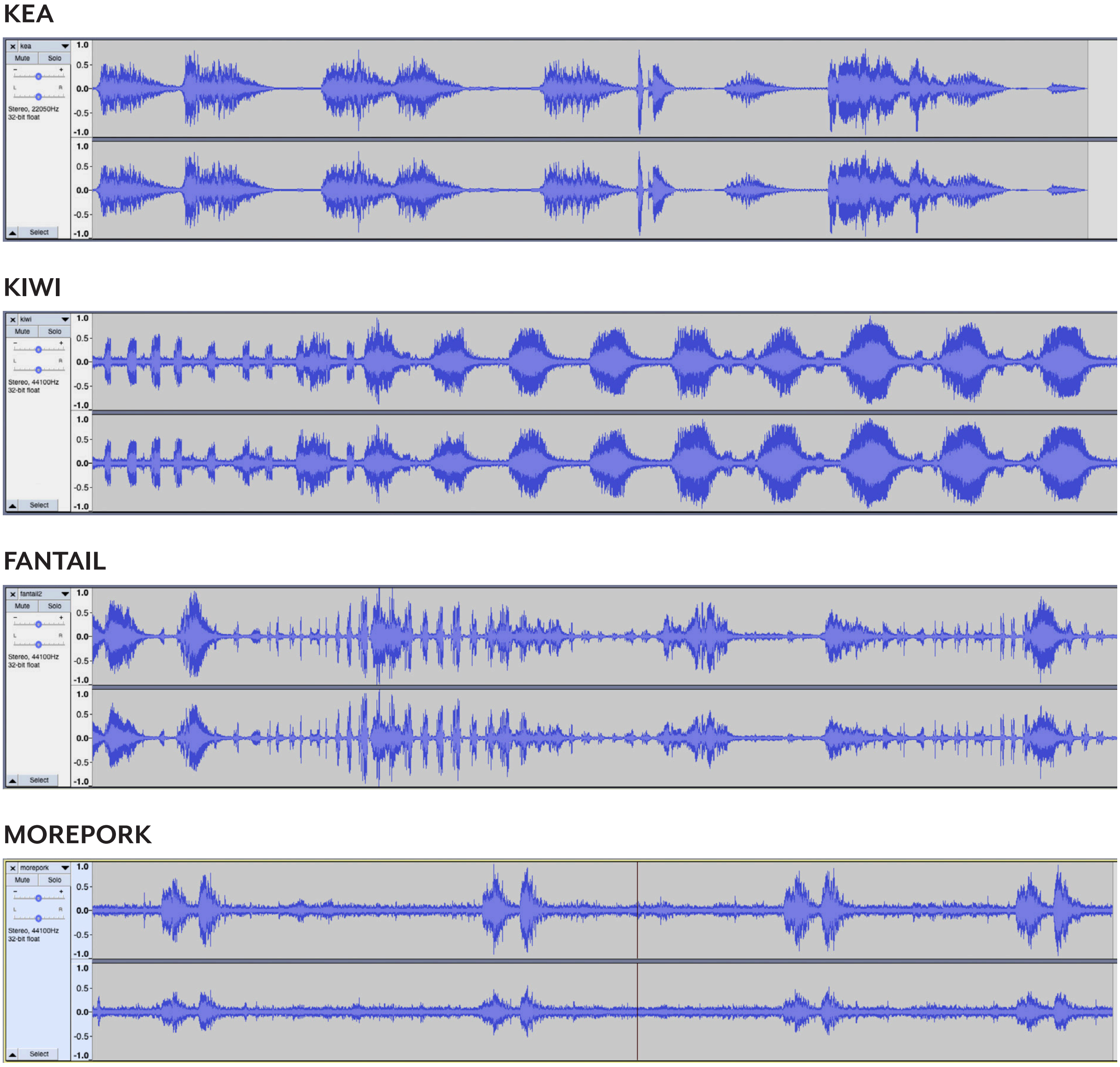
▲ **Figure 45.** *Refined colour scheme with added frequency values.*

Normalising Audio Files

At this stage of the design process, the audio processing component of the design was complete. However, because bird song recordings are often inconsistent in volume, they would sometimes not meet the volume threshold of the trigger system. This resulted in inconsistent audio-visual mapping as it would mean that the audio-visualisation would not be generated.

Ideally, the audio processing pipeline built using Houdini would have been used to adjust audio files to the appropriate volume. However, Houdini's inbuilt normalising functions were not reliable enough to balance out the varying volumes of bird song recordings. Due to this limitation, bird song recordings were normalised to a consistent volume level using an external audio editor. Any background noise was also removed to ensure that the only visual elements generated from the audio file would be relevant to the bird song being visualised.

After normalising the audio files, they were edited into fifteen-second clips to both make the data suitable for visual comparison and reduce the render time required to animate a full clip. Careful consideration was taken to ensure that no relevant sound information was removed from the audio file throughout this process.



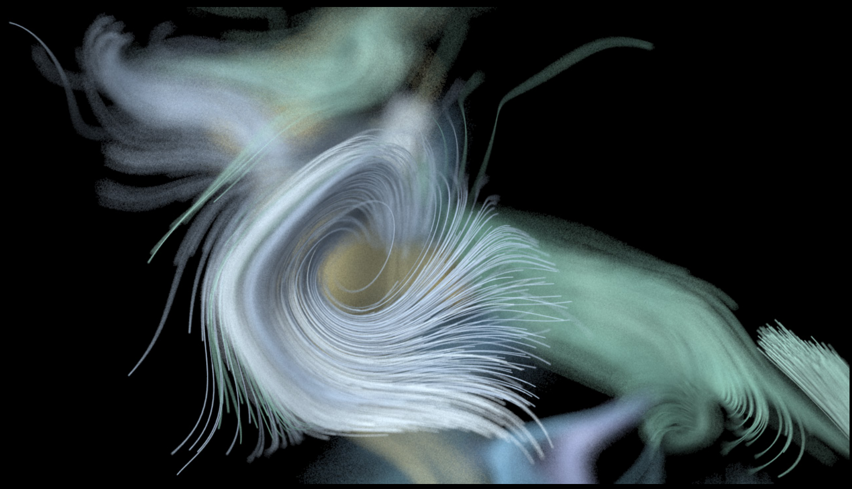
▲ **Figure 46.** *Normalising audio files using Audacity.*

Output

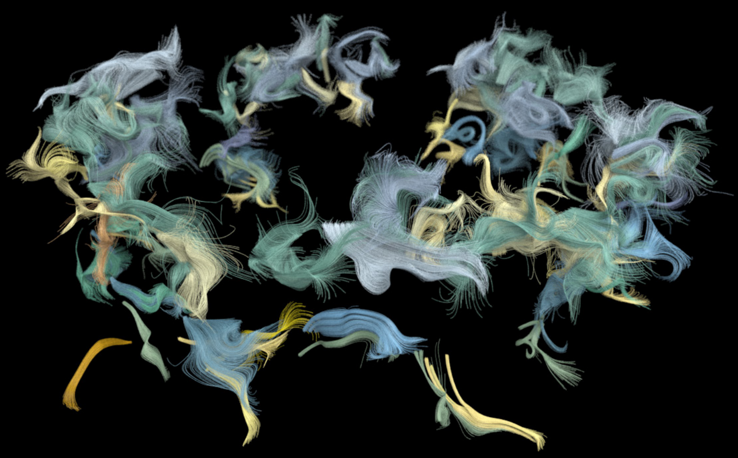
The final phase of the design development was to render the final design output, which would consist of a series of videos and images that contrast and compare various New Zealand bird songs. A multitude of approaches were explored throughout the development of the final output, exploring different camera angles, lighting options, render settings, and compositions.

Examining the generation of the shape from a side angle provided the most readable composition as the position of the form in vertical space could be clearly observed. Other experiments included examining the shape using close-up angles with depth of field, as well as moving it in a circular motion to create a loop. Referring to the criteria helped inform this decision, as viewing the shape from a side angle allowed the viewer to observe changes in pitch, patterns, and timing within the bird song most clearly. A variety of New Zealand bird songs were visualised using the system built within Houdini, which are exhibited and analysed in the following chapters.

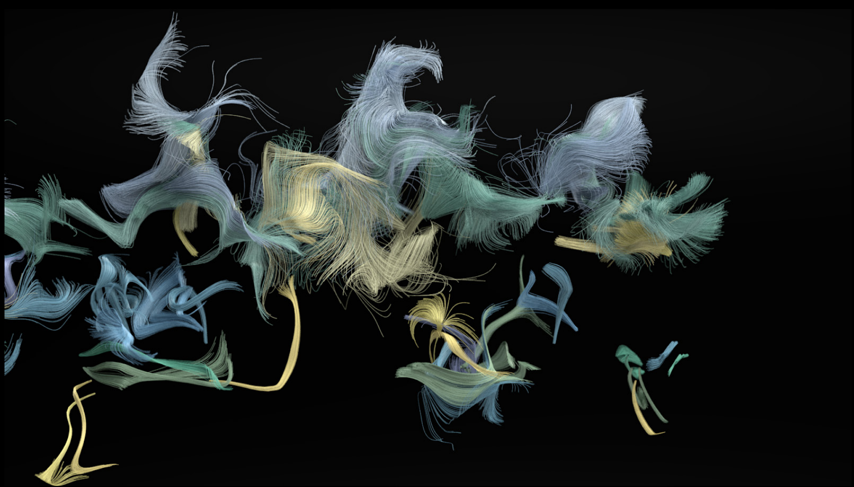
▲ **Figure 47.** *Experimenting with circular movement.*



▲ **Figure 48.** *Camera experiment - close-up detail with depth of field.*



▲ **Figure 49.** *Camera experiment - circular movement with depth of field.*



▲ **Figure 50.** *Camera experiment - side angle tracking movement.*

EVALUATION

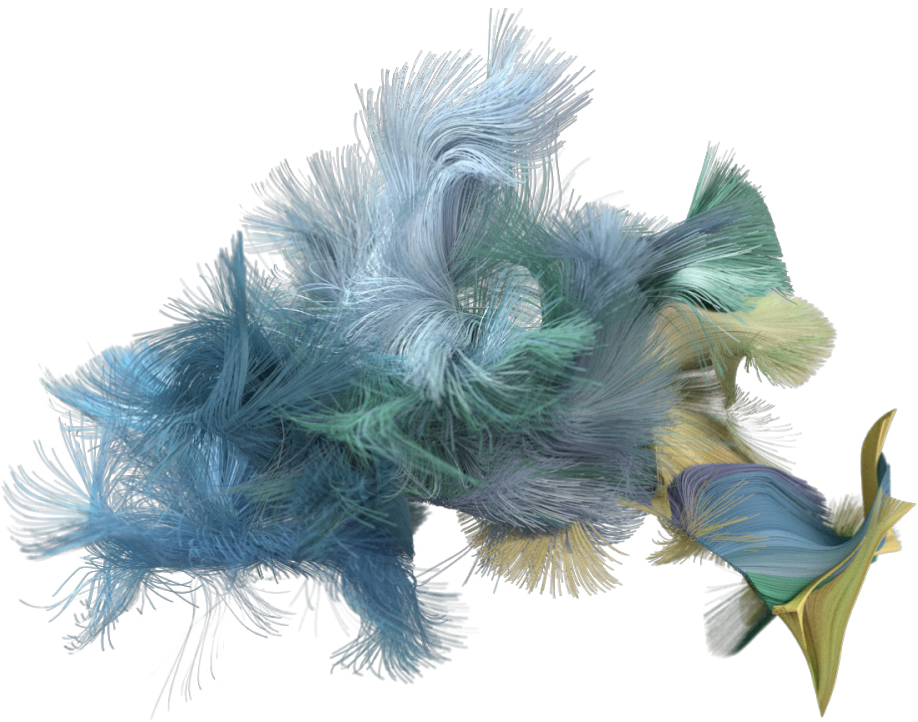
The switch to Houdini proved to be a pivotal moment in the design process, as it allowed for more defined mapping decisions. The design development of a frequency-to-colour system within Houdini was complex, but fulfilled the criteria of the system being able to identify changes in pitch and representing the specific frequency range of New Zealand bird song. There were some limitations, such as the system not being entirely accurate due to the complex overtones which exist among recorded environment sounds. However, the final result was successful in detecting changes in frequency and reflecting these changes through the perceptual mapping of form, colour, movement, and position in space.

The creation of a 3D trail system was another key feature of the design development within Houdini. They allowed for the identification of pattern and timing within the birdsong, and communicated this information to the viewer through the integration of human's phonaesthetic understanding of audio-visual connections. While the audio-visualisations can be considered abstract audio-visual forms, they also provide a visual reference to data being visualised through the shape of the forms as well as their hair or feather-like texture.

A major limitation of Houdini was that the prototyping process was slow due to the amount of sound information being processed and transformed into geometry. Earlier experiments using Touch Designer were identified as concepts which needed further development, which was largely beneficial as it allowed for more time to be spent improving the audio-visual mapping parameters. Houdini also allowed for the system's visual output to be entirely 3D and able to be examined from multiple angles, therefore meeting the criteria and improving upon the limitations of earlier design phases.

One of the most successful aspects using Houdini was that bird songs could be easily visually compared. As the system uses sound data as a design tool to generate the audio visualisations, this simply required the swapping of an audio file. This allowed for the effectiveness of the system to be measured using a variety of bird songs throughout the design process.

Following the design development, recordings were selected from both The Department of Conservation (n.d.) and McPherson (n.d.) to exhibit the audio-visual design system. These included the songs of Tui, Kiwi, Kaka, Kea, Bellbird, Fantail, Kōkako, and Morepork. These bird songs were chosen as they provided a range of contrasting sound data, with this selection of songs containing a variety in both pattern and frequency. These bird songs would be used to measure the success of the final design system by measuring the final output against the developed criteria.



▲ **Figure 51.** *Kea song render close-up.*

VI

CHAPTER SIX

FINAL DESIGN SYSTEM

▲ **Figure 52.** *Circular bird song visualisations - compilation overlay.*

OVERVIEW

The final design system uses sound data from New Zealand birds to generate audio-visual artworks which attempt to express unique elements of each song. The system works by detecting the frequency of a bird song over time and assigning a defined colour value to the shapes being generated. The colour scheme is based upon Isaac Newton's pitch-to-colour theory, and also integrates the mel scale system for estimating pitch from frequency. The colours are adapted from their original shades, to match the shades of colour found amongst the feathers of native New Zealand birds.

The subtle changes in frequency which occur among the bird song can be observed visually as the shapes generate across the screen from left to right. The generated shapes vary in form depending upon the value of the frequency, with higher frequencies generating spikier, expanded shapes and lower frequencies generating smoother, condensed shapes. This links to the phonaesthetic connections humans make between sonic and visual elements such as the maluma/takete phenomenon outlined by Styles and Gawne (2017). The position on the shape in vertical space is also dictated by the value of the frequency, with a frequency value of around 8,000 Hz being positioned at the top of the screen, and a frequency value of around 1,000 Hz being positioned at the bottom. Frequencies in between these values are also positioned appropriately. For example, a frequency value of 4,500 Hz would appear in the middle of the screen.

A total of eight New Zealand bird songs were chosen to observe the visual variance which occurs when comparing different bird calls. As well as general identifiable songs, various types of calls from the same species were compared, such as male and female, territorial and communicative, and species from different locations.

VIDEO WORK

To view the corresponding videos for each bird song visualisation, please refer to the files referenced in the following section. A range of camera angles have been included. While the linear visualisations provide the most legible translation of sound, front-views and close-ups of each bird song have also been included to allow the visualised sound data to be examined from multiple angles. Please note, one audio file was selected for animation from each species of New Zealand bird. This was to allow for a consistent data-set which can be compared through the examination of both the sound and the video. It is recommended to view the video files using Quicktime or VLC. The file structure is as follows:

Synaesthetic-Sanctuary-Animations

Close-ups

- bellbird-closeup.mp4
- fantail-closeup.mp4
- kaka-closeup.mp4
- kea-closeup.mp4
- kea-closeup-2.mp4
- kiwi-closeup.mp4
- kōkako-closeup.mp4
- morepork-closeup.mp4
- tui-closeup.mp4

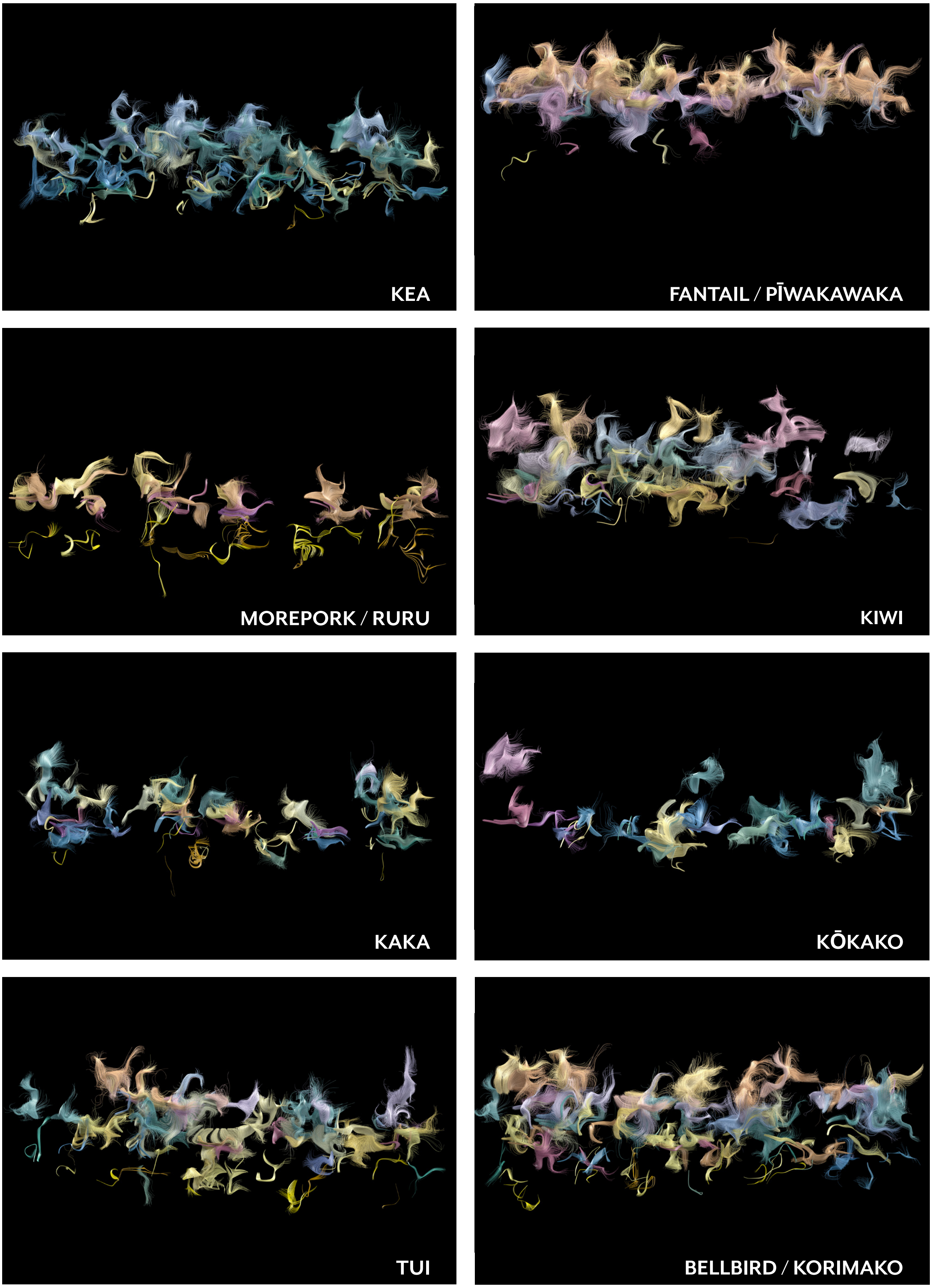
Front-view

- bellbird-frontview.mp4
- fantail-frontview.mp4
- kaka-frontview.mp4
- kea-frontview.mp4
- kiwi-frontview.mp4
- kōkako-frontview.mp4
- morepork-frontview.mp4
- tui-frontview.mp4

Linear

- bellbird-linear.mp4
- fantail-linear.mp4
- kaka-linear.mp4
- kea-linear.mp4
- kiwi-linear.mp4
- kōkako-linear.mp4
- morepork-linear.mp4
- tui-linear.mp4

▼ **Figure 53.** *Bird song comparisons.*



KEA

Scientific name : Nestor notabilis

Location : Forested and Alpine regions of the South Island of New Zealand

Population : 3000-7000 individuals

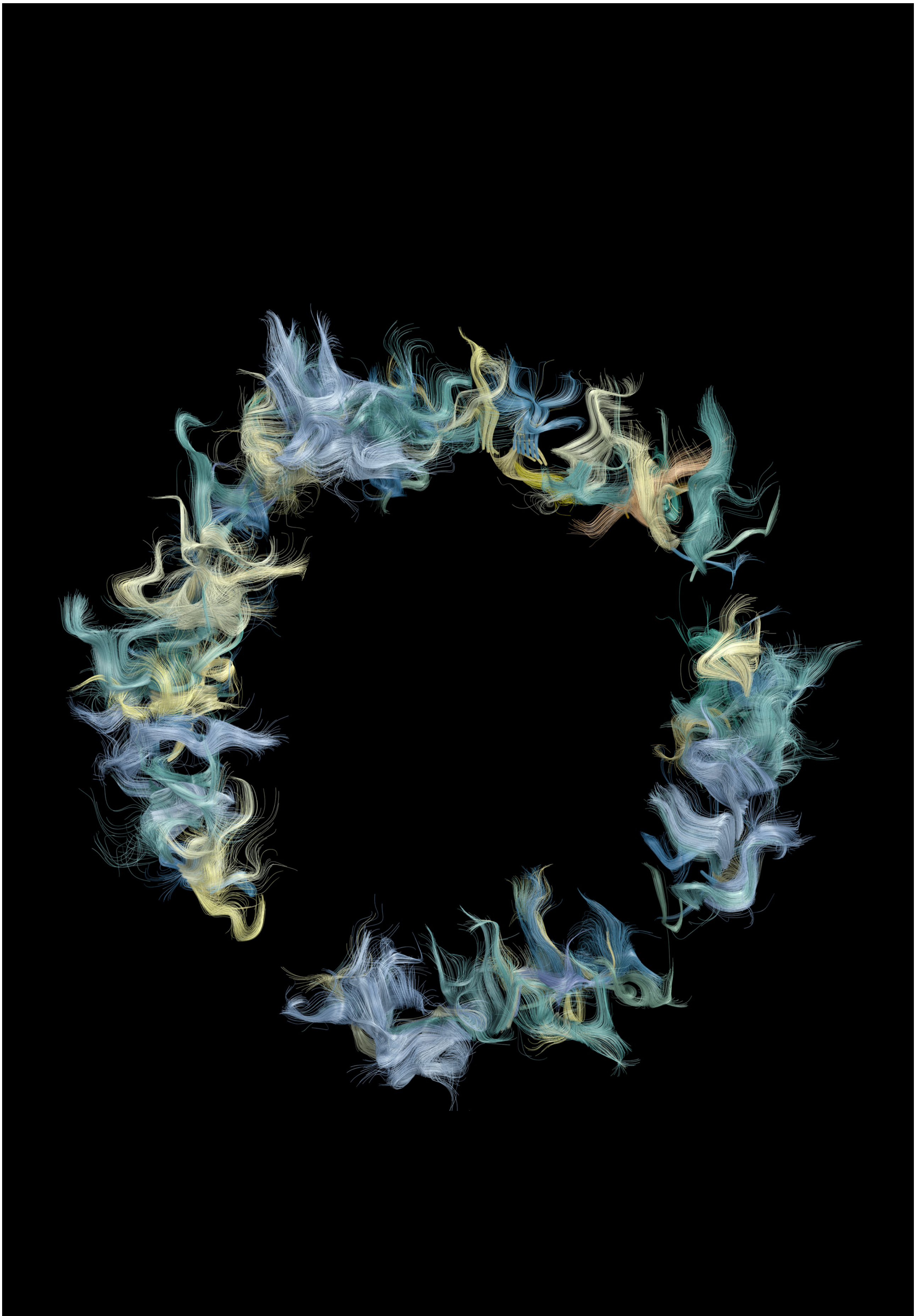
Conservation Status : Threatened - Nationally endangered

Mass : 460g

The most commonly heard Kea call consists of a long, loud, high-pitched descending cry which may be either continuous or broken into segments (Kemp, 2013).

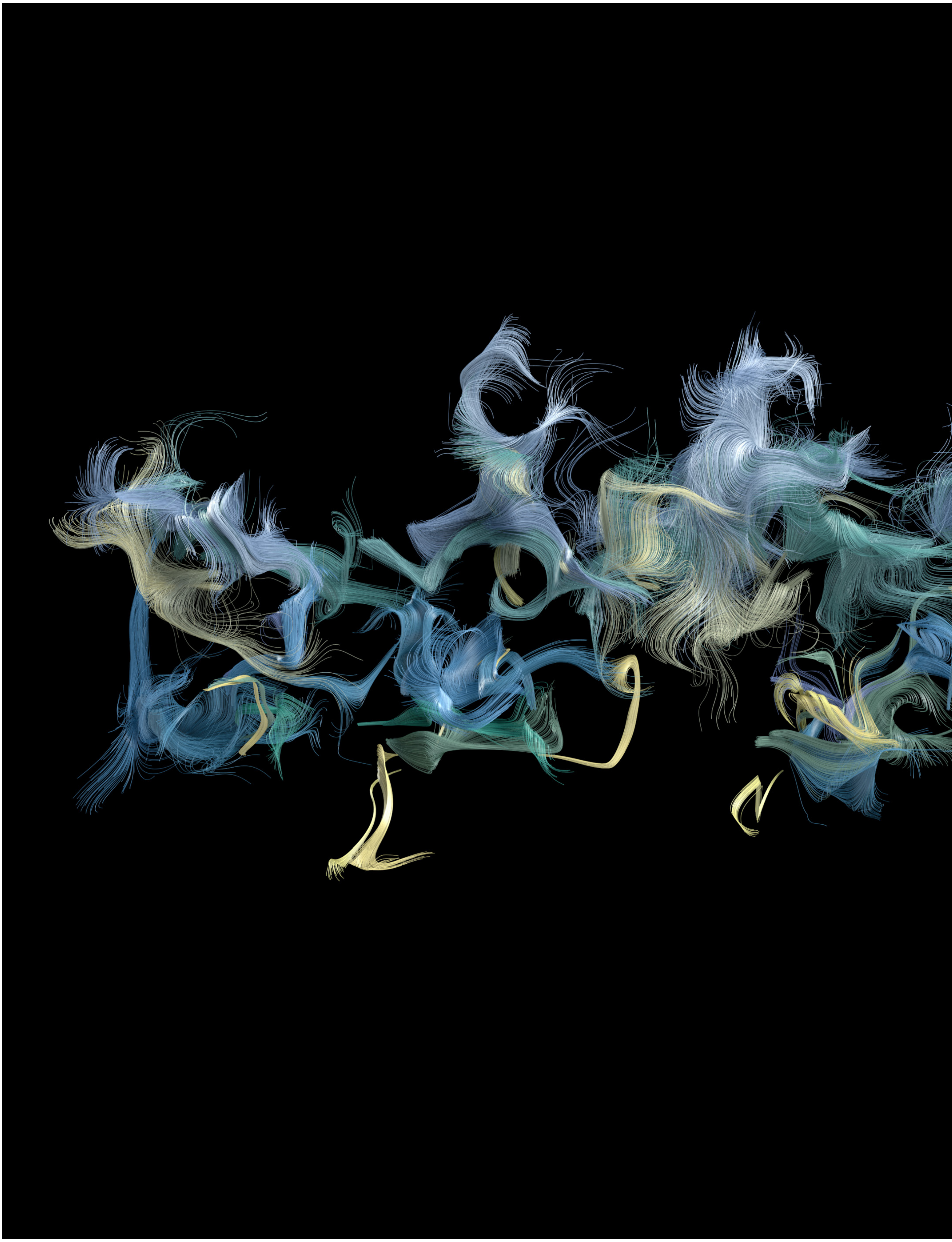
▲ **Figure 54.** *Kea, Johannes Keulemans - 1872.*

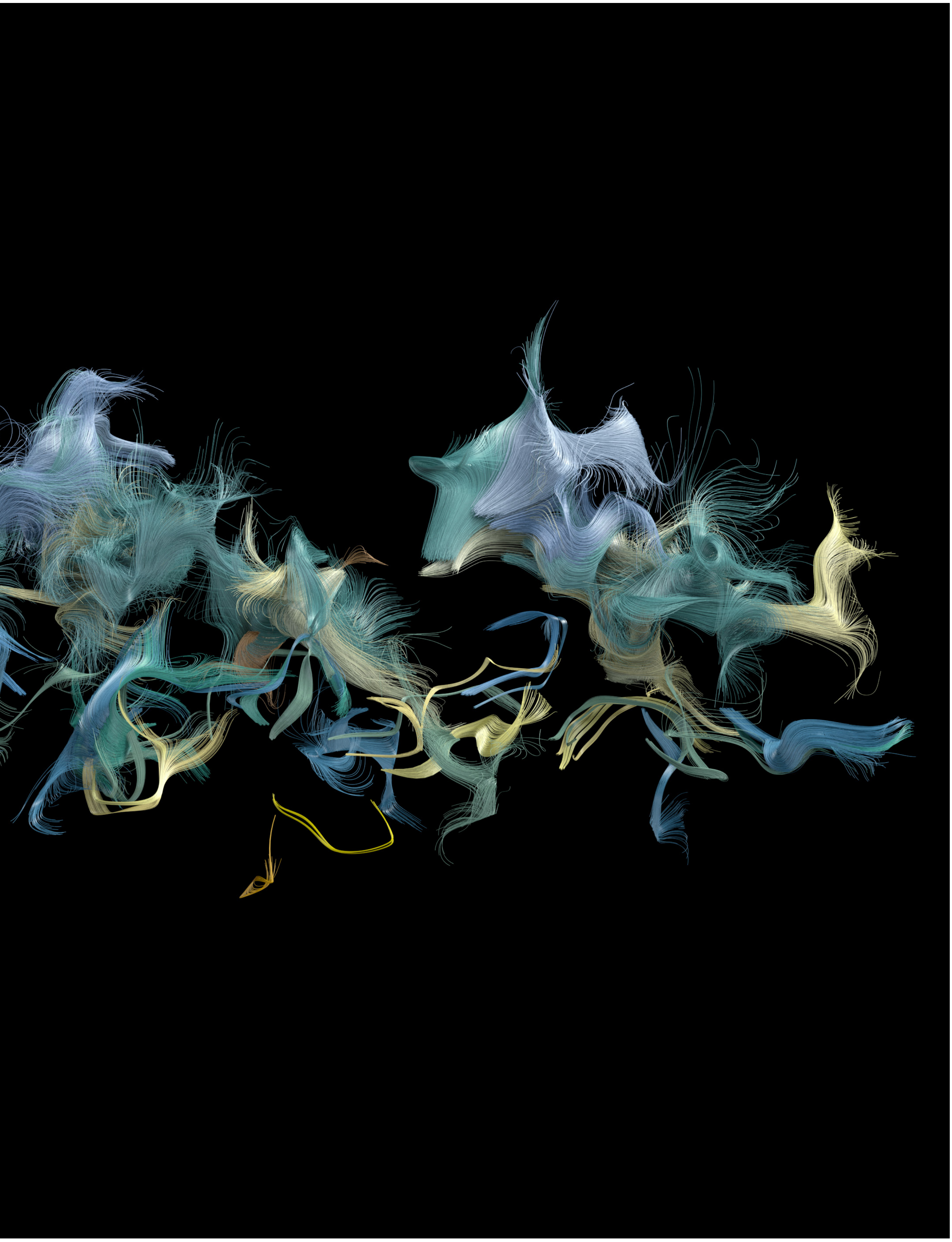
Note. From *A History of New Zealand Birds* by J. Keulemans & W. Buller, 1888, London, United Kingdom: (n.p.). In the public domain.



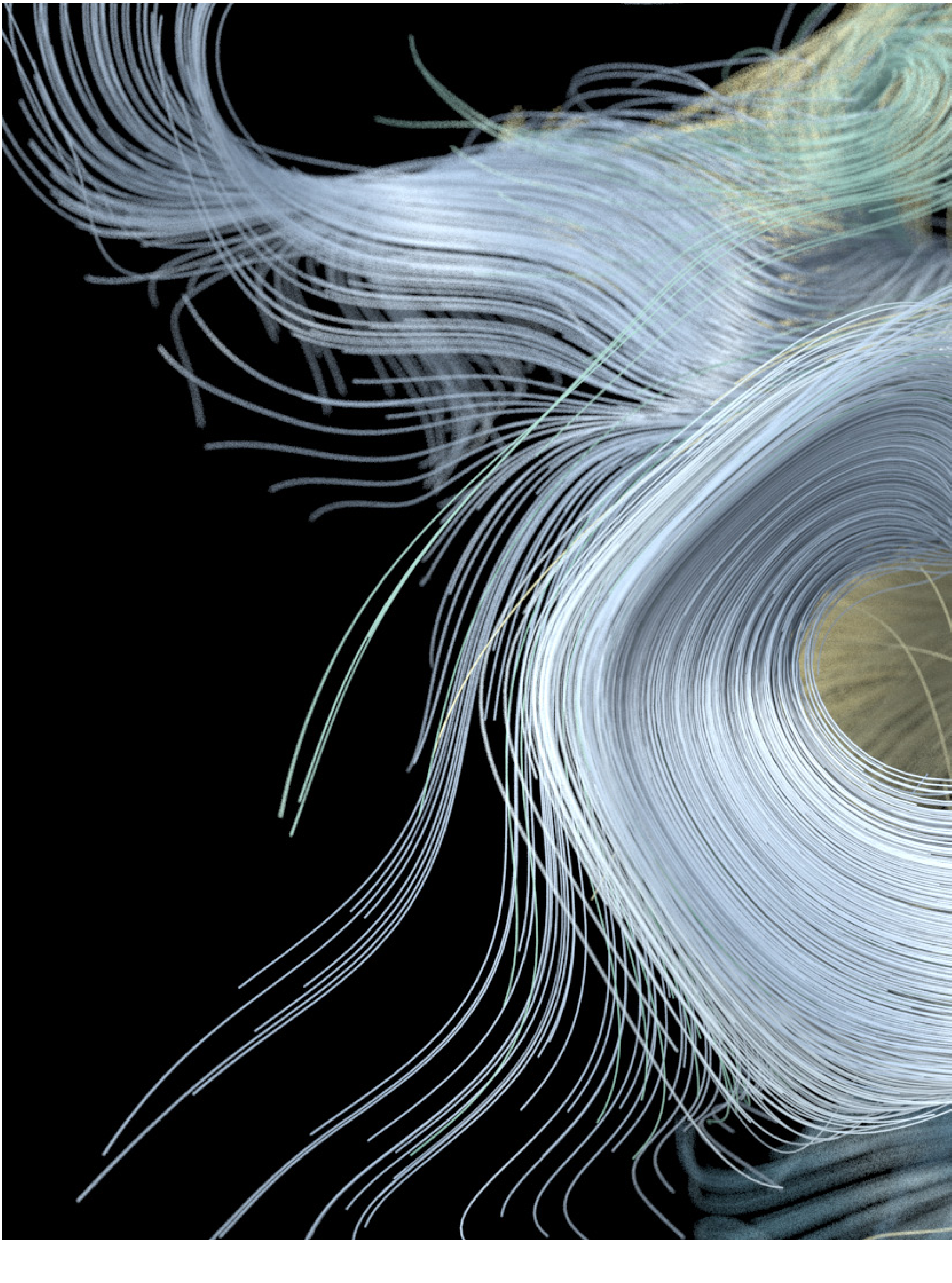
▲ **Figure 55.** *Kea song - circular.*

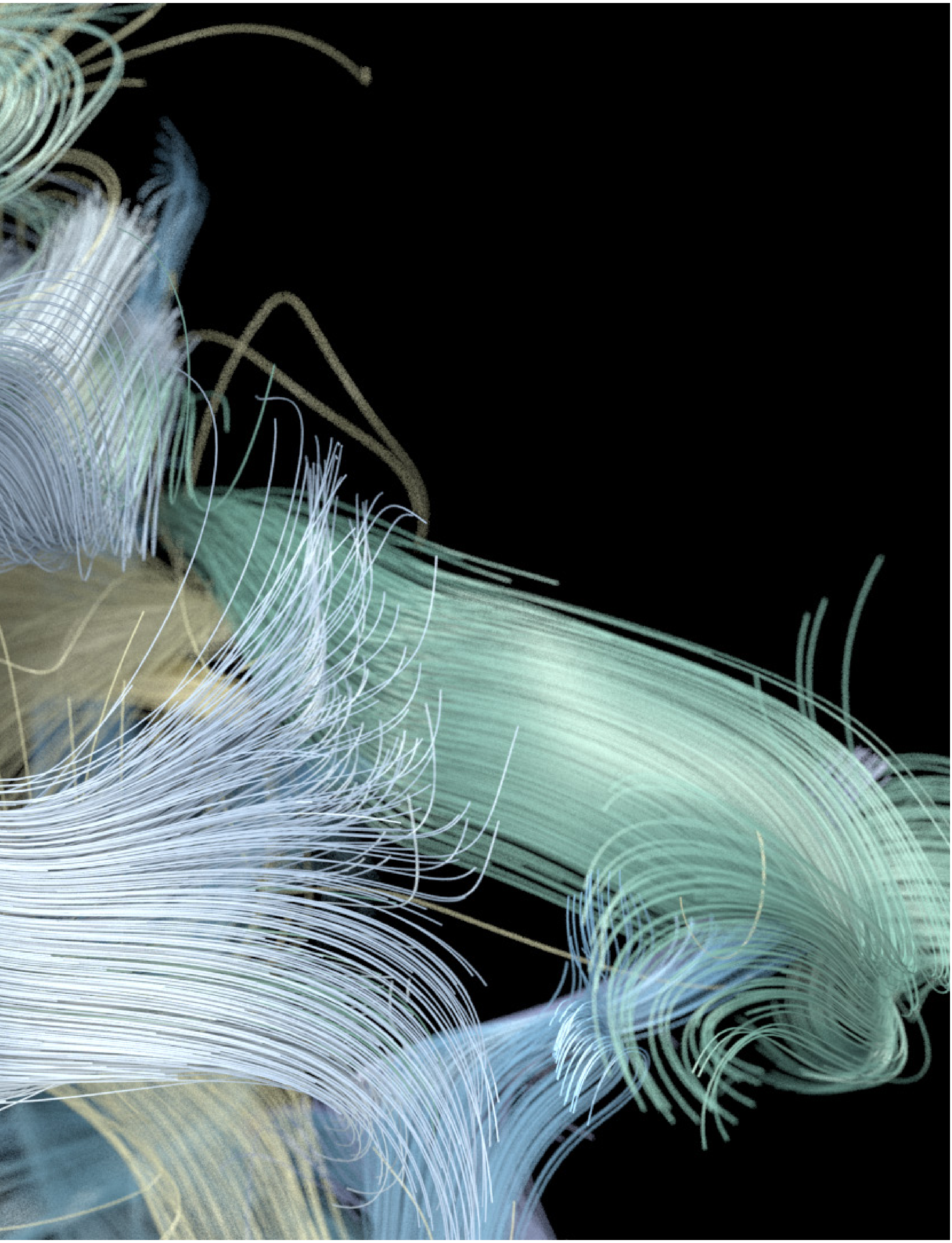
▼ **Figure 56.** *Kea song - linear.*

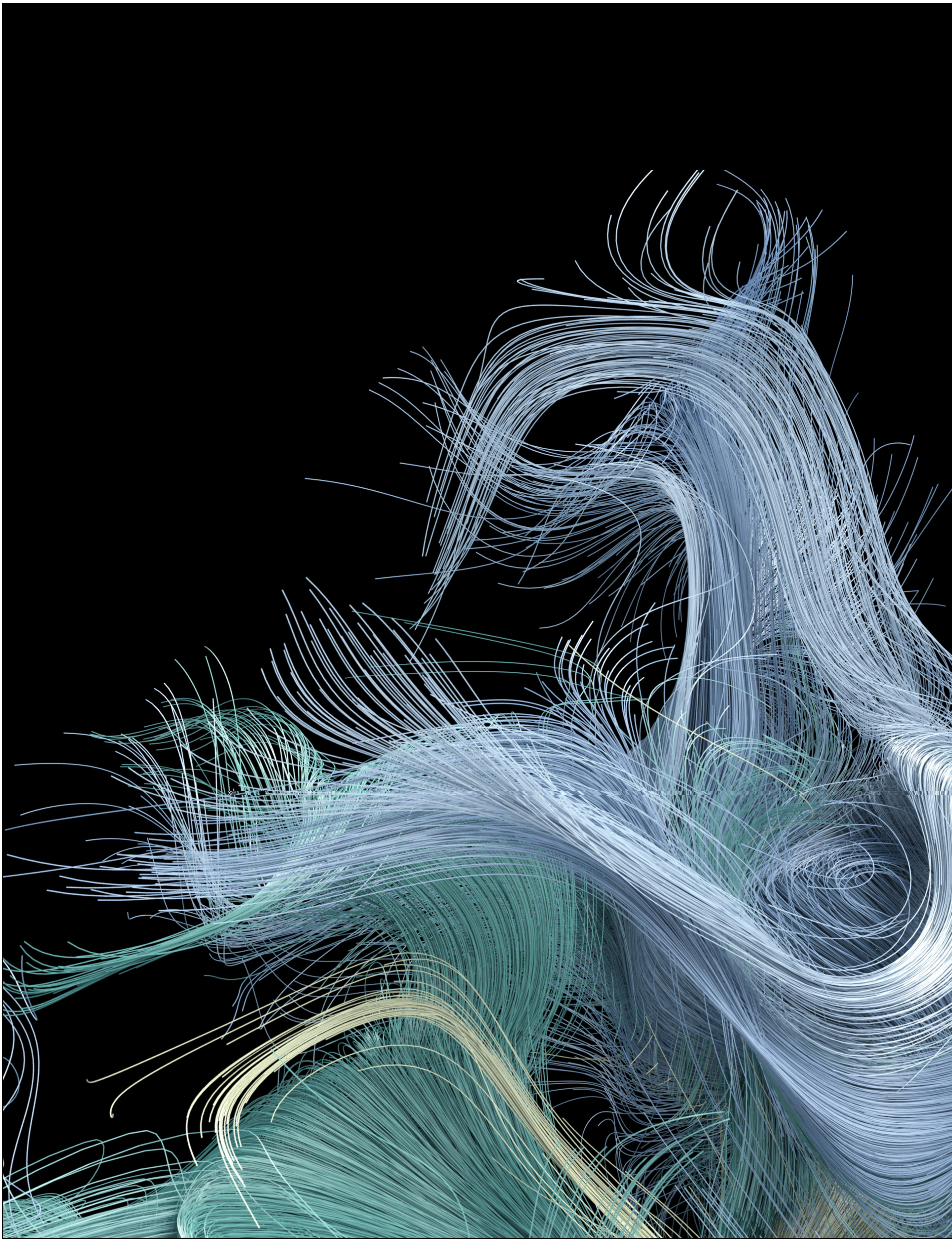




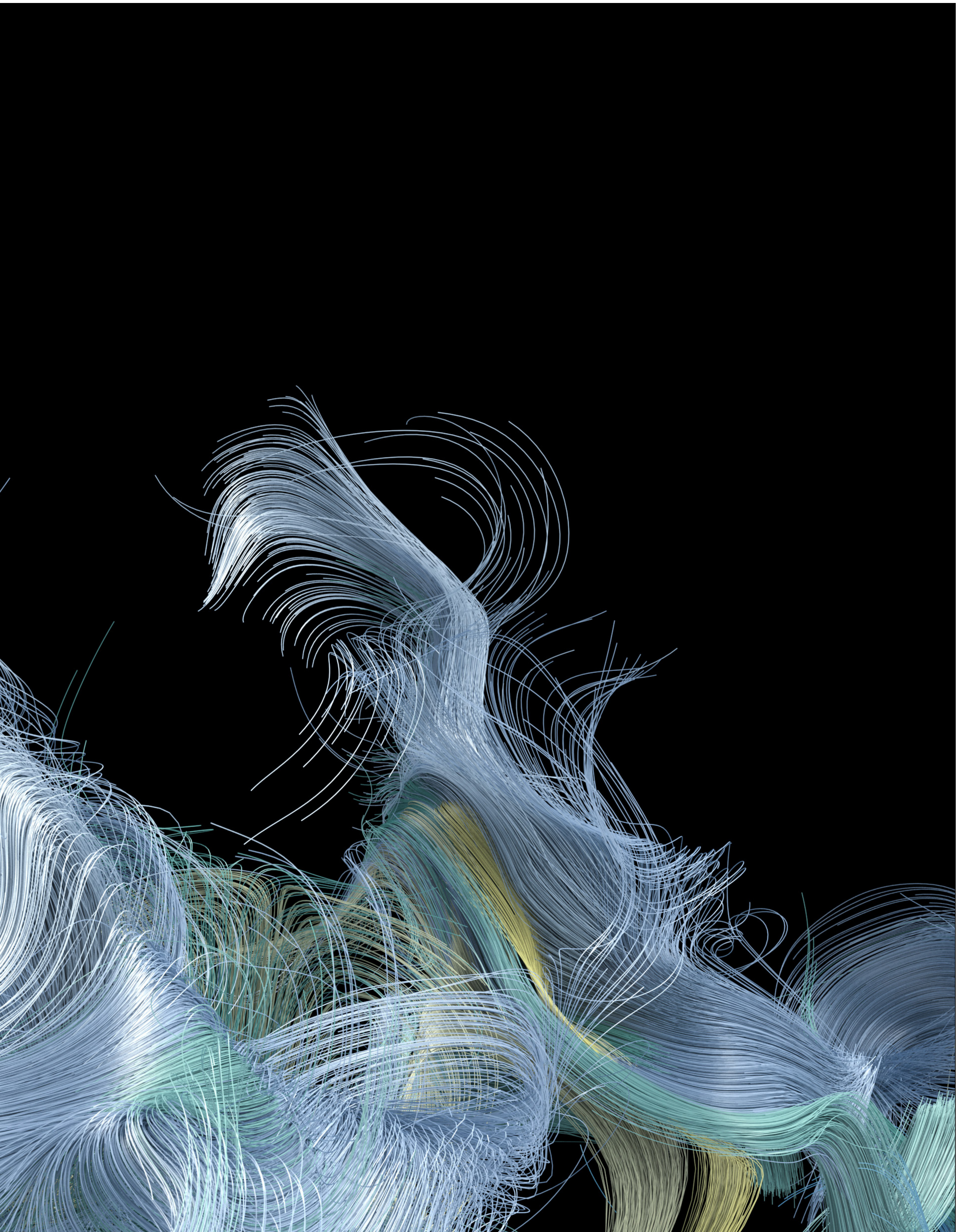
▼ **Figure 57.** *Kea song - detail 01.*







▼ **Figure 58.** *Kea song - detail 02.*



MOREPORK / RURU

Scientific name : Ninox novaeseelandiae

Location : Forests throughout mainland New Zealand and offshore islands

Population : Widespread and moderately abundant

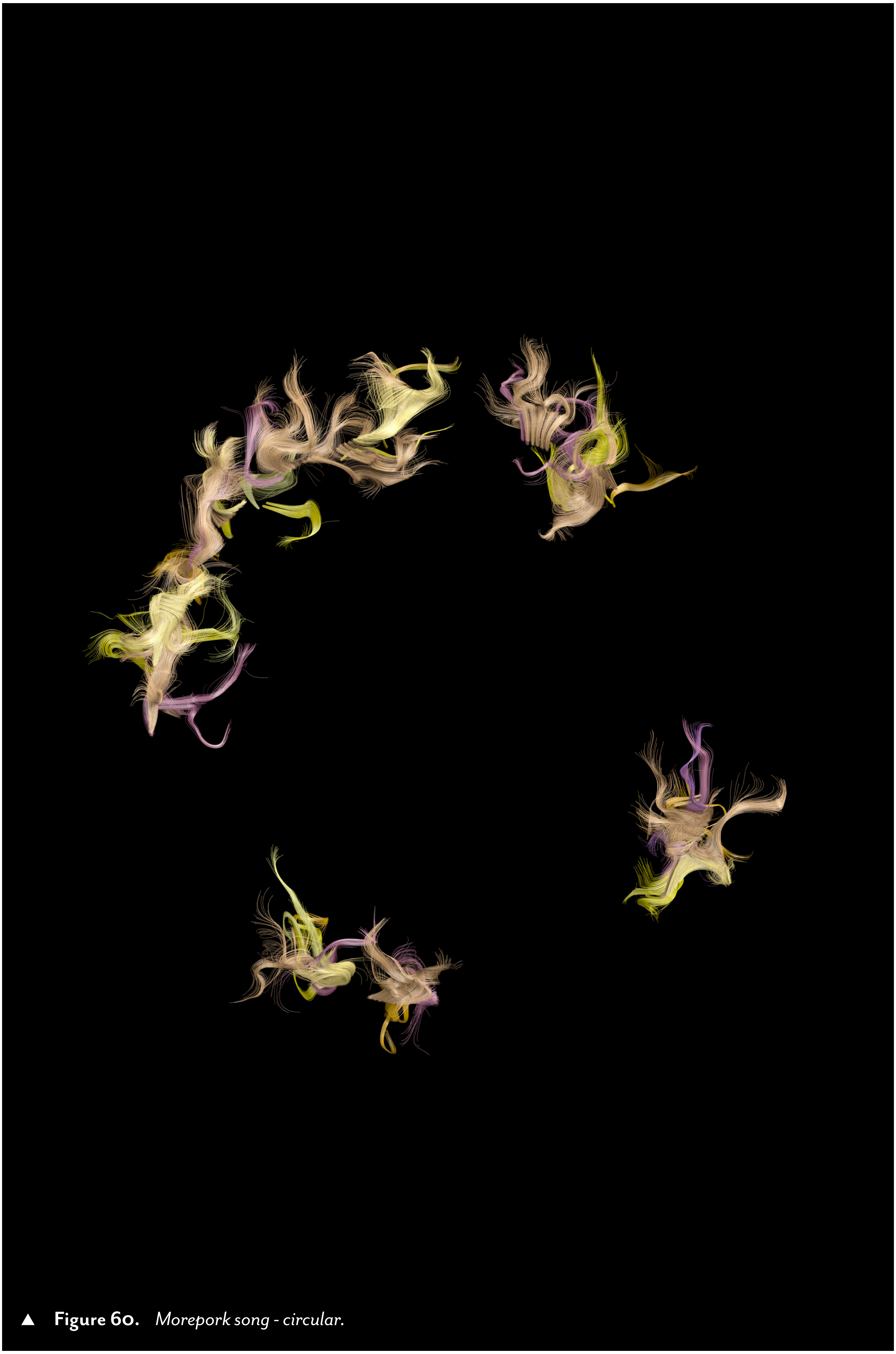
Conservation Status : Not threatened

Mass : 175g

The song of the Morepork occurs only at night. The onomatopoeic "more-pork" noise is the most commonly heard sound of the morepork. Other Morepork songs consist of a "quork quork", or a rising "quee" which is often mistaken for a Kiwi (Seaton & Hyde, 2013).

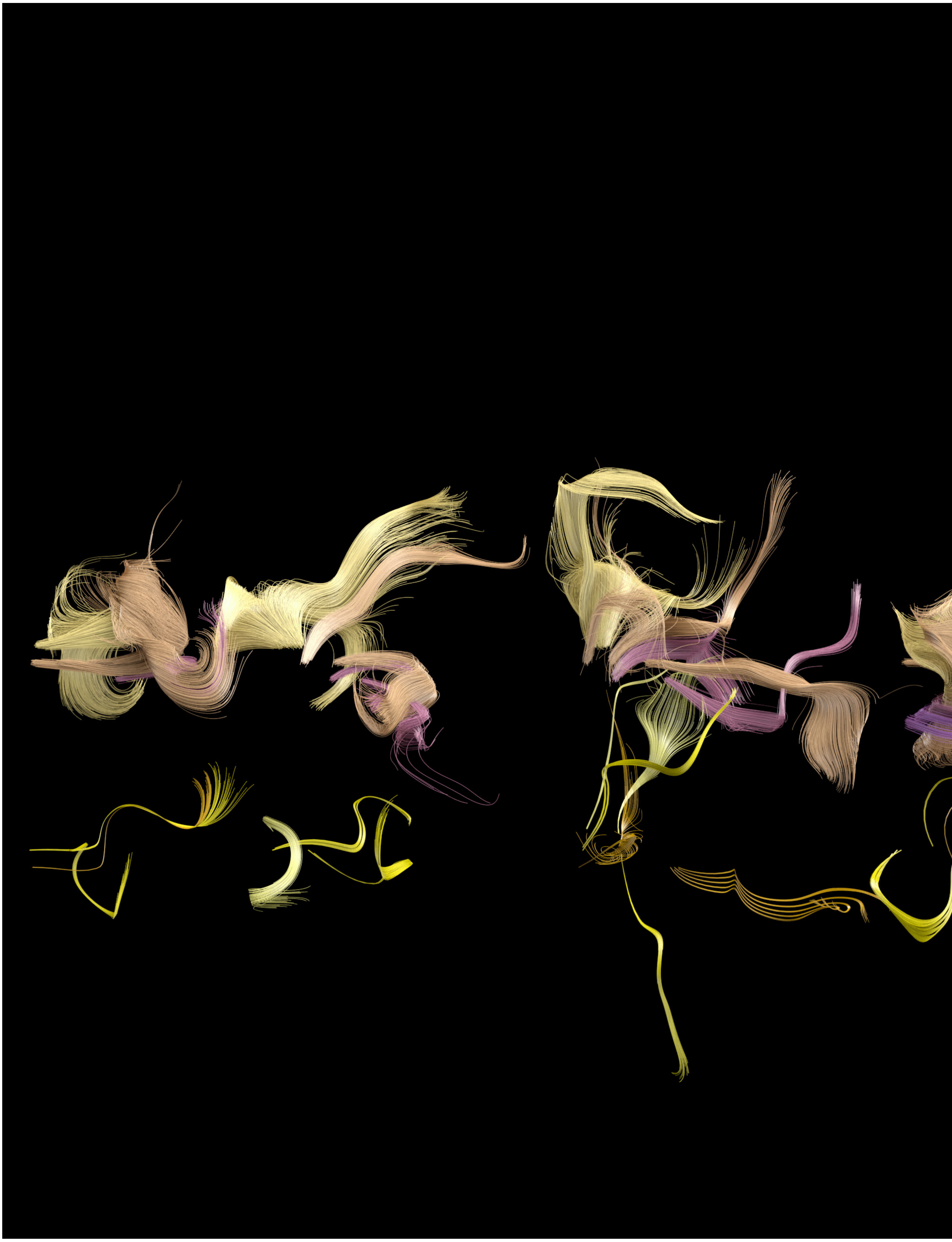
▲ **Figure 59.** *Morepork, Johannes Keulemans - 1872.*

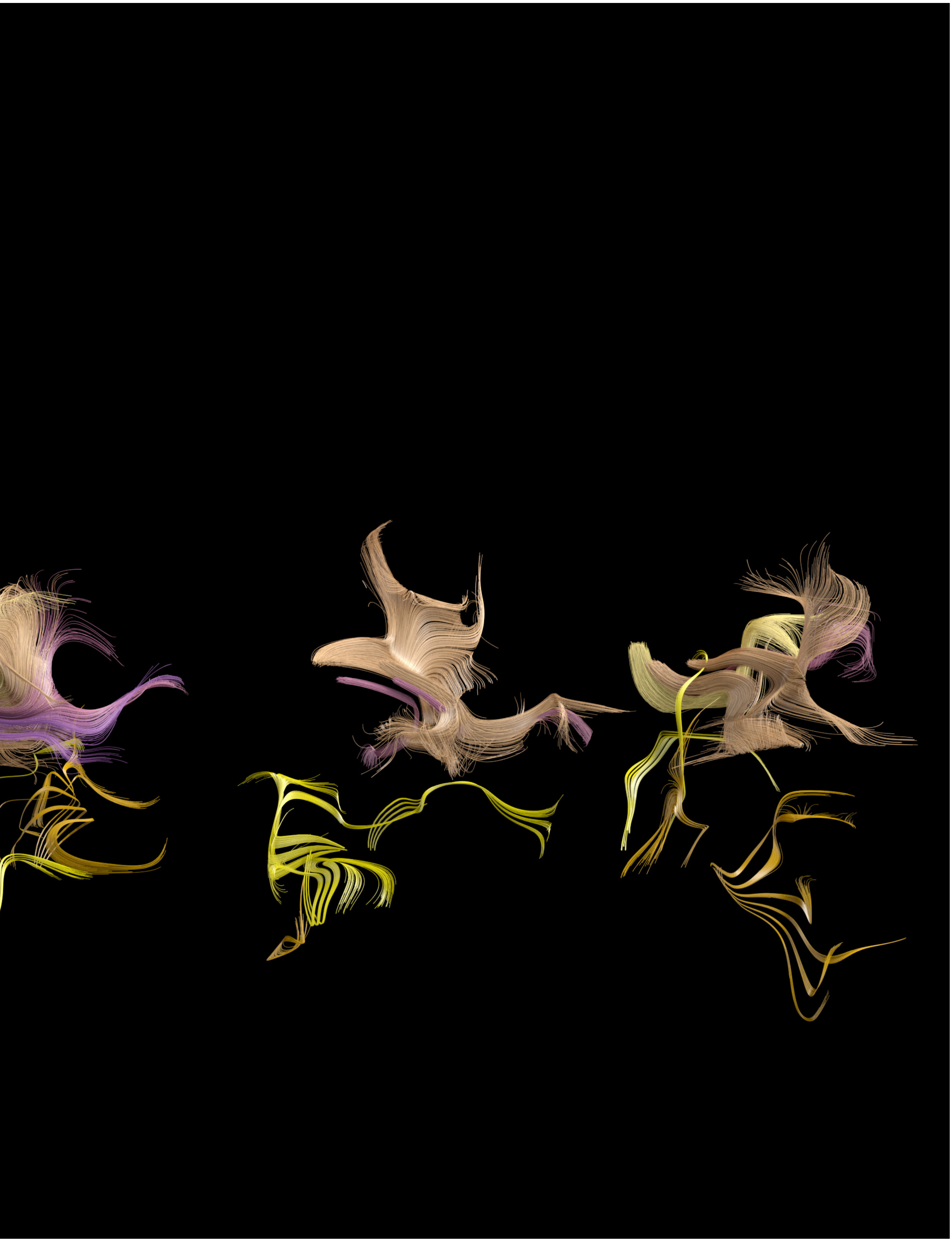
Note. From *A History of New Zealand Birds* by J. Keulemans & W. Buller, 1888, London, United Kingdom: (n.p.). In the public domain.



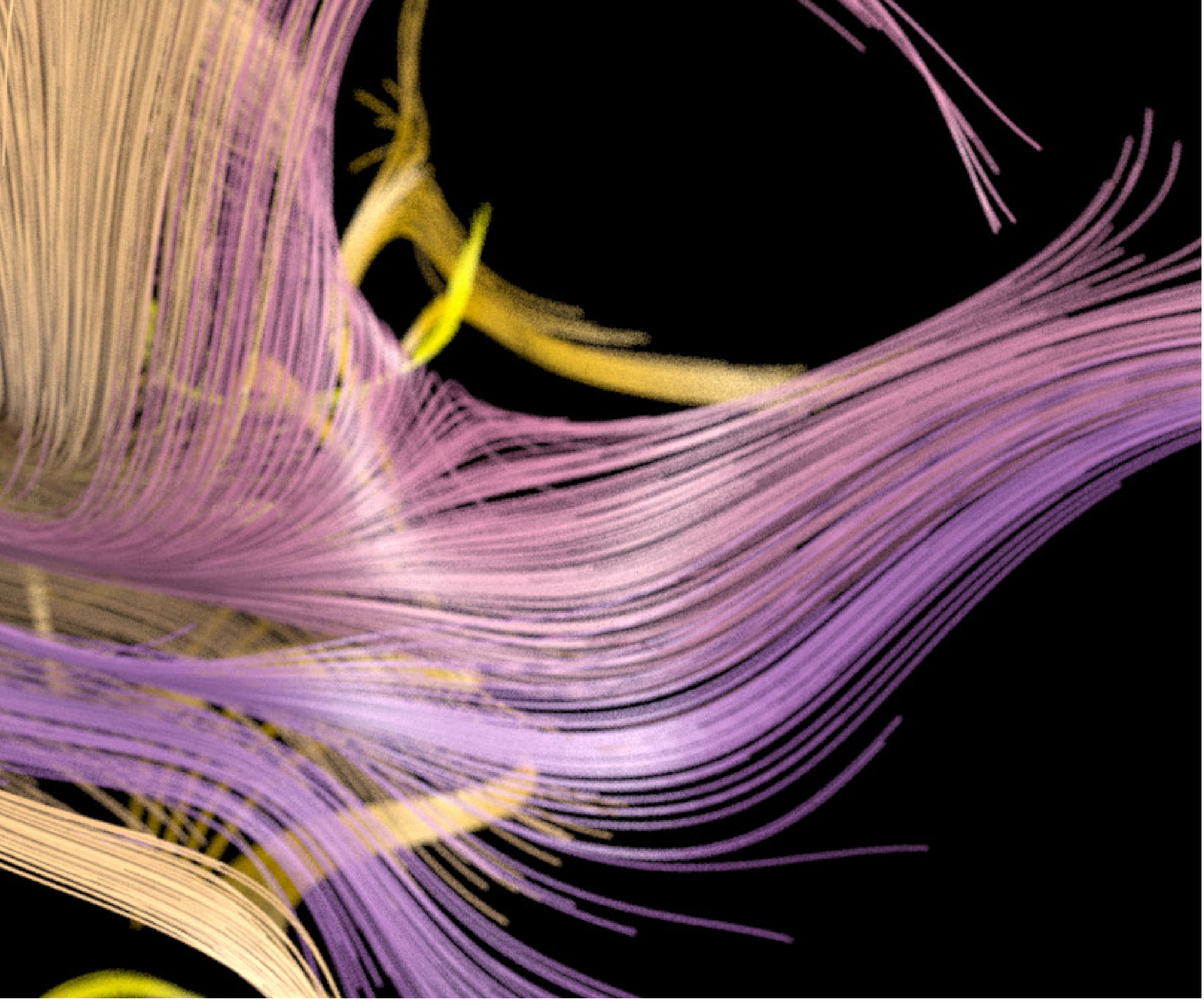
▲ **Figure 60.** *Morepork song - circular.*

▼ **Figure 61.** *Morepork song - linear.*

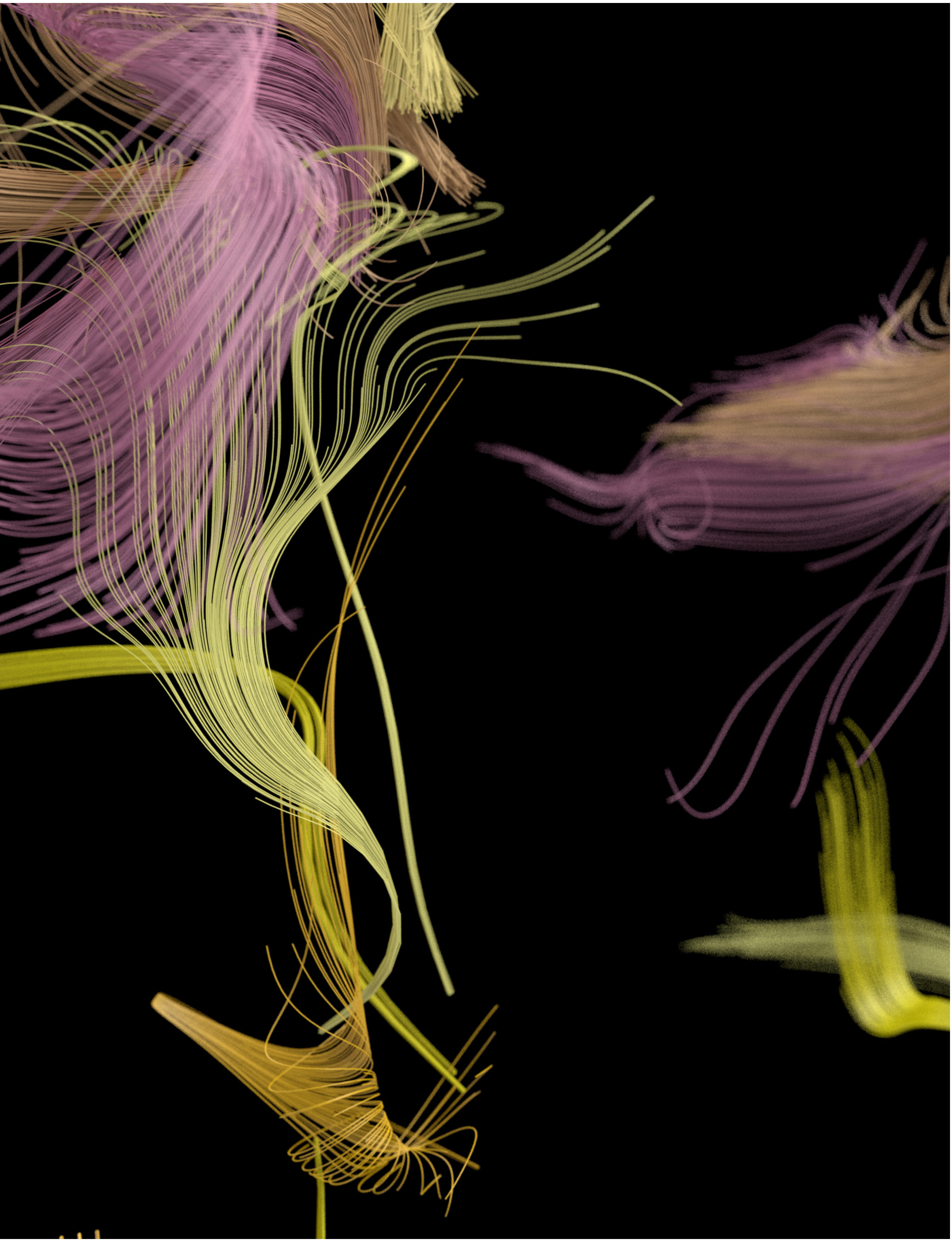




▼ **Figure 62.** *Morepork song - detail 01.*



▼ **Figure 63.** *Morepork song - detail 02.*



KAKA

Scientific name : Nestor meridionalis

Location : Large forested areas in the North and South Islands

Population : Fewer than 10,000 individuals

Conservation Status : North Island - At risk (Recovering)

South Island - Nationally Vulnerable

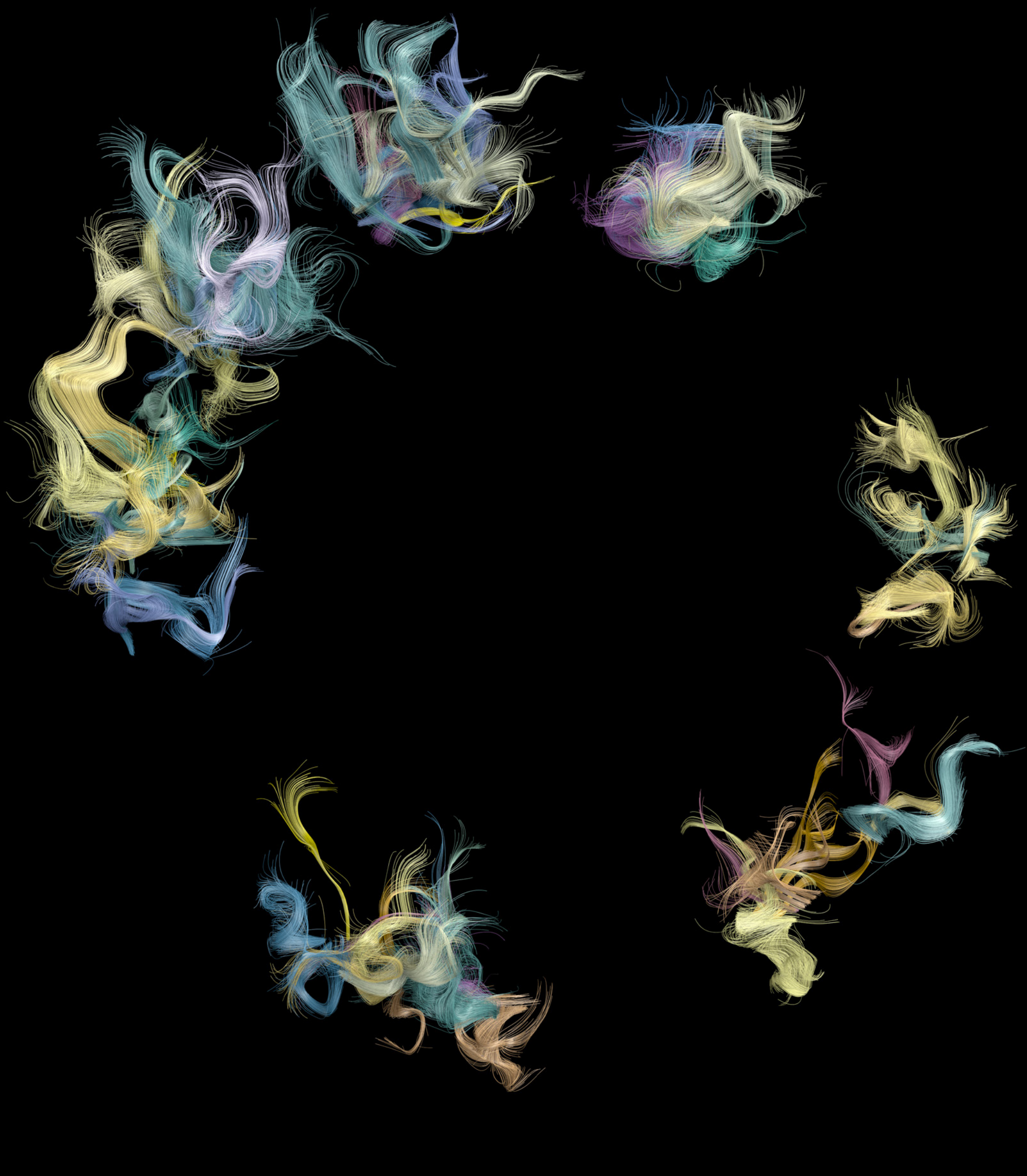
Chatham Islands - Extinct

Mass : 340-400g

The song of the Kaka is similar to that of the Kea, except calls are more varied and may consist of grating "skraarks". Kaka's can also make a range of loud musical whistles which can vary depending on their locations (Moorhouse, 2013).

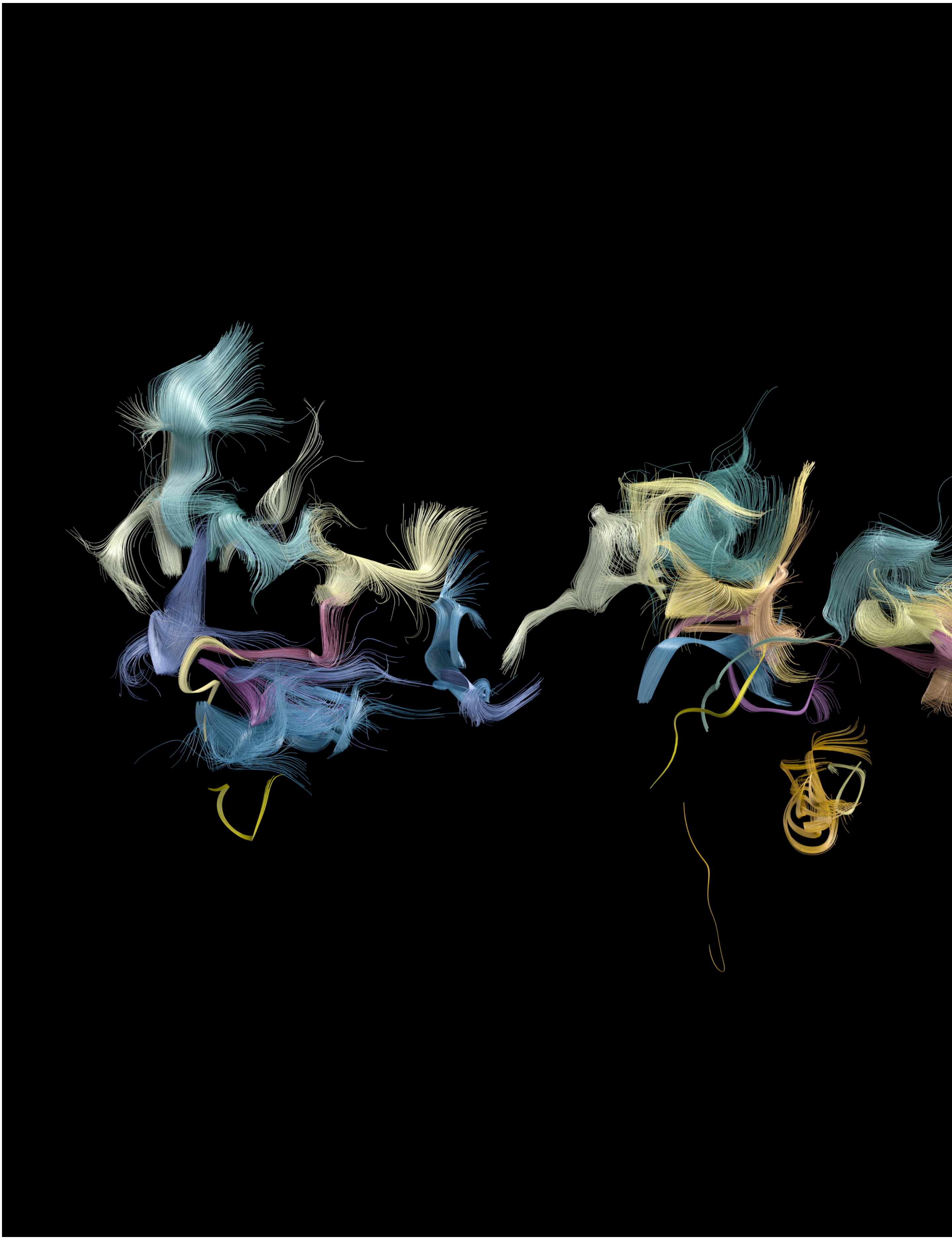
▲ **Figure 64.** *Kaka, Johannes Keulemans - 1888.*

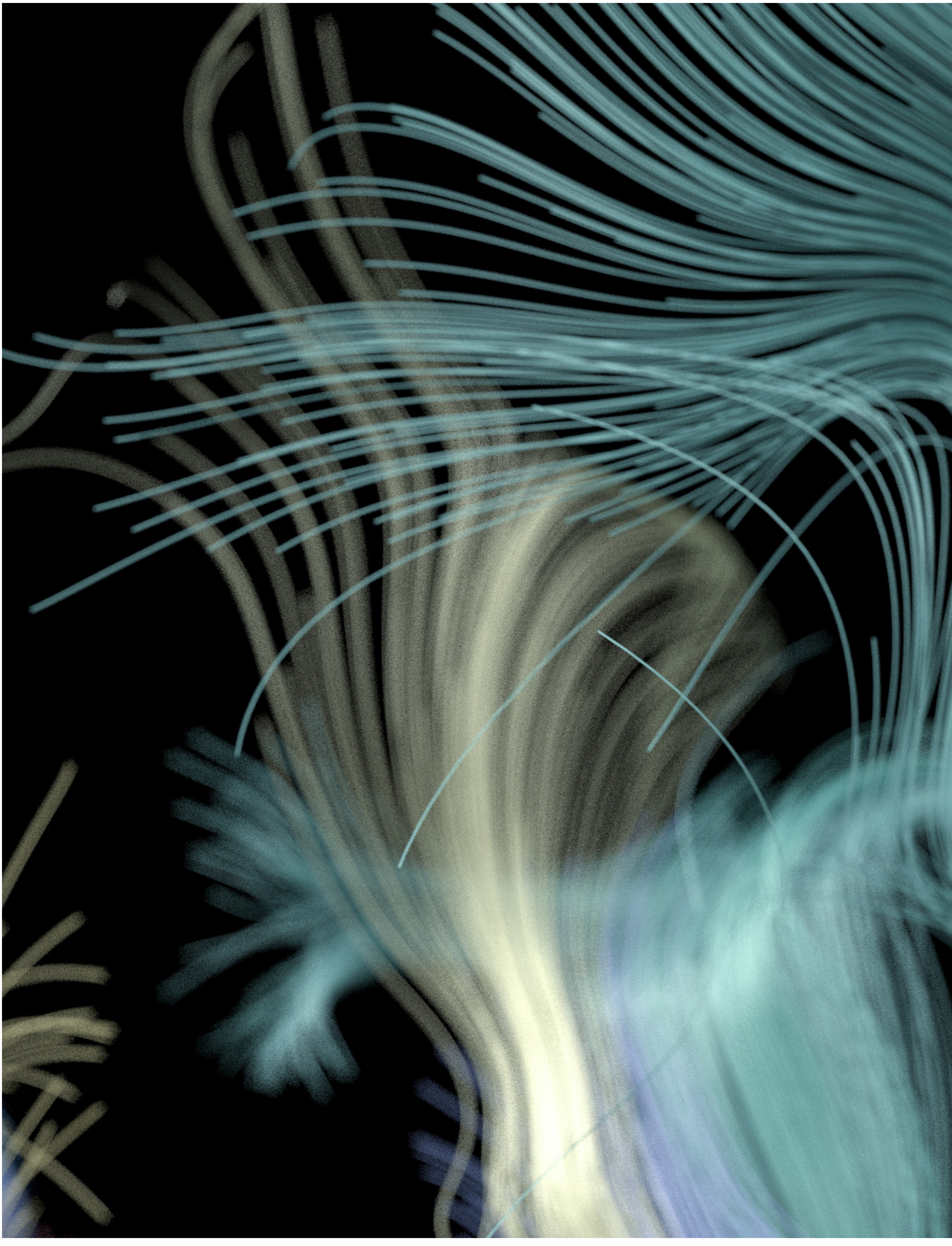
Note. From *A History of New Zealand Birds* by J. Keulemans & W. Buller, 1888, London, United Kingdom: (n.p.). In the public domain.



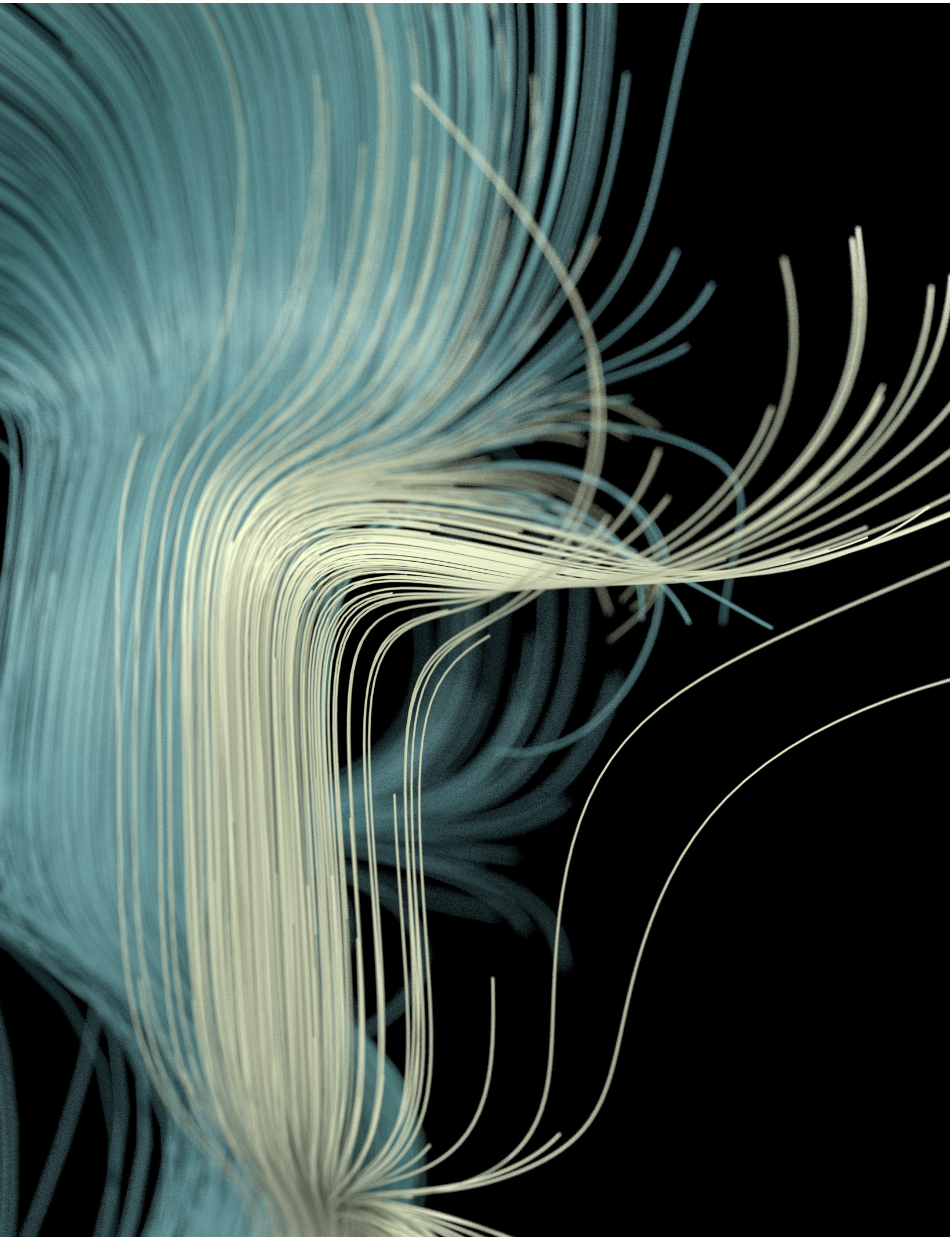
▲ **Figure 65.** *Kaka song - circular.*

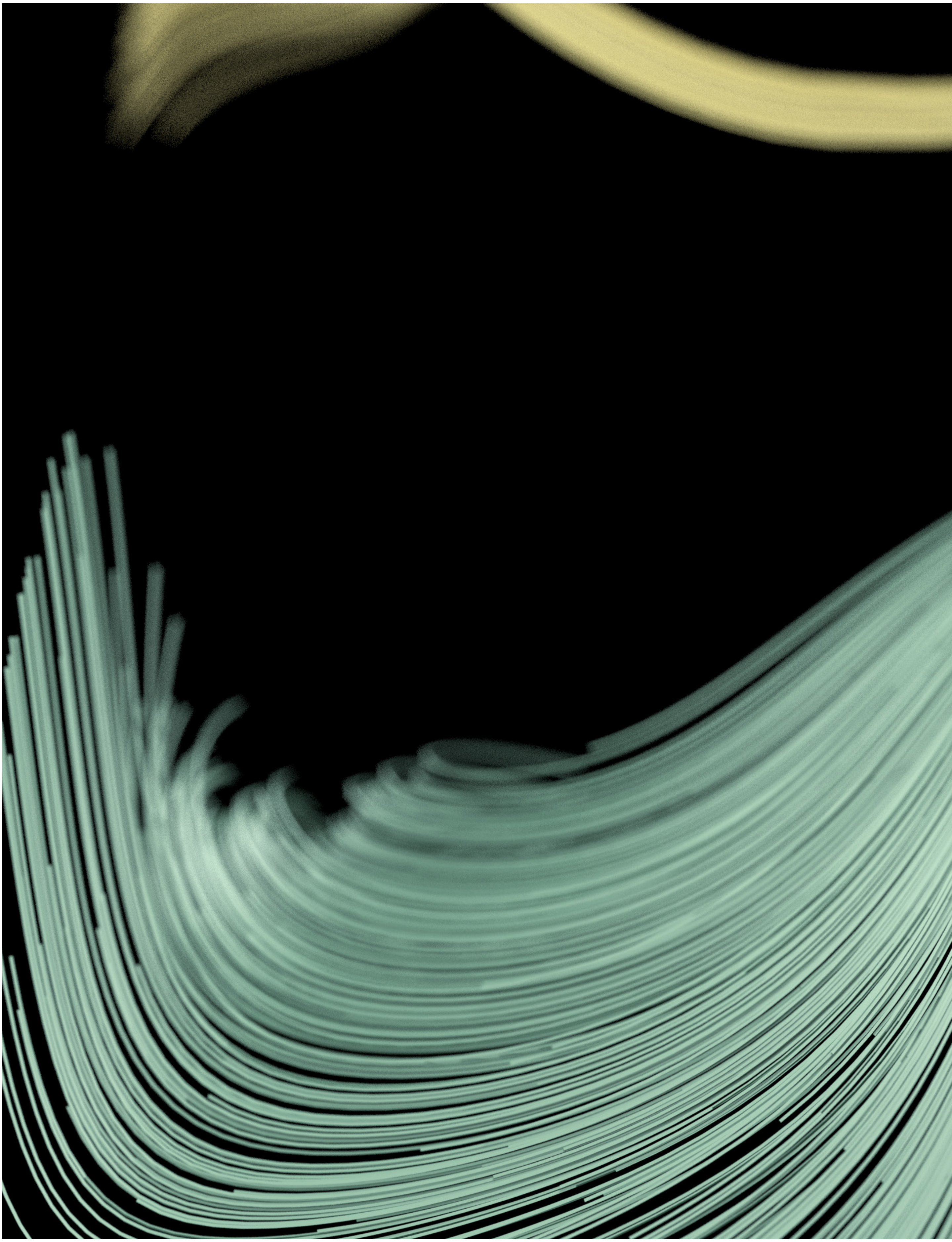
▼ **Figure 66.** *Kaka song - linear.*



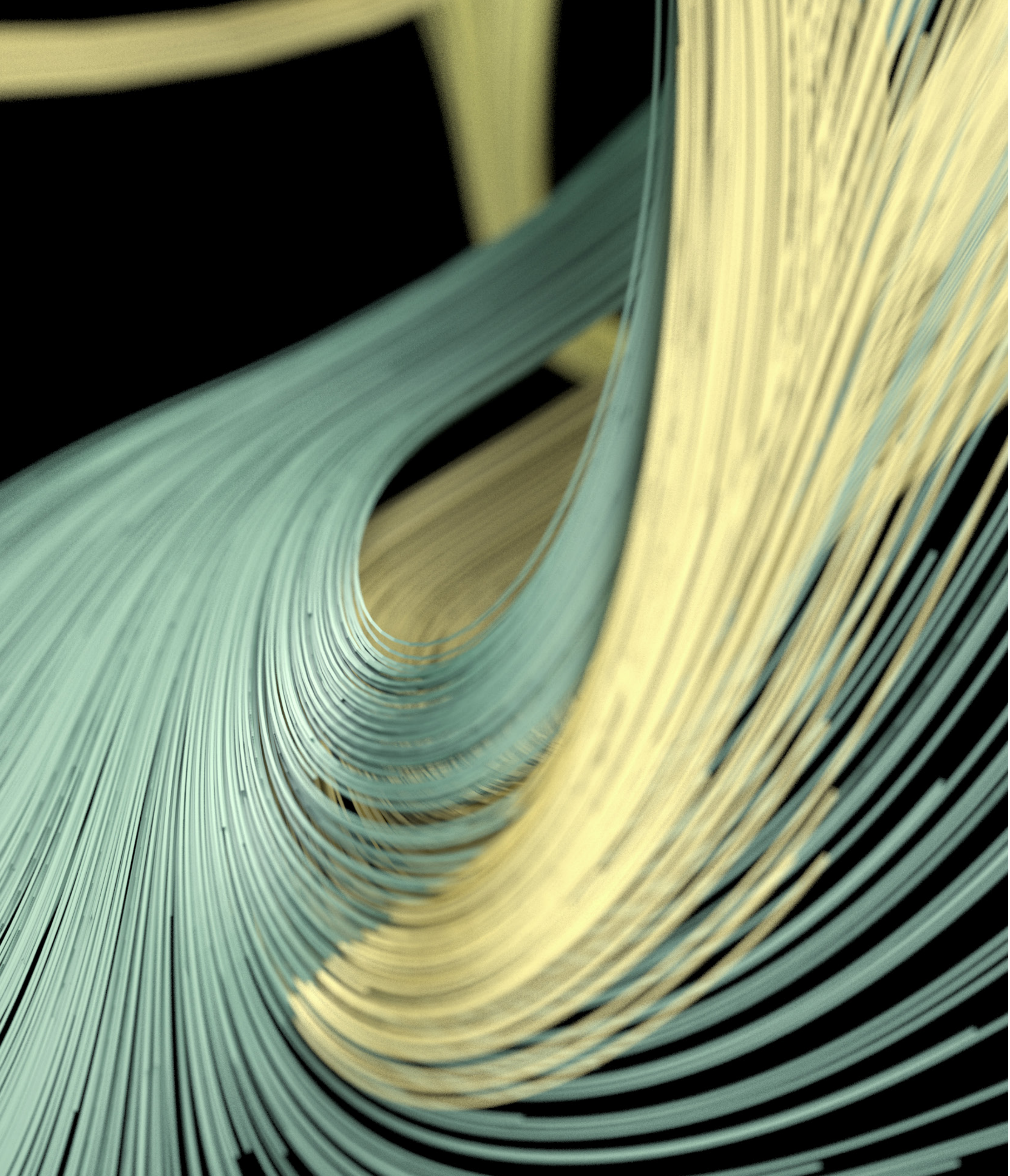


▼ **Figure 67.** *Kaka song - detail 01.*





▼ **Figure 68.** *Kaka song - detail 02.*



TUI

Scientific name : Prosthemadera novaeseelandiae

Location : Native forests, rural gardens and suburban parks across New Zealand

Population : Widespread and locally abundant

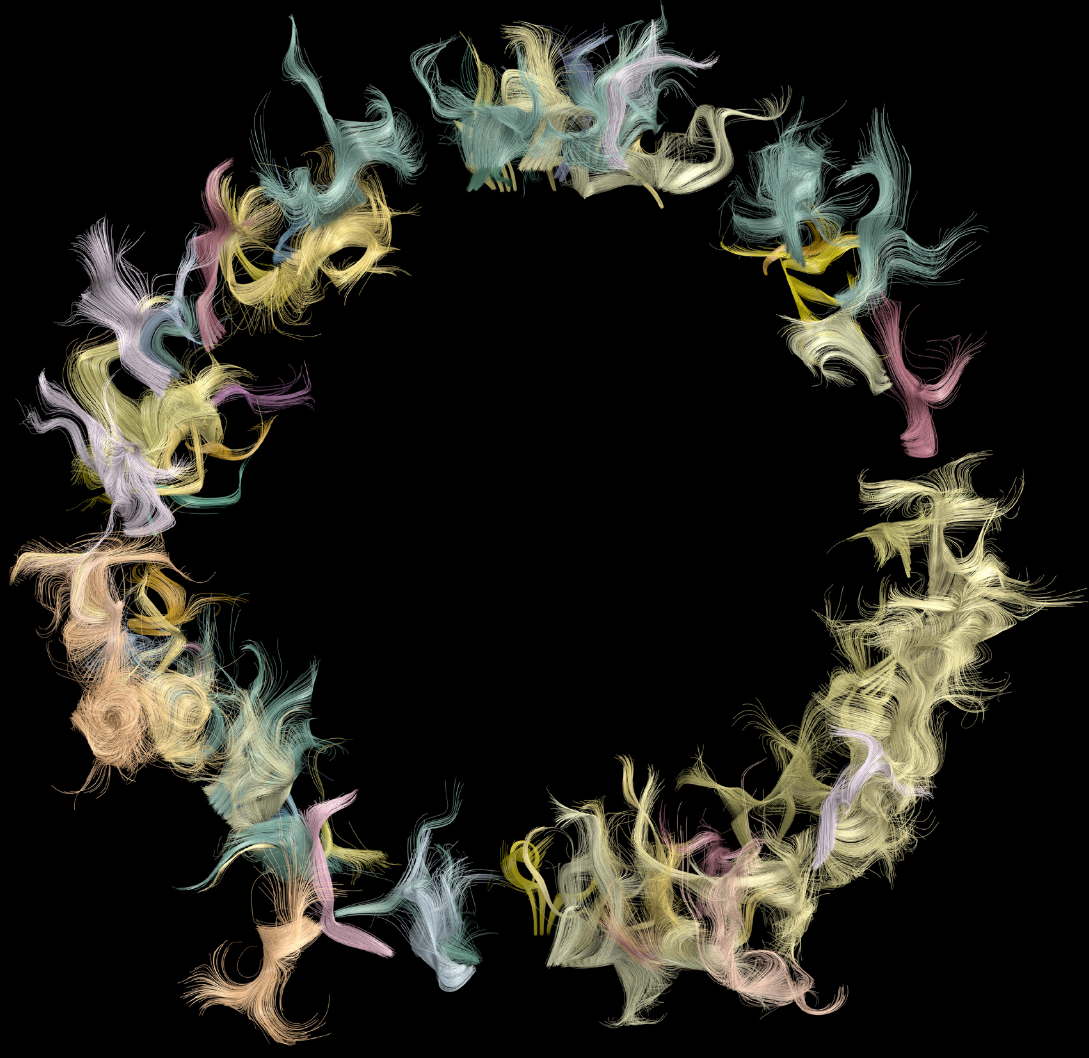
Conservation Status : Not threatened

Mass : 120g (male), 90g (female)

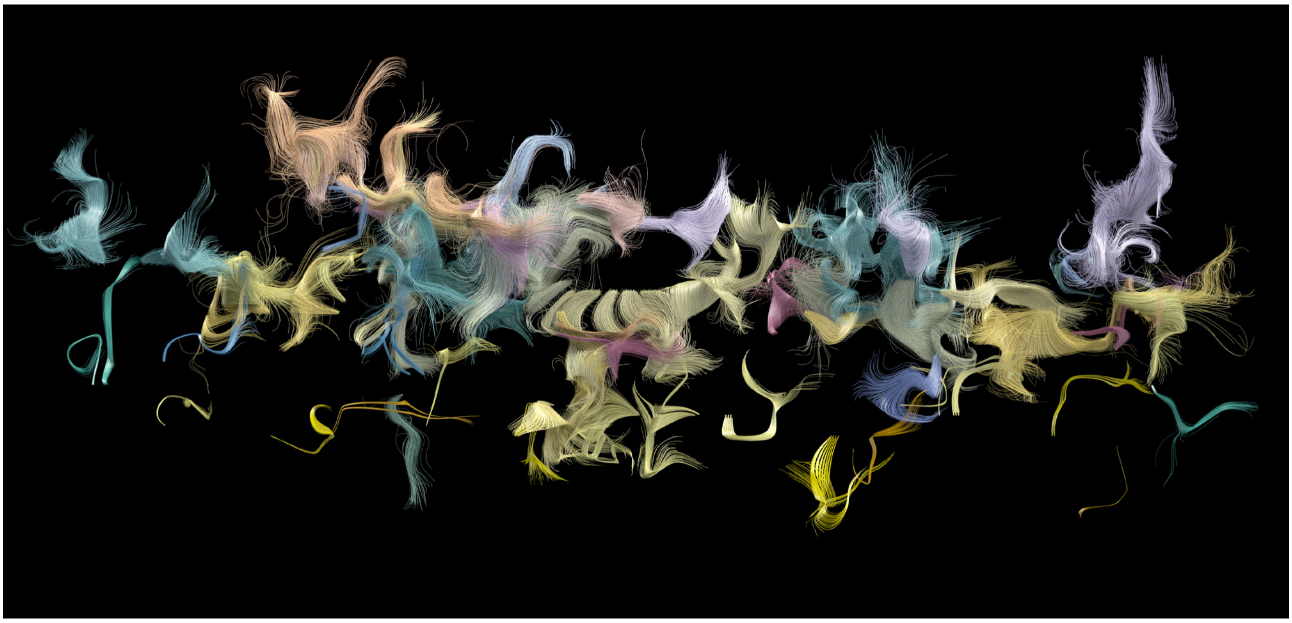
Tui songs consist of a loud and complicated mixture of tuneful notes which are interspersed with loud grunts, coughs and wheezes (Robertson, 2013).

▲ **Figure 69.** *Tui, Johannes Keulemans - 1888.*

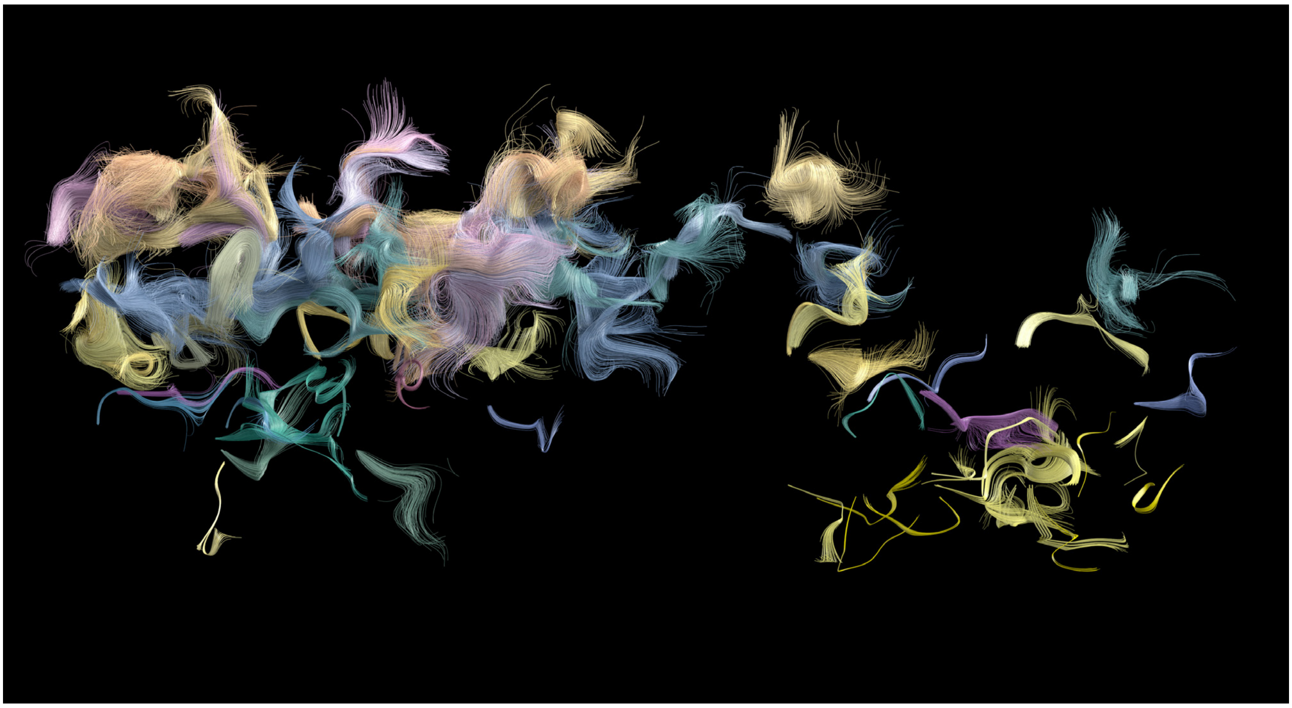
Note. From *A History of New Zealand Birds* by J. Keulemans & W. Buller, 1888, London, United Kingdom: (n.p.). In the public domain.



▲ **Figure 70.** *Tui song - circular.*



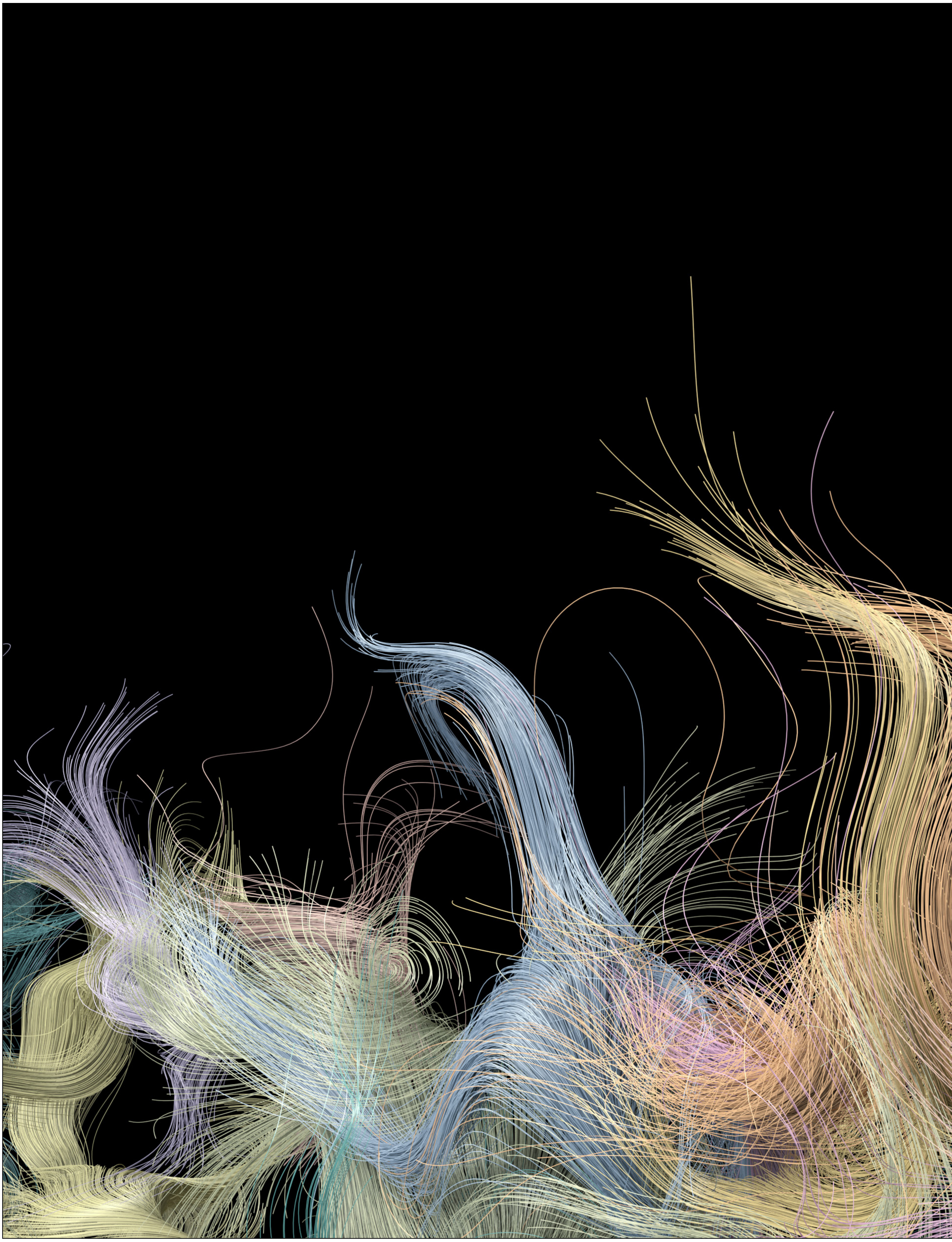
▲ **Figure 71.** *Tui song - linear.*

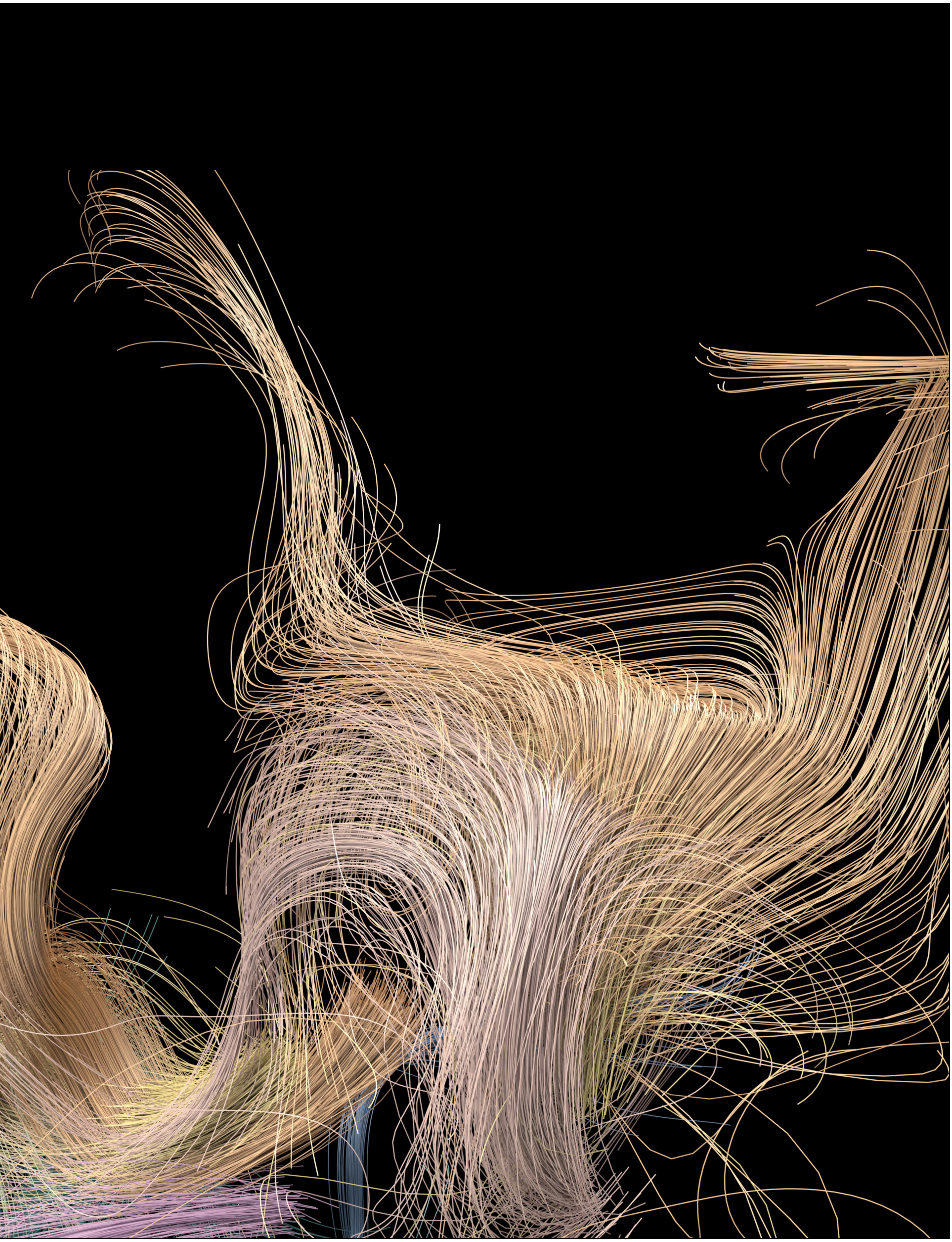


▲ **Figure 72.** *Tui territorial call - linear.*

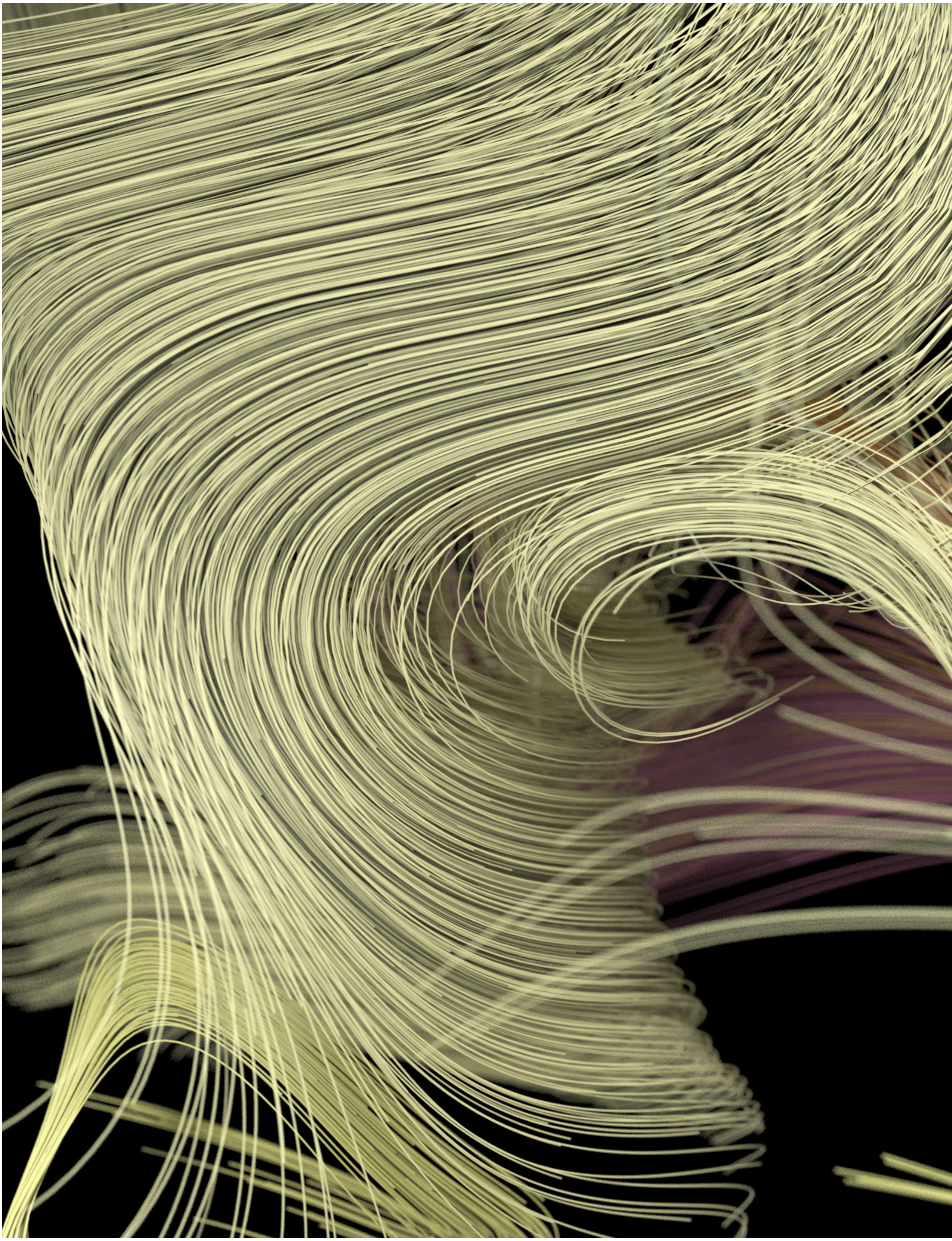
▲ **Figure 73.** *Tui communicative call - linear.*

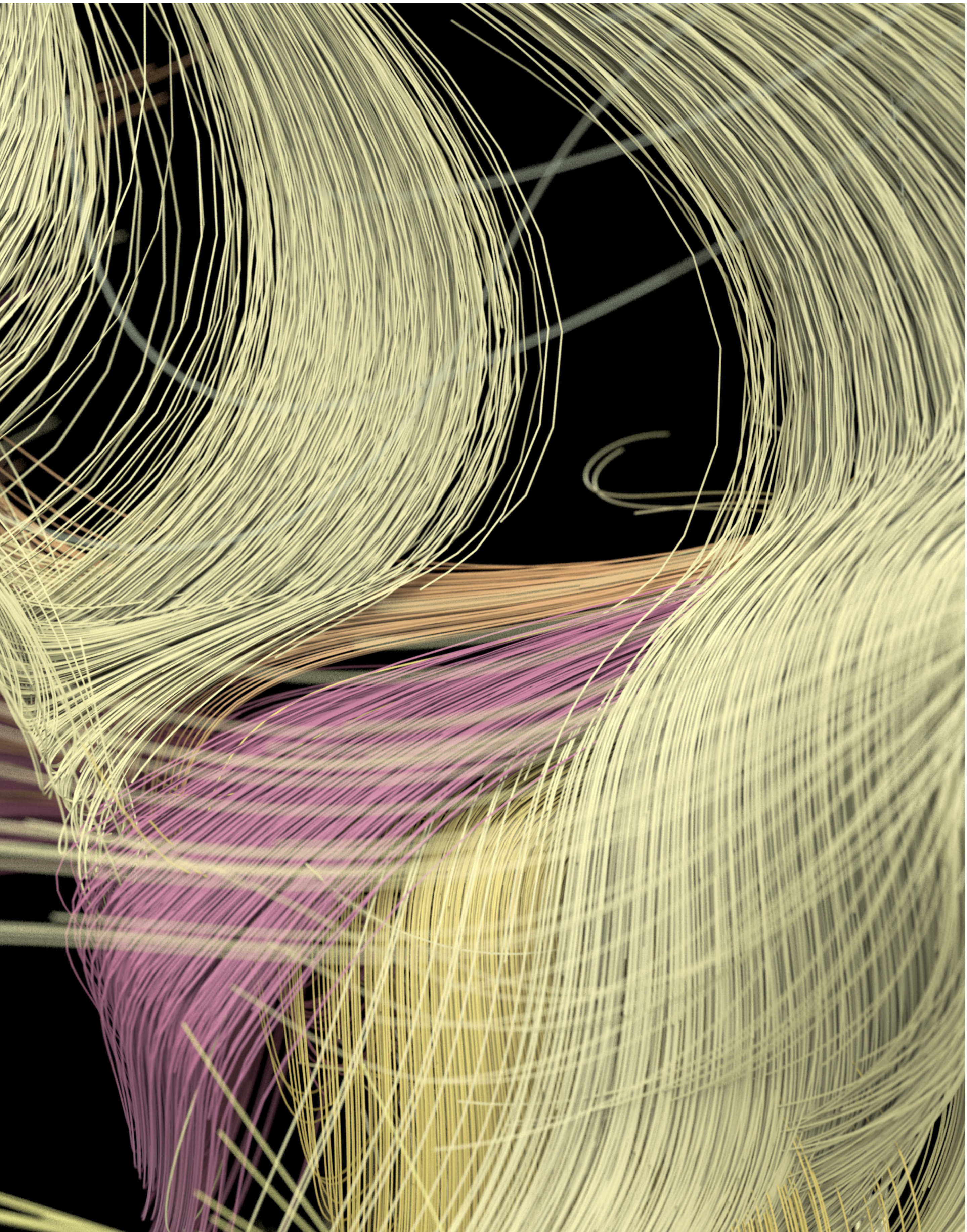
▼ **Figure 74.** *Tui song - detail 01.*





▼ **Figure 75.** *Tui song - detail 02.*





FANTAIL / PĪWAKAWAKA

Scientific name : Rhipidura fuliginosa

Location : Widespread across New Zealand

Population : Widespread and locally common

Conservation Status : Not threatened

Mass : 8g

The distinctive song of the Fantail consists of a chattering, tweeting rhythm with an extremely high pitch (Powlesland, 2013).

▲ **Figure 76.** *Fantail, Johannes Keulemans - 1888.*

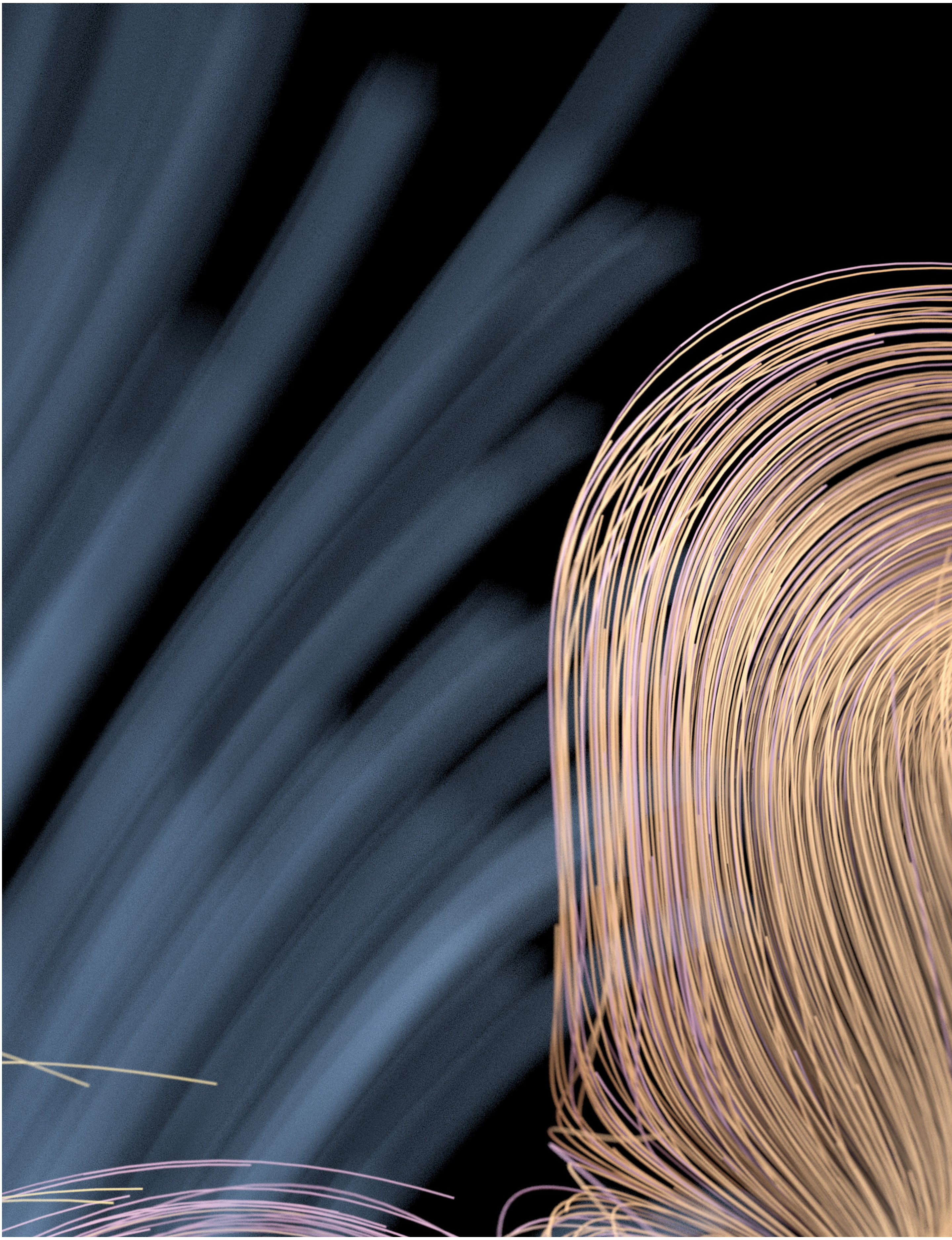
Note. From *A History of New Zealand Birds* by J. Keulemans & W. Buller, 1888, London, United Kingdom: (n.p.). In the public domain.

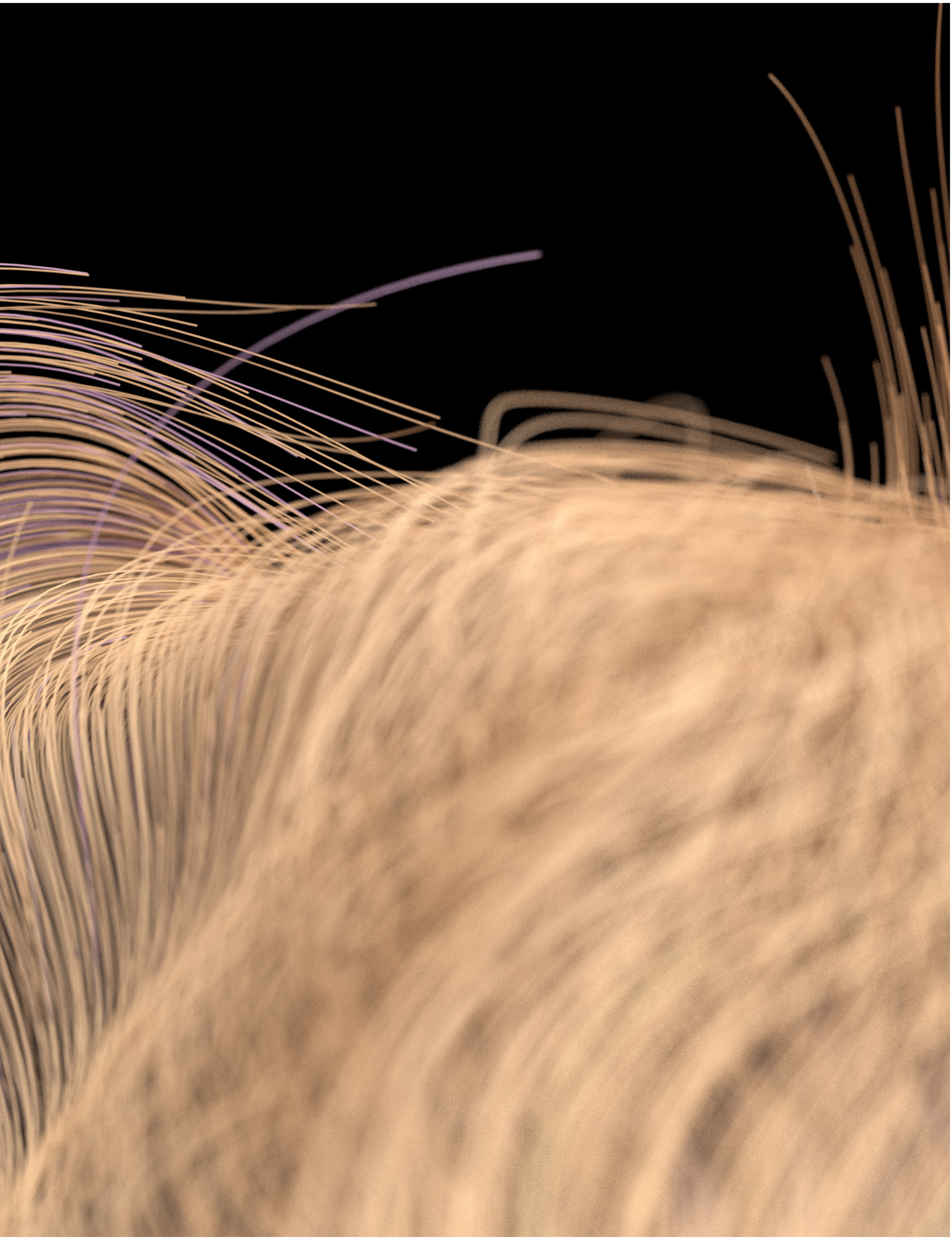
▲ **Figure 77.** *Fantail song - circular.*

▲ **Figure 78.** *Fantail song South Island - linear.*

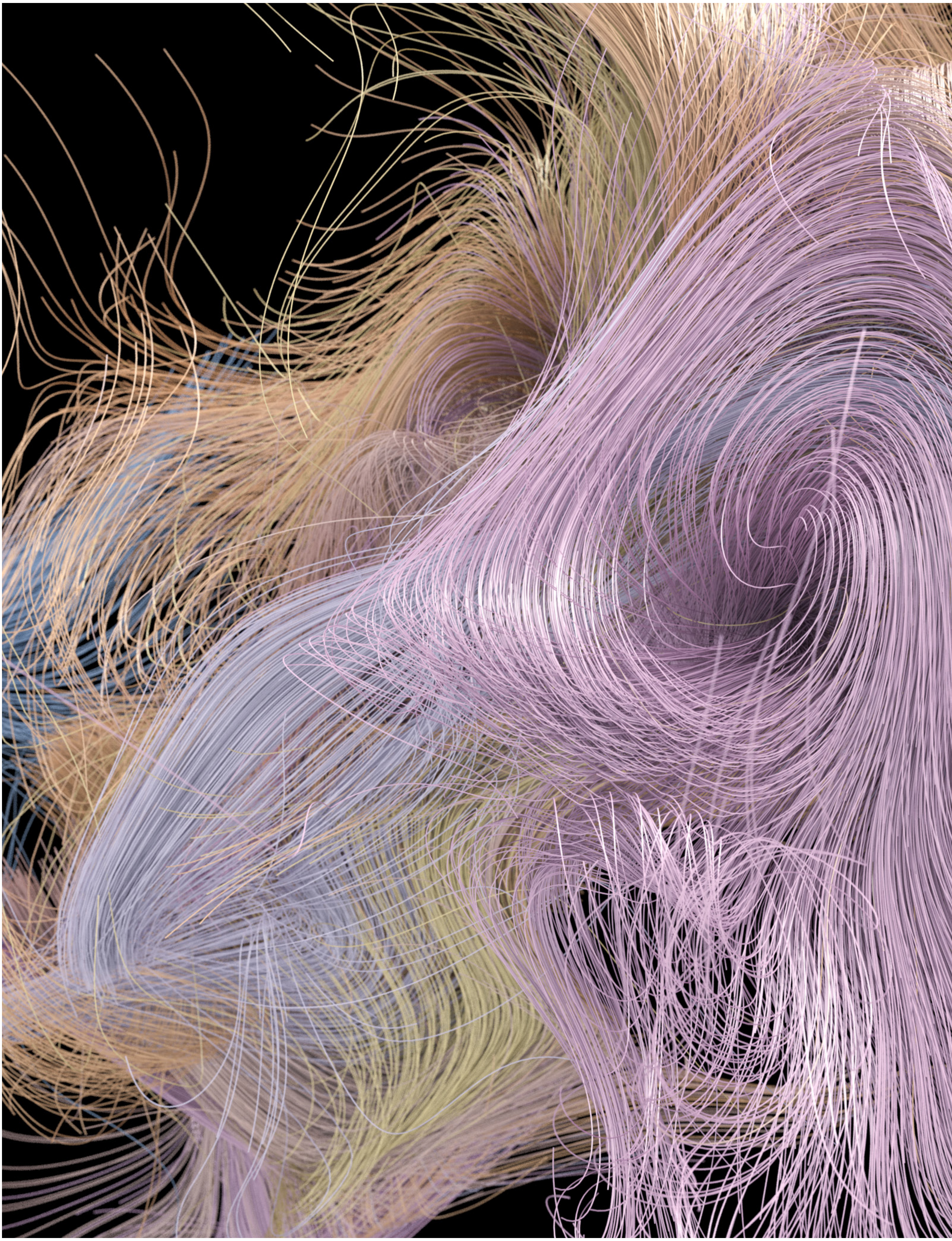
▲ **Figure 79.** *Fantail song North Island - linear.*

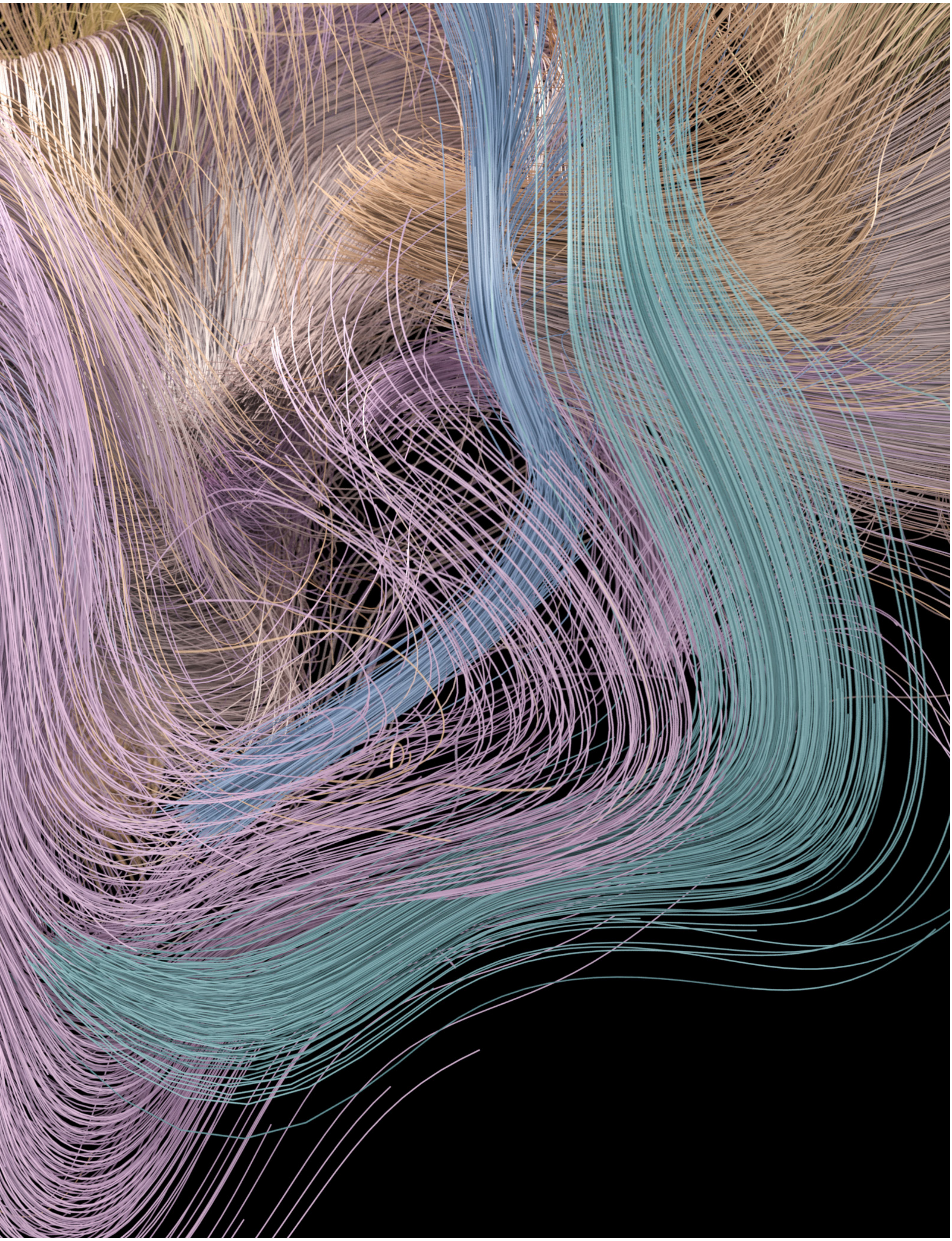
▼ **Figure 80.** *Fantail song - detail 01.*





▼ **Figure 81.** *Fantail song - detail 02.*





NORTH ISLAND BROWN KIWI

Scientific name : Apteryx mantelli

Location : Native and exotic forests, shrubs and farmlands in Northland and some offshore islands

Population : About 25,000 individuals

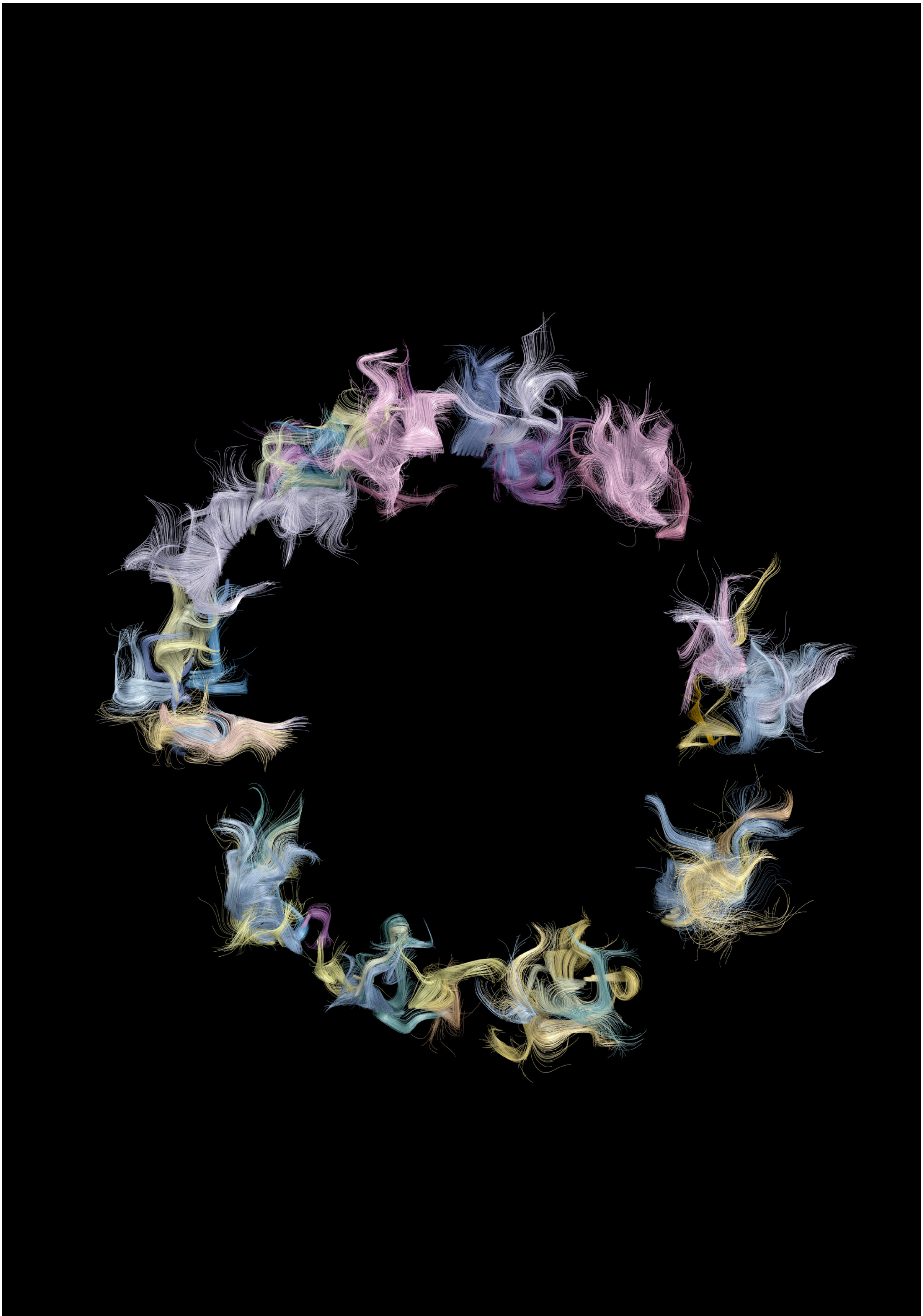
Conservation Status : Declining

Mass : 2kg (male), 2.4kg (female)

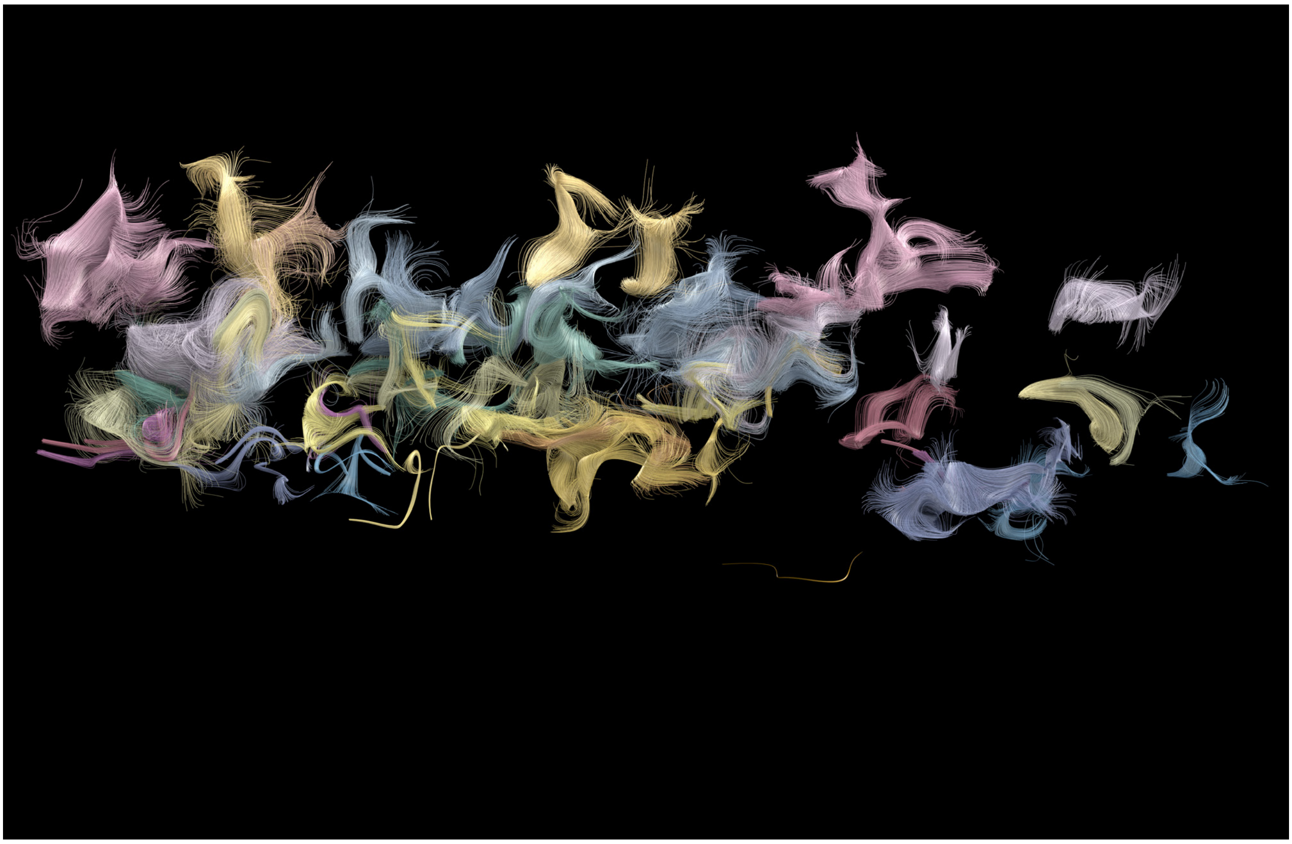
The song of the male Kiwi consists of a number of high-pitched, ascending notes, while the female's song consists of a mixture of high-pitched notes and repeating guttural screams (Robertson, 2013).

▲ **Figure 82.** *North Island Brown Kiwi, Johannes Keulemans - 1872.*

Note. From *A History of New Zealand Birds* by J. Keulemans & W. Buller, 1888, London, United Kingdom: (n.p.). In the public domain.



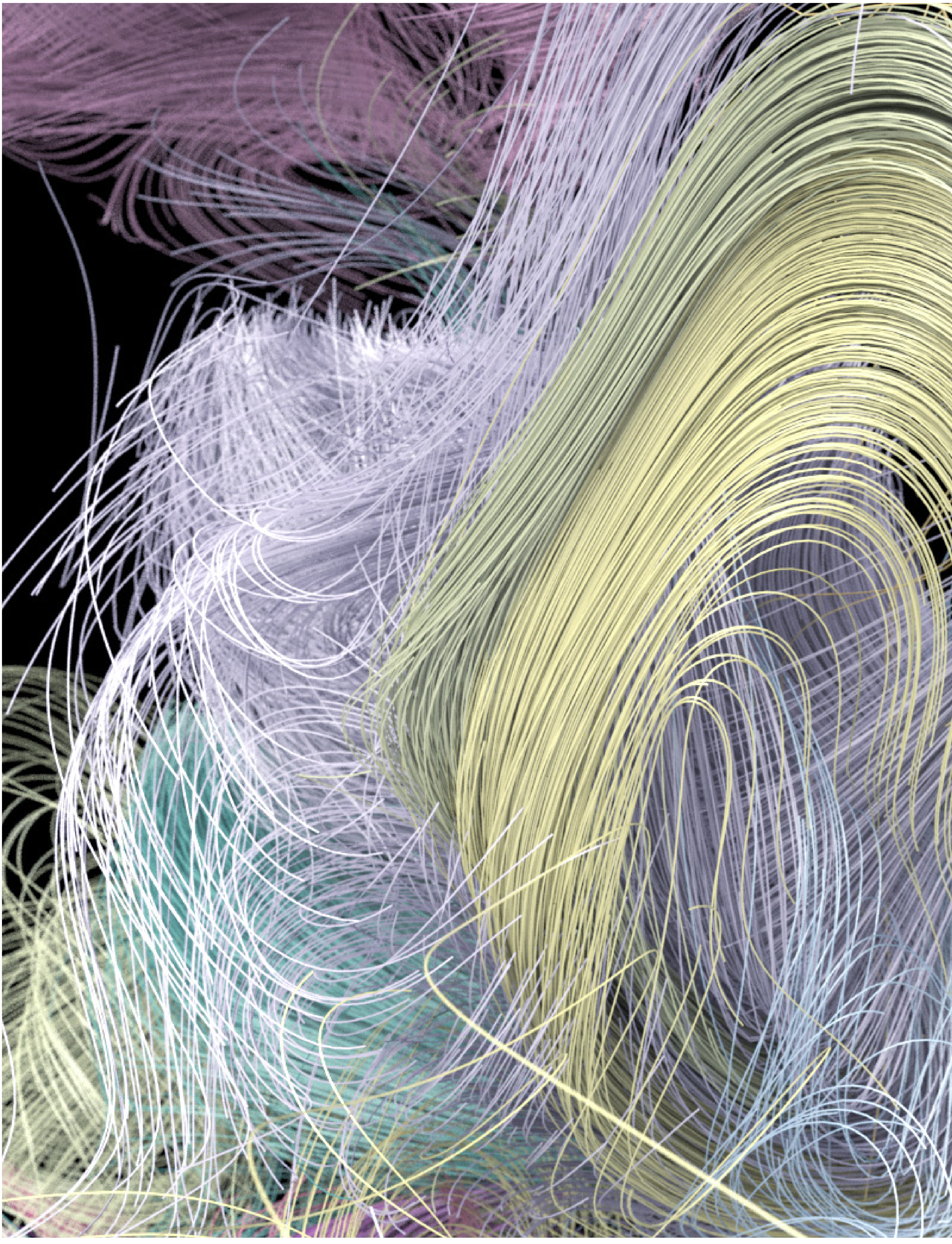
▲ **Figure 83.** *Kiwi song - circular.*

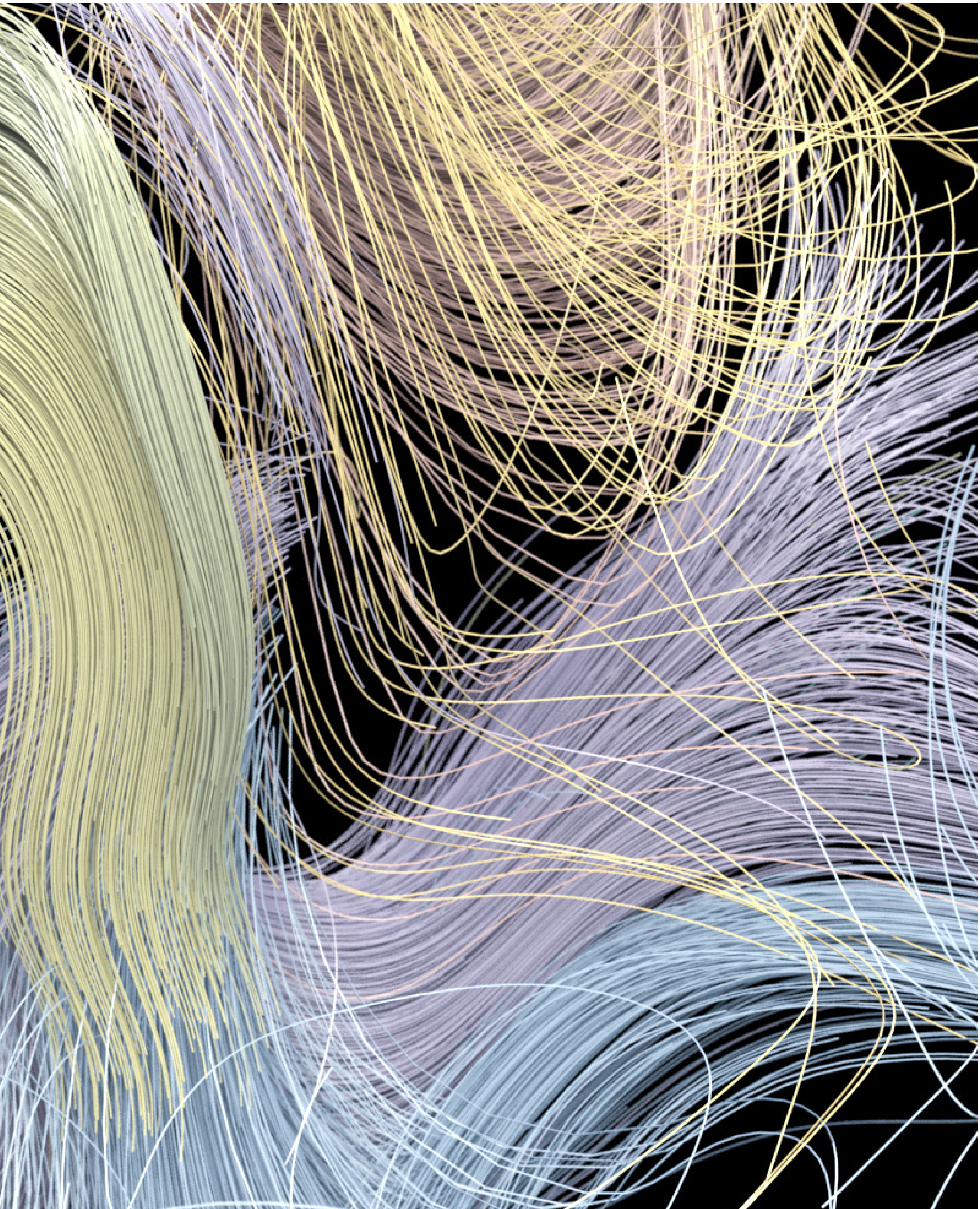


▲ **Figure 84.** *Kiwi song female - linear.*

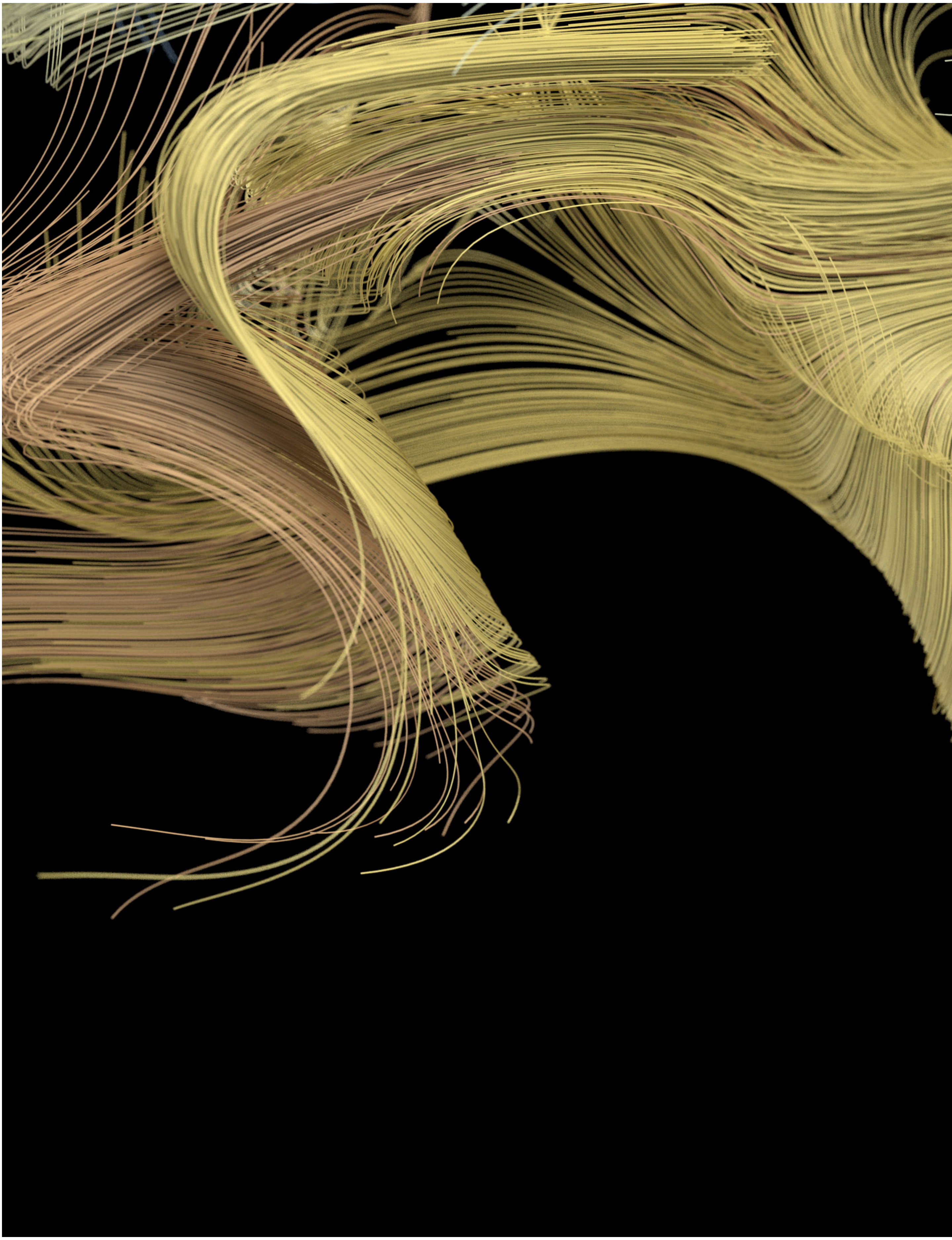
▲ **Figure 85.** *Kiwi song male - linear.*

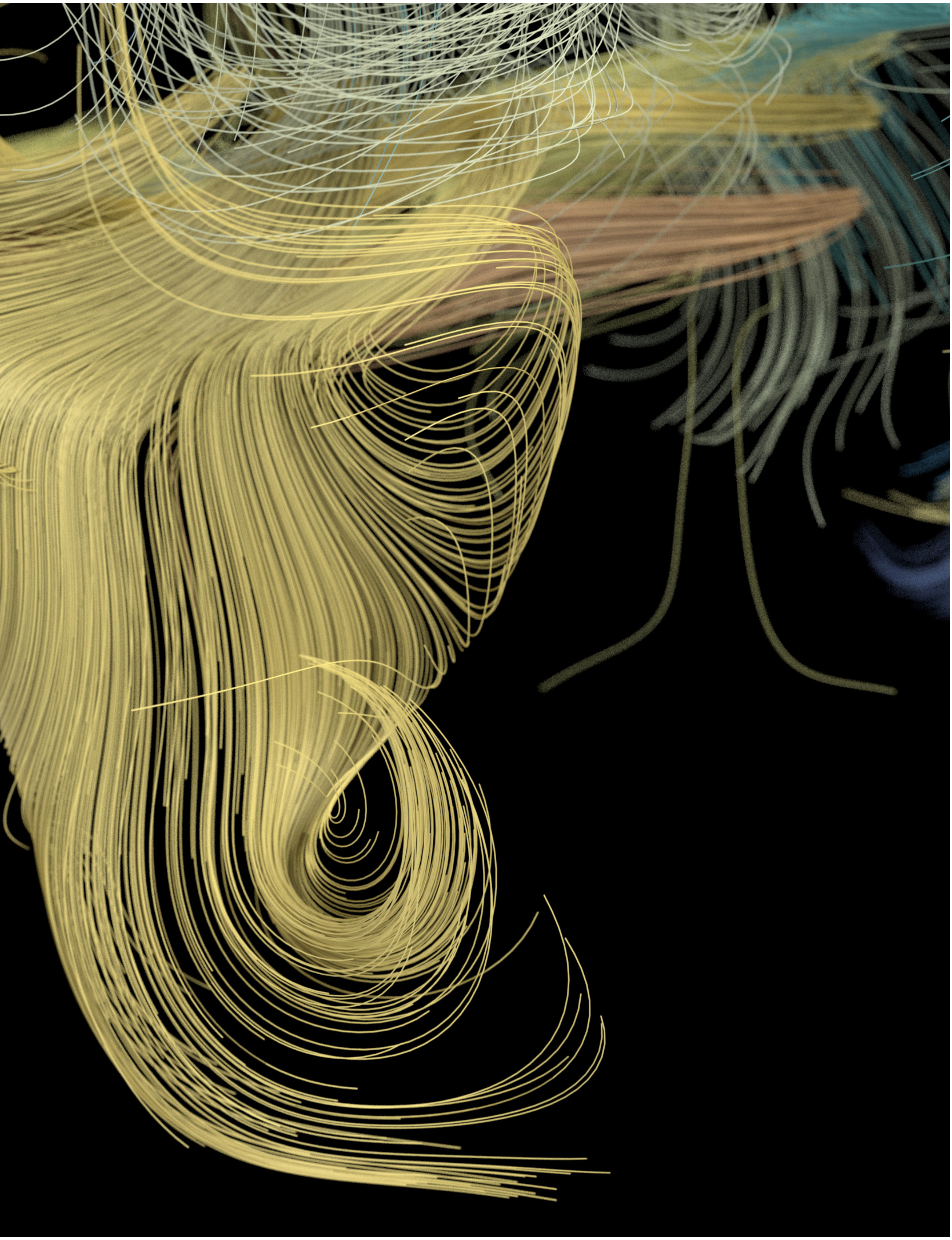
▼ **Figure 86.** *Kiwi song - detail 01.*





▼ **Figure 87.** *Kiwi song - detail 02.*





NORTH ISLAND KŌKAKO

Scientific name : Callaeas wilsoni

Location : Confined to a few scattered forests in the Northern half of the North Island

Population : About 3000 individuals

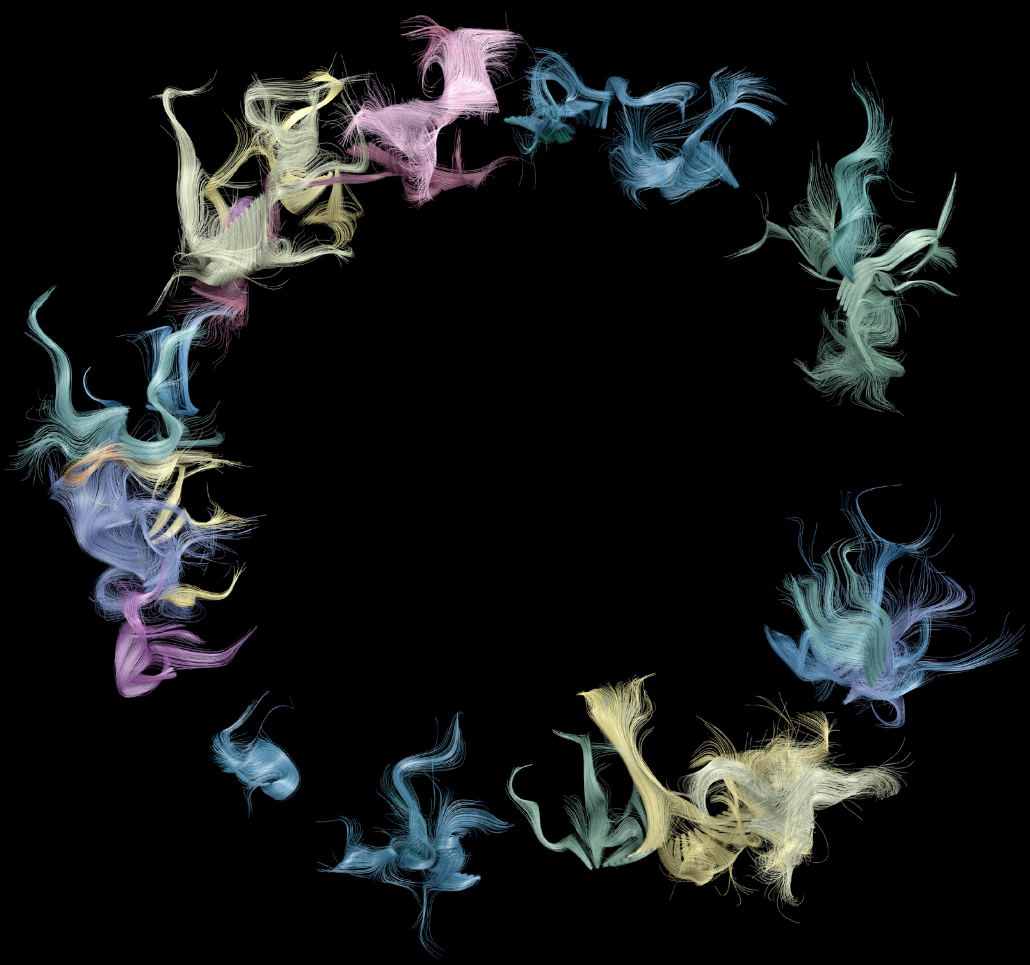
Conservation Status : Recovering

Mass : 233g (male), 218g (female)

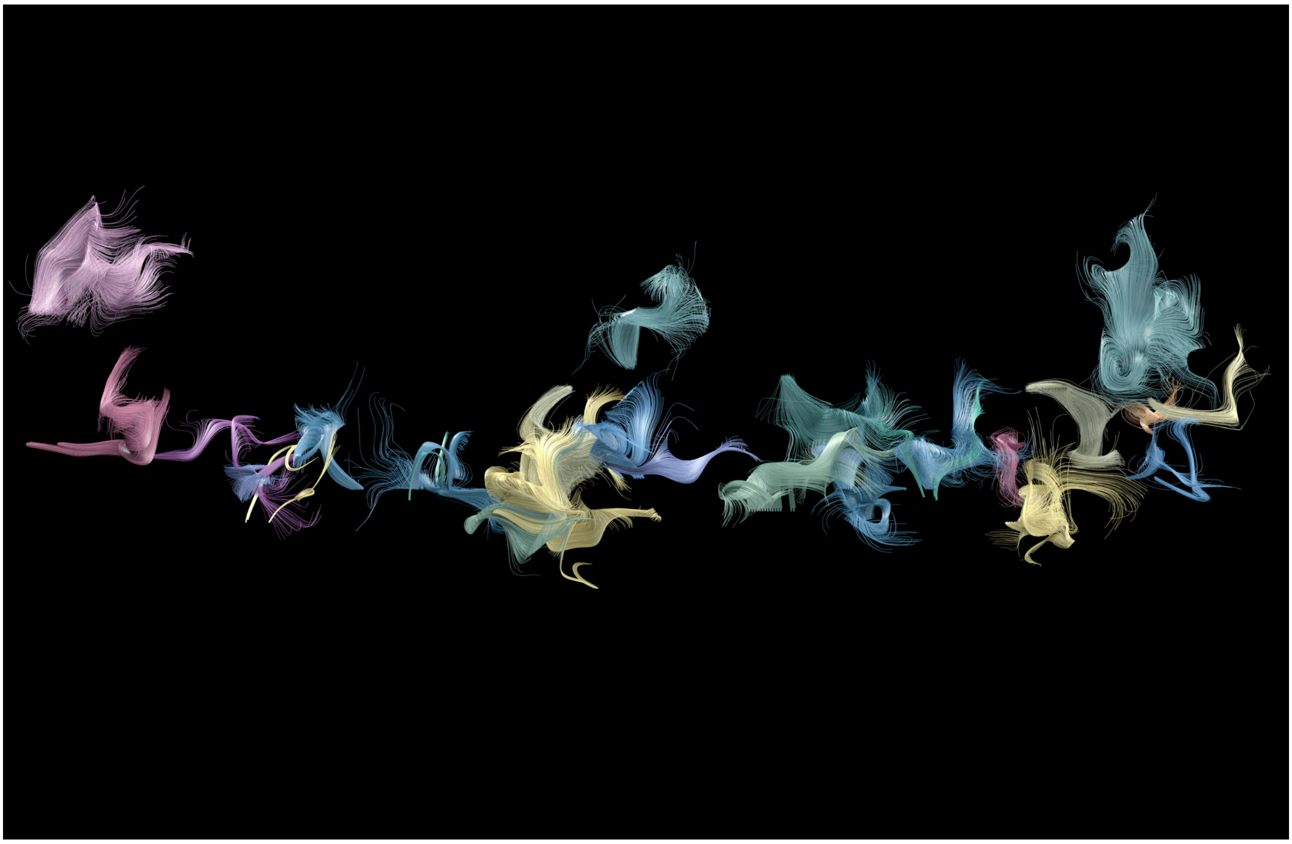
The song of the Kokako consists of slow, sonorous, sustained organ-like notes typically sung as a duet and often from a high perch (Innes, 2013).

▲ **Figure 88.** *North Island and South Island Kōkako, Johannes Keulemans - 1872.*

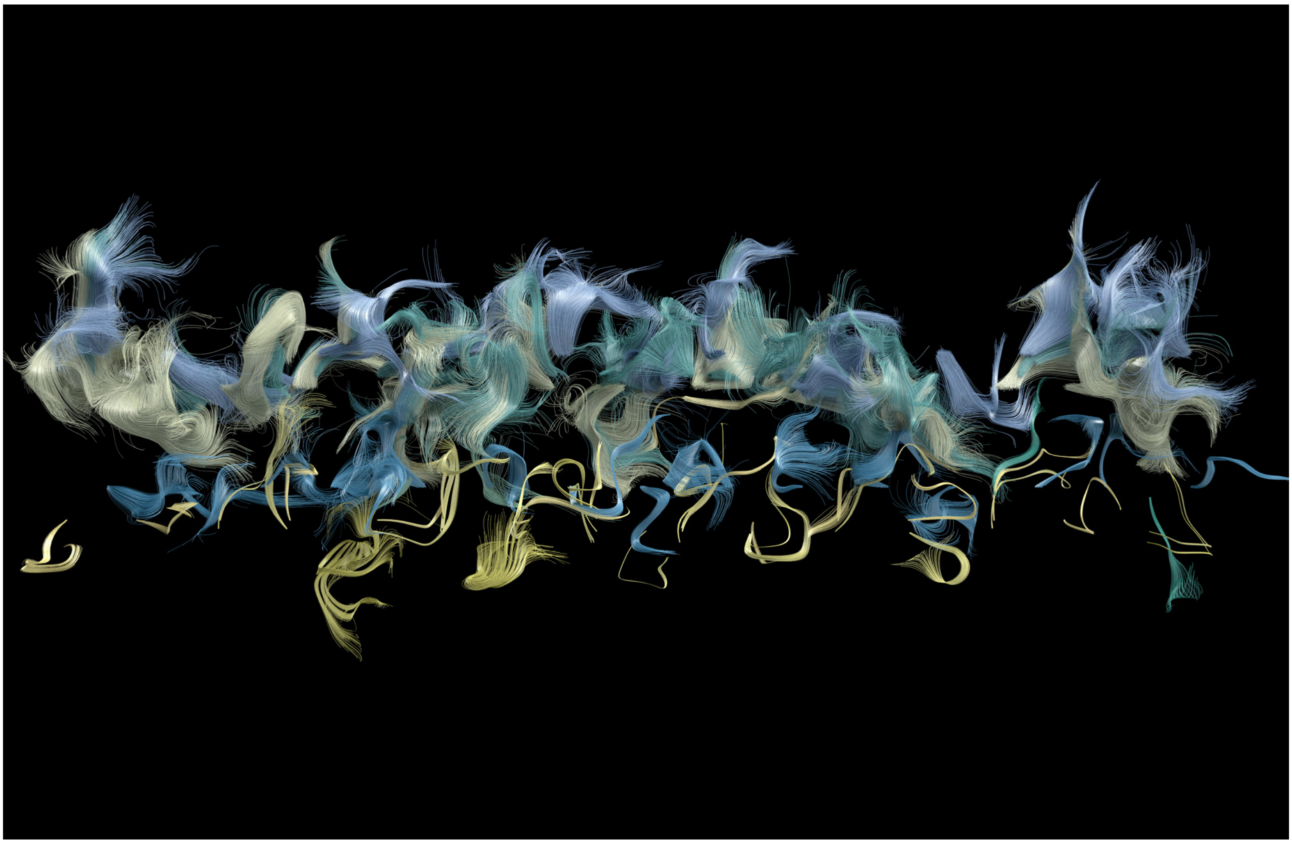
Note. From *A History of New Zealand Birds* by J. Keulemans & W. Buller, 1888, London, United Kingdom: (n.p.). In the public domain.



▲ **Figure 89.** *Kōkako song - circular.*

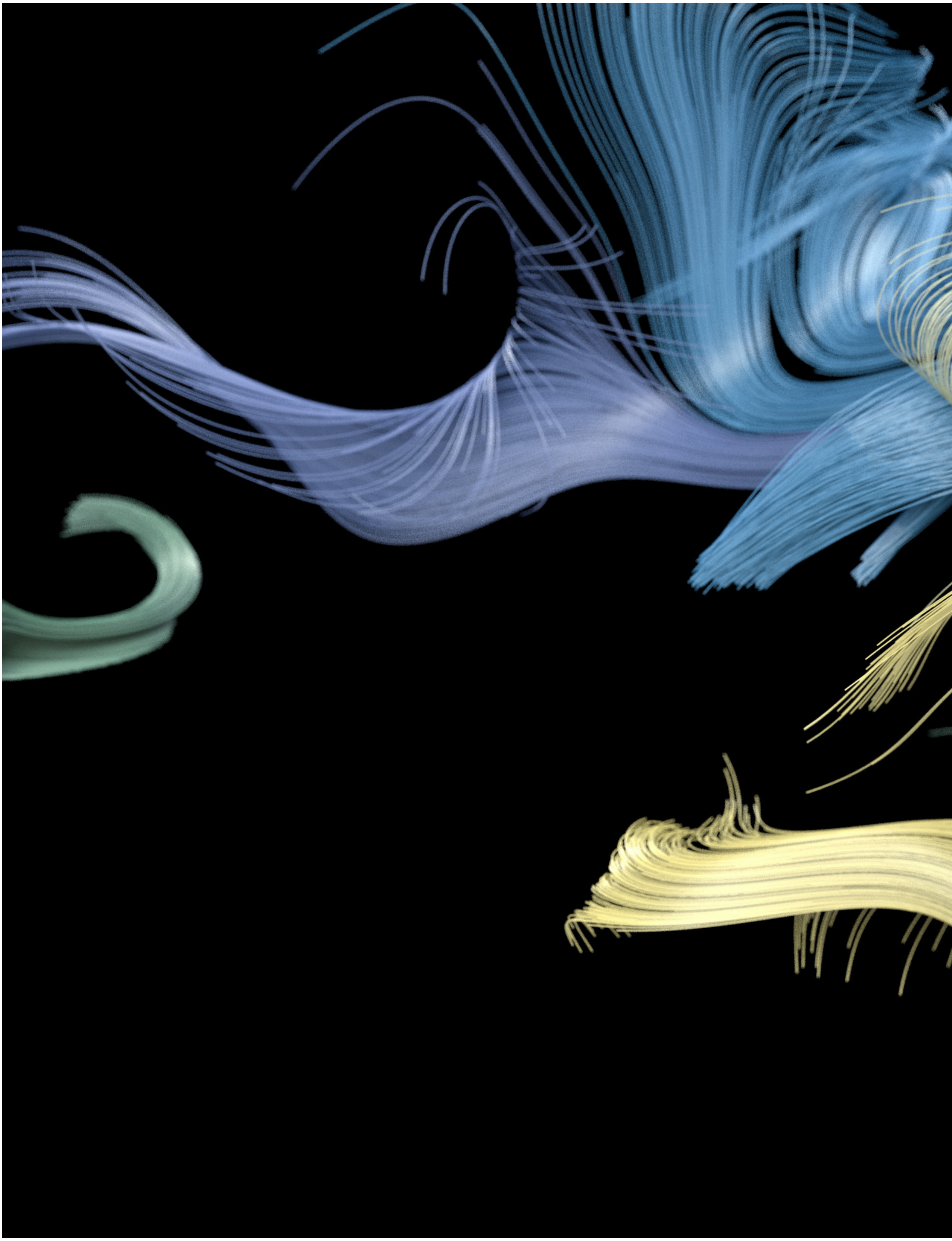


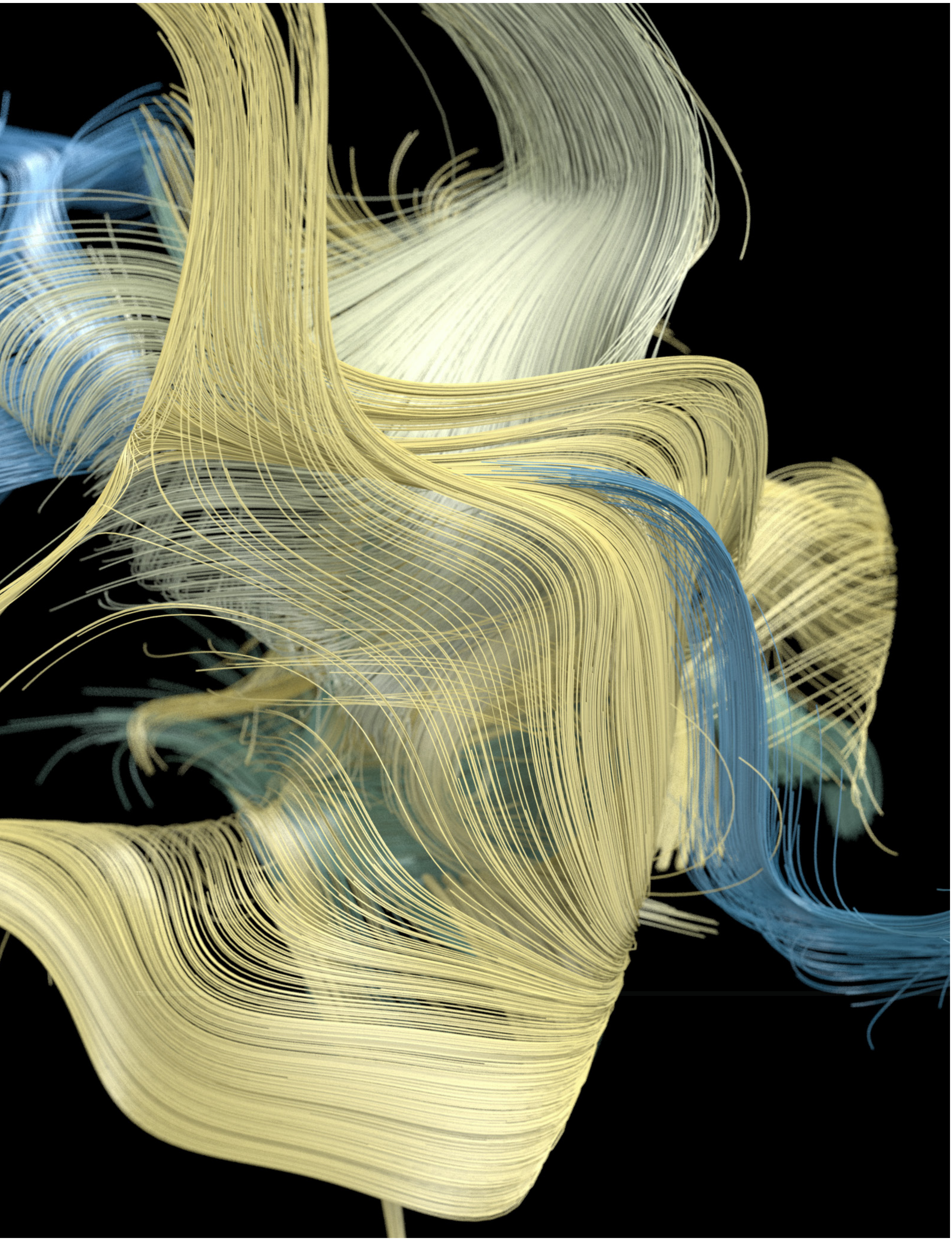
▲ **Figure 90.** *Kōkako song female - linear.*



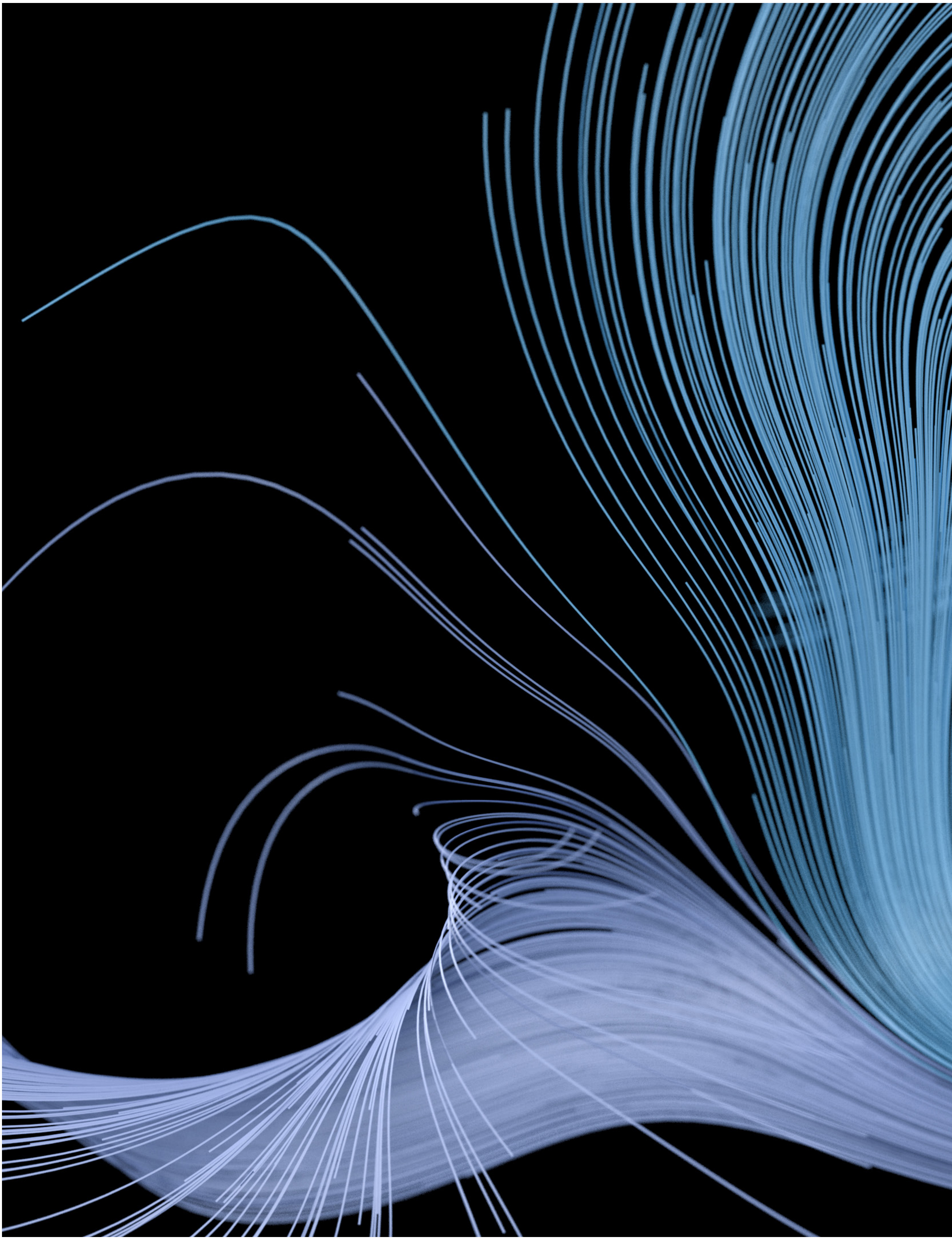
▲ **Figure 91.** *Kōkako alarm call - linear.*

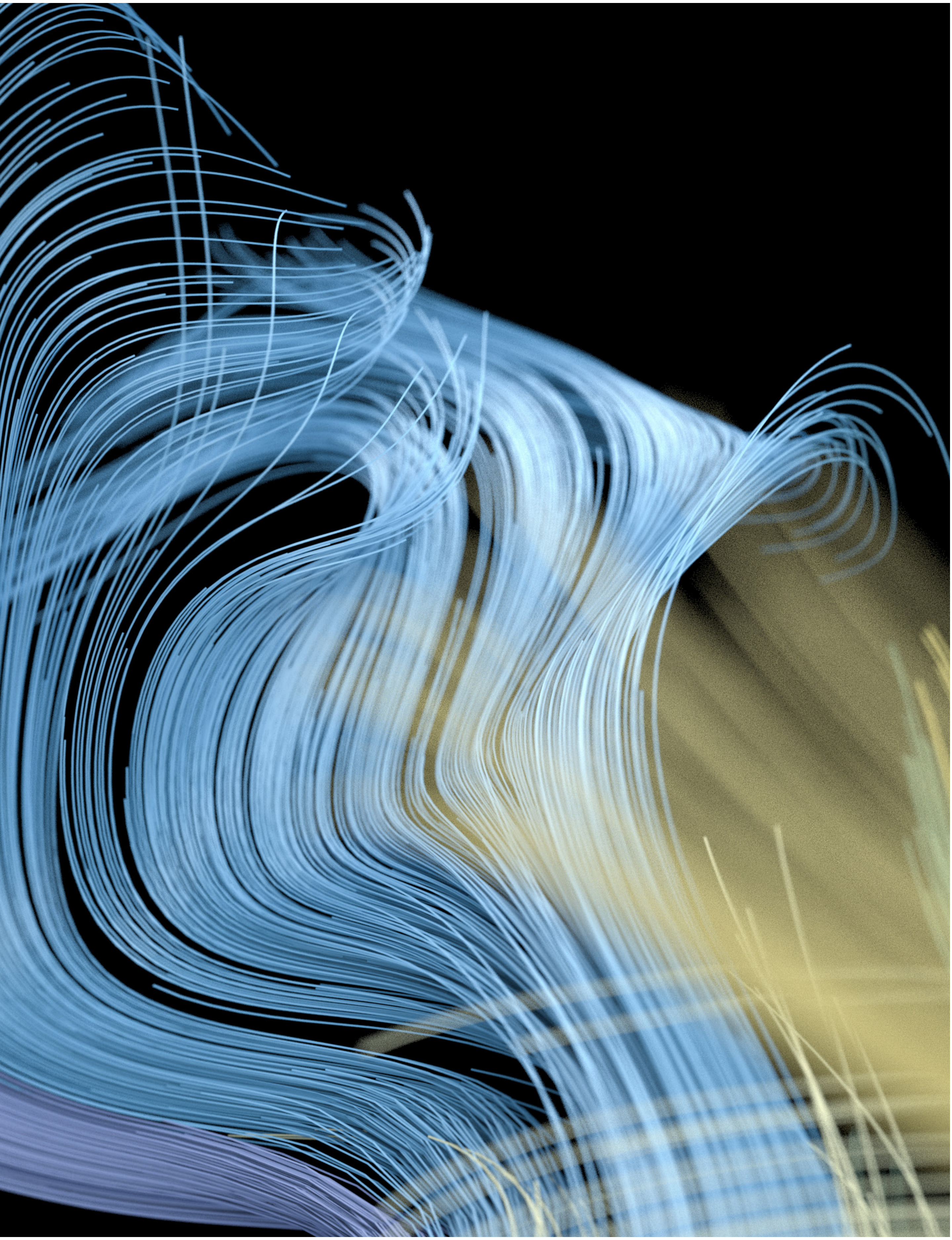
▼ **Figure 92.** *Kōkako song - detail 01.*





▼ **Figure 93.** *Kōkako song - detail 02.*





BELLBIRD / KORIMAKO

Scientific name : Anthornis Melanura

Location : Native and exotic forests, scrub, parks and gardens across New Zealand

Population : Common and abundant in predator-free offshore islands

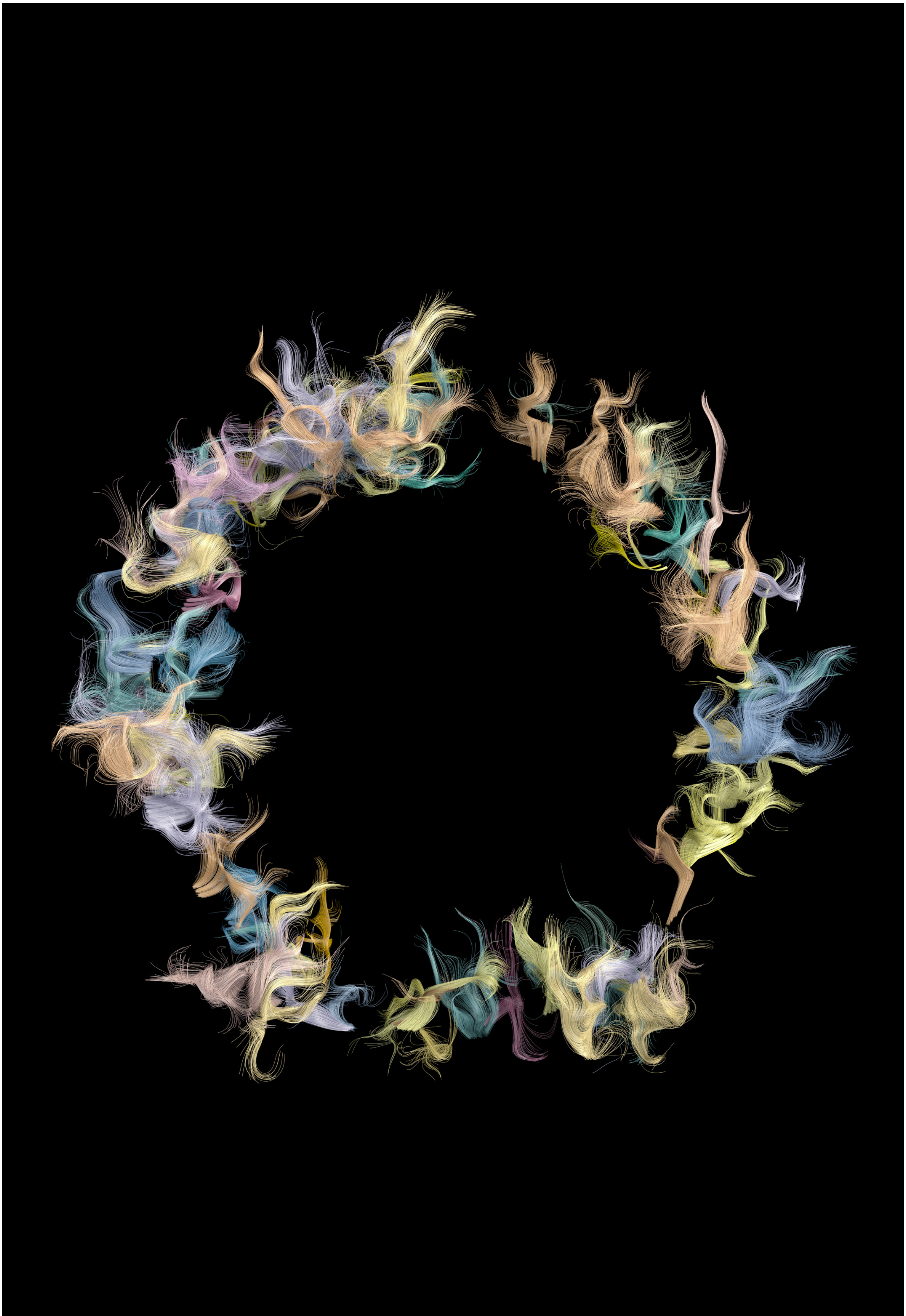
Conservation Status : Not threatened

Mass : 34g (male), 26g (female)

The song of a Bellbird is very similar to that of a Tui, with a tuneful mixture of sporadic notes. The major difference is that it lacks the grunts, coughs, and wheezes which are prevalent in the Tui song (Sagar, 2013).

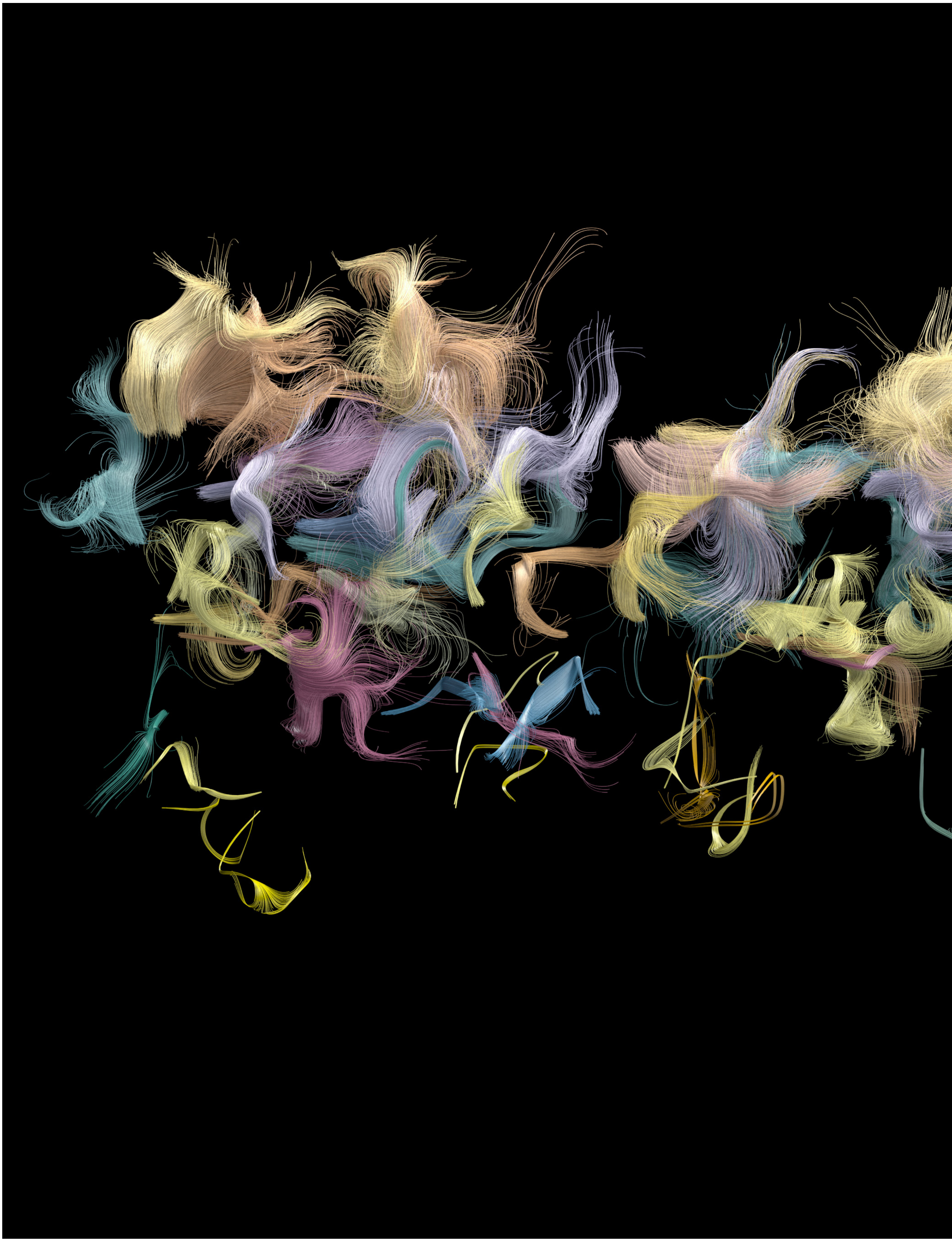
▲ **Figure 94.** *Bellbird, Johannes Keulemans - 1888.*

Note. From *A History of New Zealand Birds* by J. Keulemans & W. Buller, 1888, London, United Kingdom: (n.p.). In the public domain.

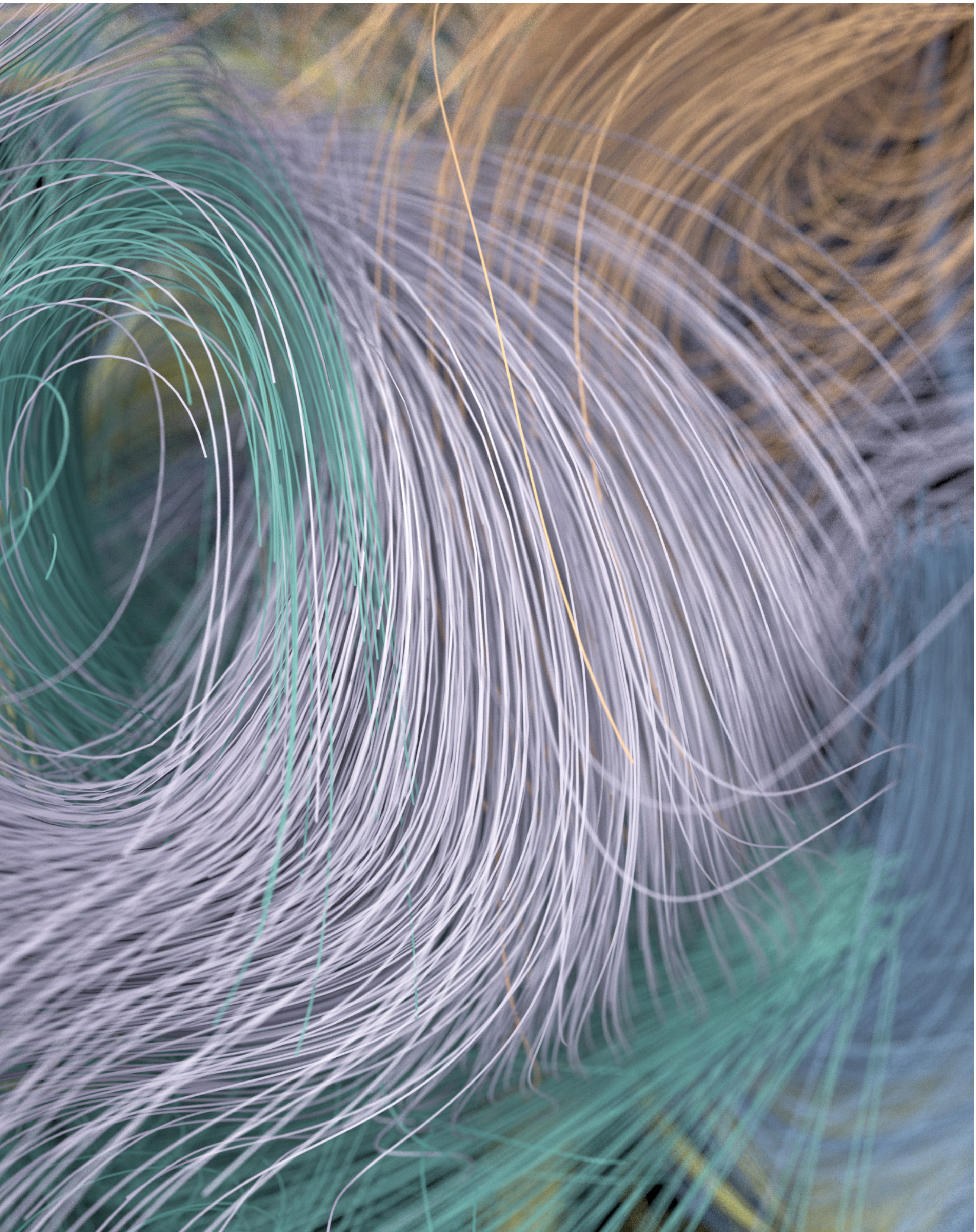


▲ **Figure 95.** *Bellbird song - circular.*

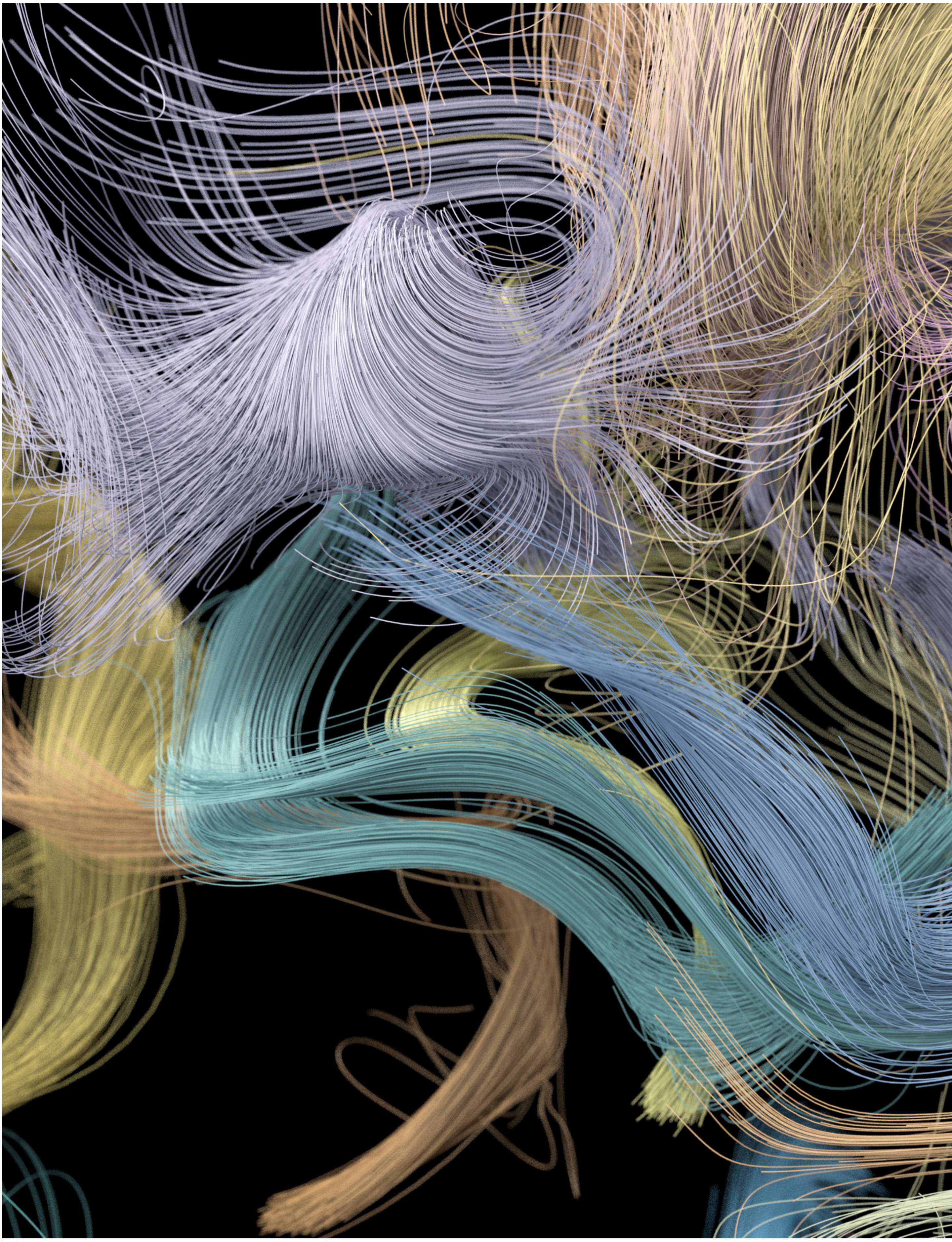
▼ **Figure 96.** *Bellbird song - linear.*

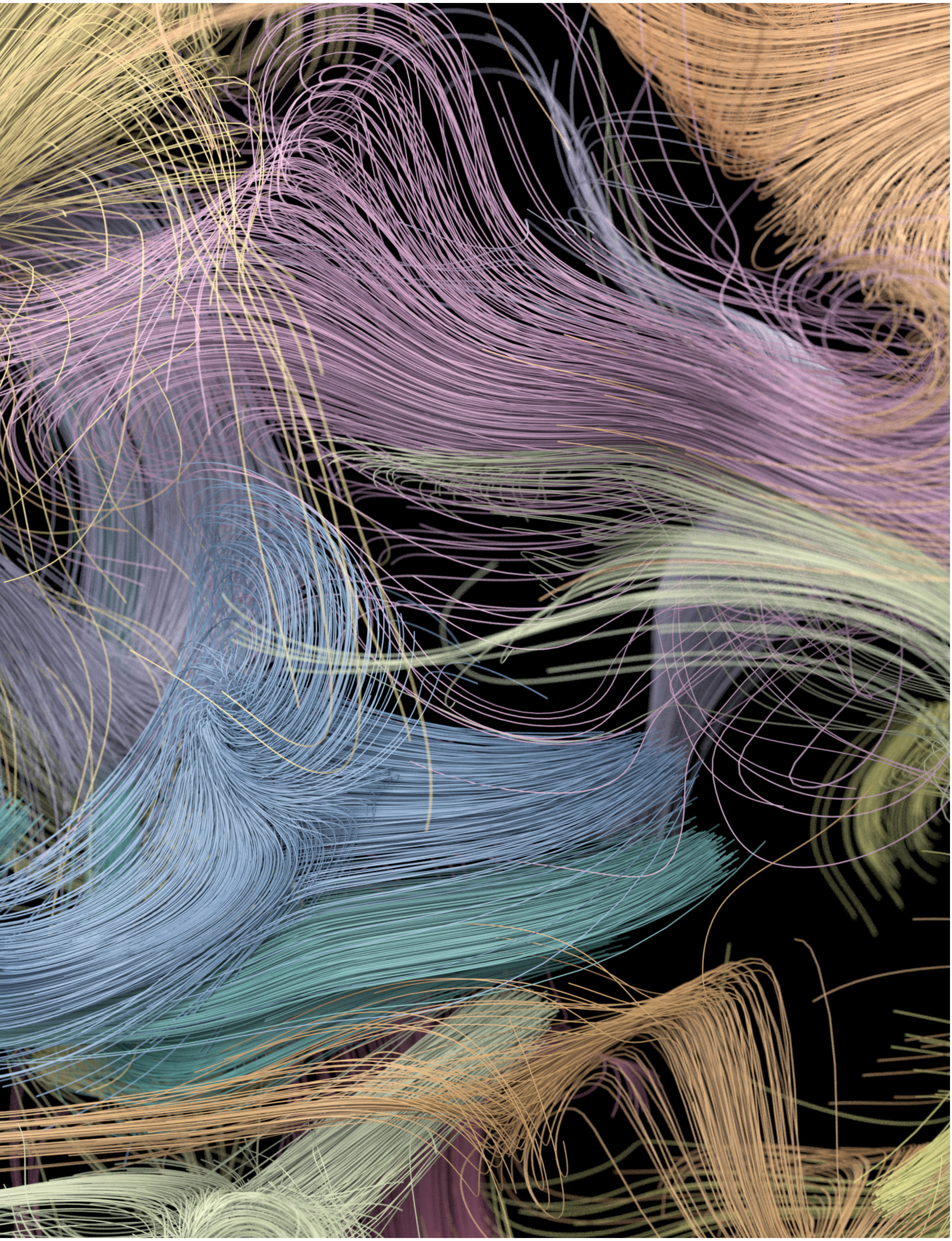


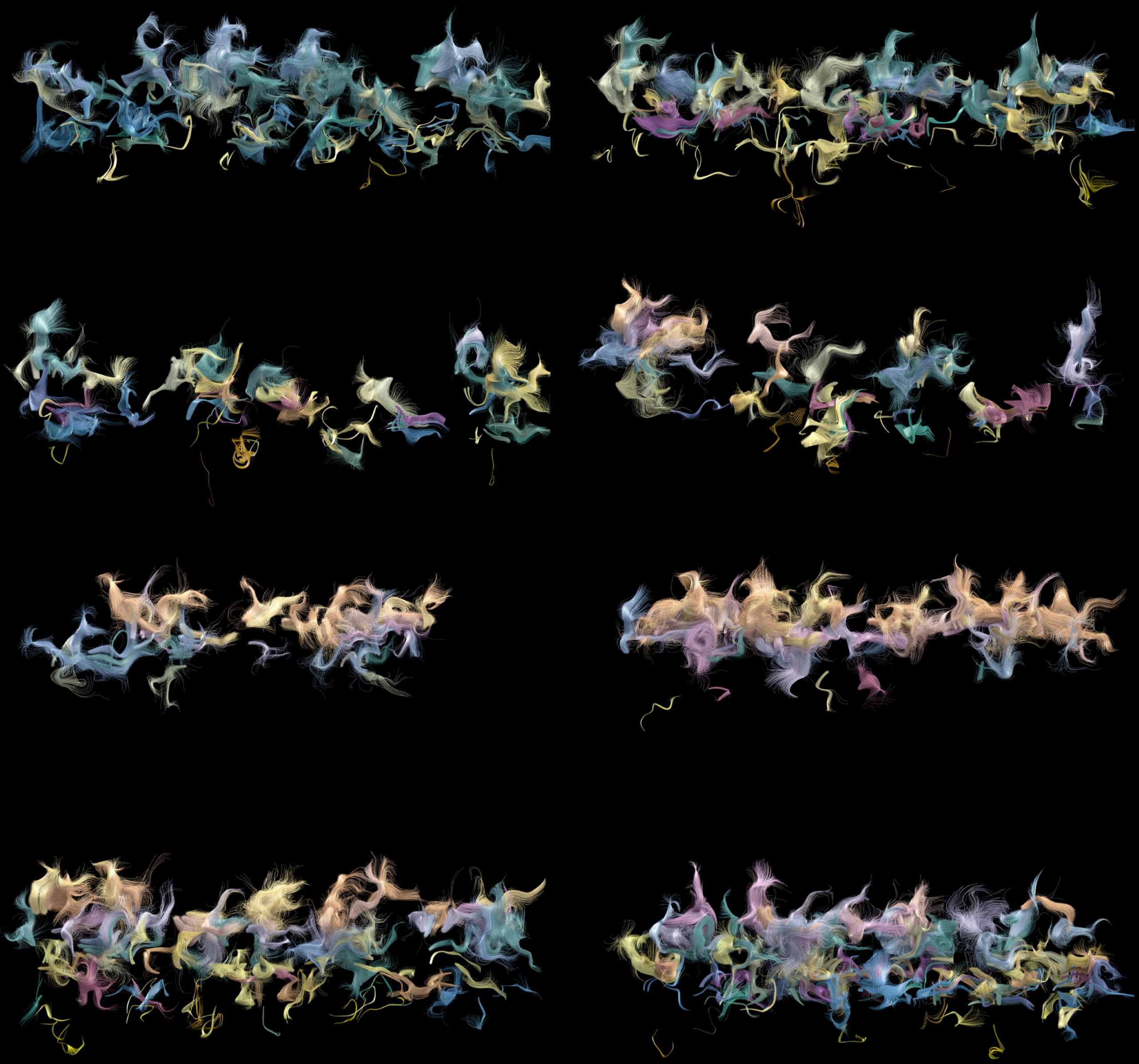
▼ **Figure 97.** *Bellbird song - detail 01.*

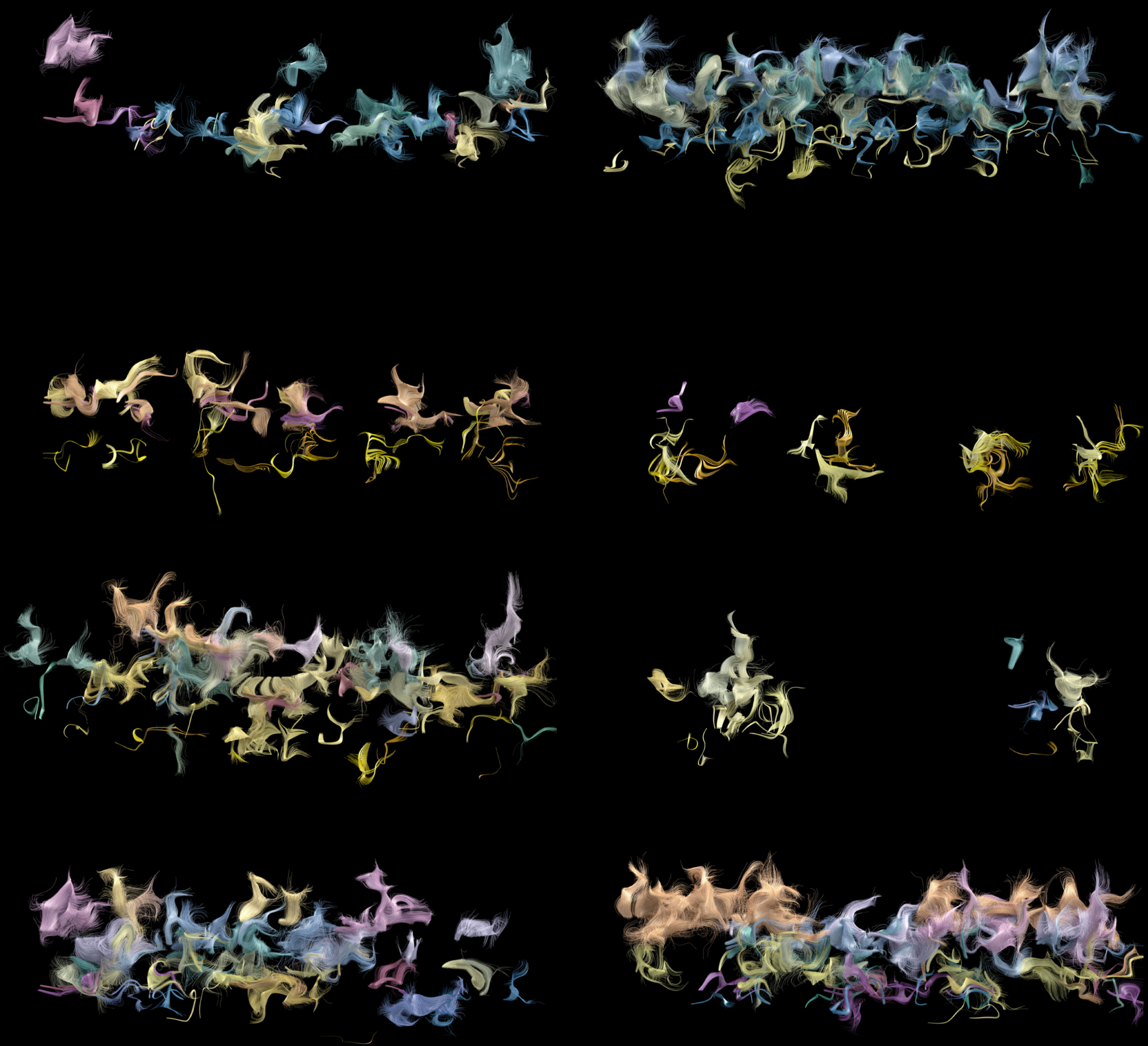


▼ **Figure 98.** *Bellbird song - detail 02.*









▲ **Figure 99.** *Bird song visualisation overall comparison.*

VII

CHAPTER SEVEN

DISCUSSION & CONCLUSION

OVERVIEW

This research project sought to answer the question:

How can sound data be used as a design tool to create an audio-visual design system capable of comparing New Zealand bird songs through an artistic format?

This research has explored how an audio-visual design system can be developed and implemented using New Zealand bird song as a design tool. It has investigated the potential of using generative design software such as Touch Designer and Houdini to create an audio-visual design system which can be used to generate visualisations which are mathematically based on the audio input signal. These visualisations can also be considered expressive audio-visual artworks in themselves.

The iterative design process led towards the development of a specific set of criteria to measure the effectiveness of the final design, which is outlined and discussed below. This criteria was developed throughout the design process, with additions being made at the end of each design phase.

- The system should generate 3D visualisations so that the resulting form can be examined from multiple angles.

- The system should be able to contrast and compare different New Zealand bird songs and generate identifiable visual differences or similarities, which have a direct relation to the difference in sound.

- Sound data of New Zealand bird song should be able to be used as an input to generate comparable forms using the final system.

- The audio-visual mapping choices should consider human's phonaesthetic understanding of audio-visual connections, such as the maluma/takete phenomenon.

- The system should be able to identify changes in pitch and have enough visual variance to represent the frequency range of New Zealand bird song.

- The movement of the visual elements should have a direct relation to the sound.

- Visual embellishments of the data should consider the appearance of New Zealand birds to allow people to draw distinct associations to the data being visualised.

- The audio-visual mapping decisions should integrate the full colour spectrum to create as much variance between different frequencies as possible.

Human's phonaesthetic understanding of sound, the history of artistic audio-visual design, as well as scientific methods for analysing sound inspired and helped inform audio-visual mapping decisions throughout various stages of the design process. Audio-visual concepts investigated within the background research such as maluma/takete and Newton's colour theory were also integrated into the design process.

The frequency-to-colour system developed using Houdini's CHOP network can successfully identify when a change in frequency occurs. To create variance between visual elements across the wide range of pitches present within New Zealand bird song, aesthetic elements such as colour, size, and shape have been manipulated using audio properties. The mel scale was used to represent perceived changes in frequency based upon human hearing. The perceived pitch of the mapped frequency was then estimated using the mel scale and assigned a colour value based on Newton's colour theory. While the frequency-to-colour system may not be entirely accurate in terms of its pitch-to-colour correspondence, it is successful in detecting changes in frequency and reflecting those changes through visual elements.

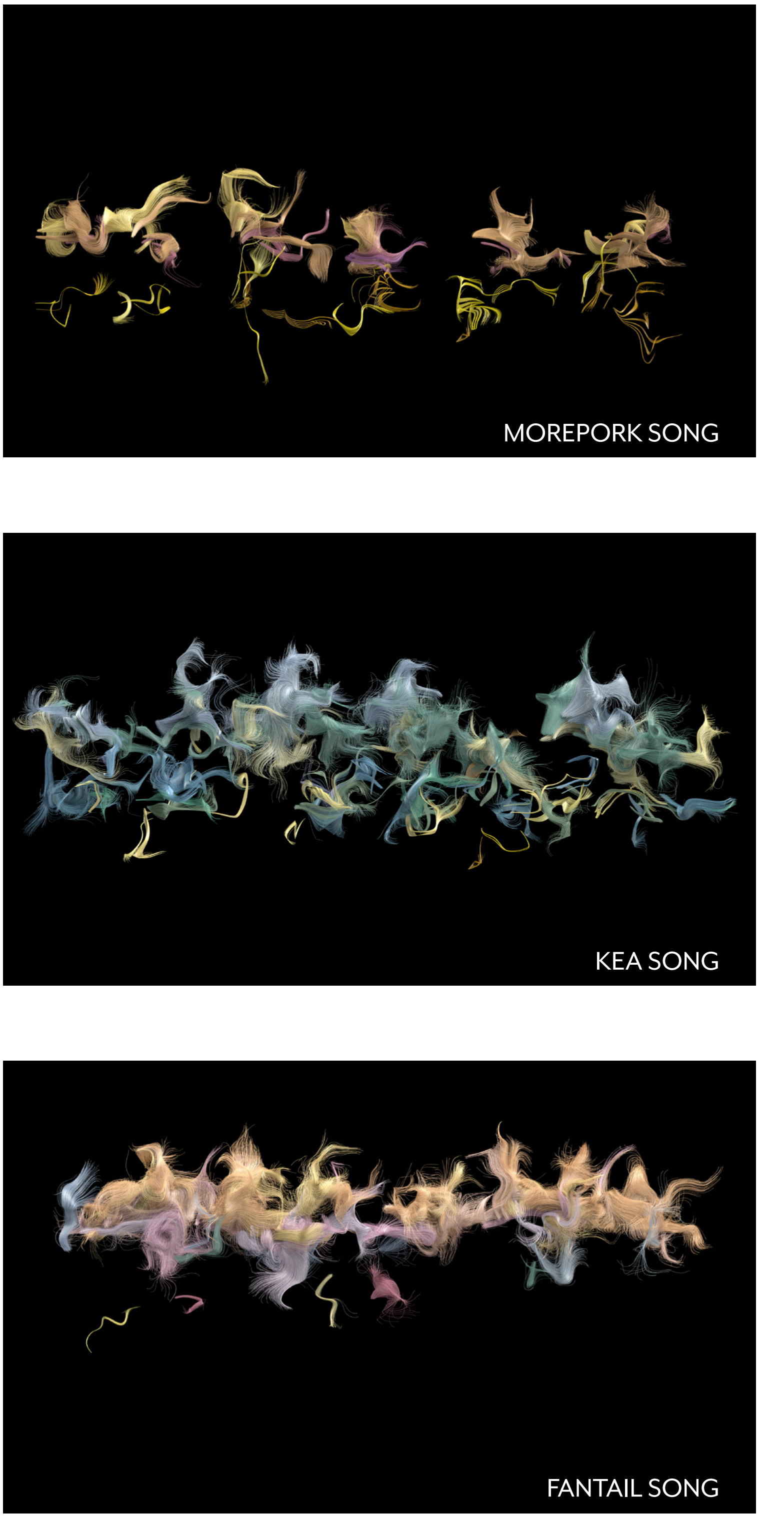
The visual representation of the design can be considered abstract, however the aesthetic qualities of the design reference the physical appearance of New Zealand birds through colour and texture. As identified by Kirk (2016), visual embellishments which are compatible with the data being visualised allow people to draw distinct associations to the data being visualised, as well as engaging audiences through visual intrigue. Following the recommendations of Kirk (2016), colour values based upon Isaac Newton's colour/pitch theory have been adapted so that each colour is based on the feathers of native New Zealand birds.

The use of generative design techniques created an opportunity to manipulate 3D geometry using audio files as a design tool. As Wang et al. (2017) explain, the use of 3D techniques allows for a more accurate representation of the energy change curve within bird songs, as well as clearer monitoring of any changes in signal energy over time. Through generating a 3D form using Houdini, it was possible to observe the audio-visualisation from multiple angles. Areas of interest could also be focused on and examined in more detail.

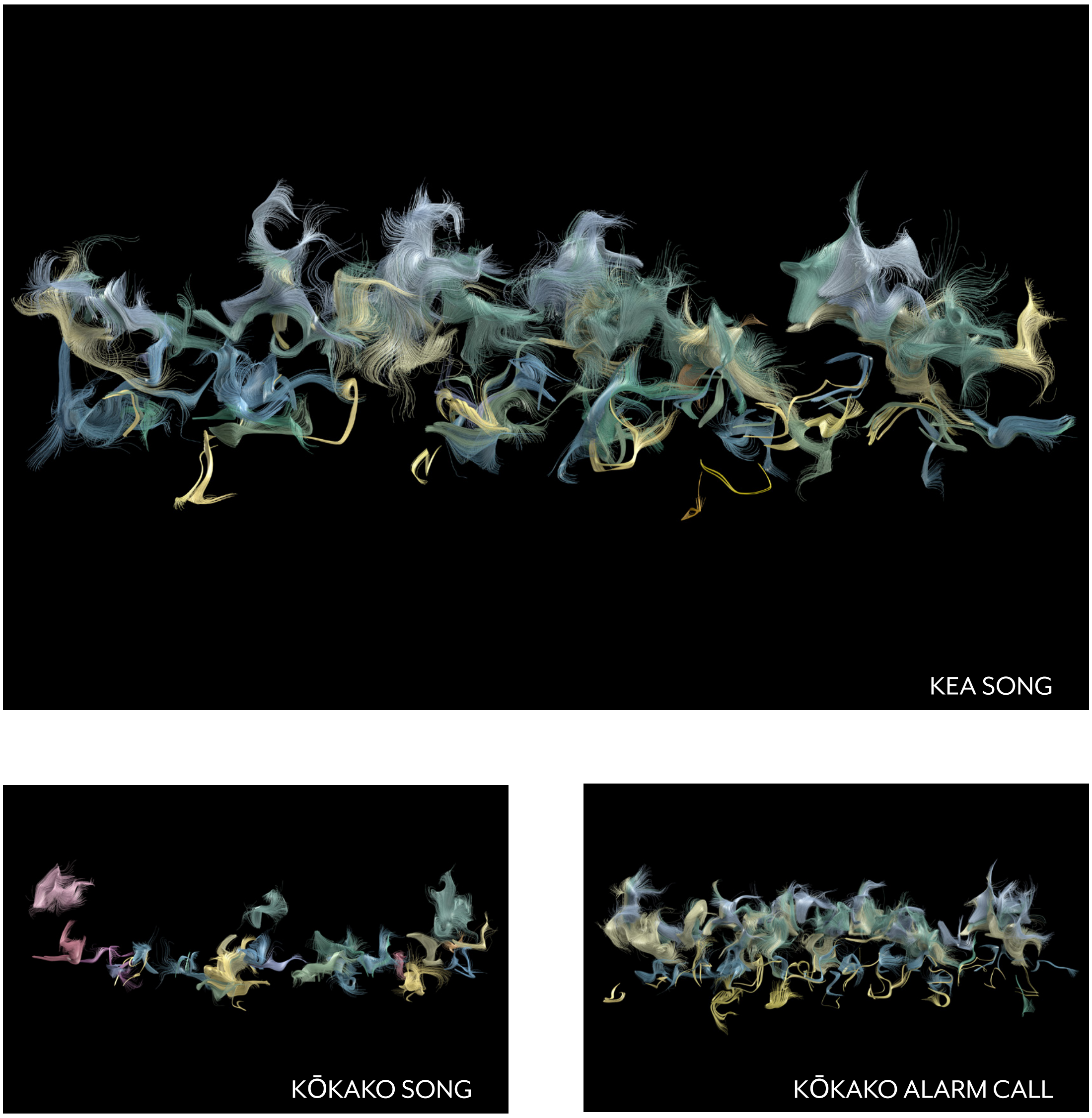
A significant amount of the research process was dedicated to the learning of a variety of generative software, specifically Touch Designer and Houdini in order to create a wider scope for the design output. Developing an initial workflow with Touch Designer, which is optimised for real-time visuals, made the process of learning Houdini easier as the software is relatively similar.

The final system can contrast and compare various New Zealand bird songs by using auditory data as a design tool. As demonstrated in Figure 100, identifiable differences can be visually observed, such as the descending mid-range song of the Kea, the repetitive lower pitched song of the Morepork, or the high pitched song of the Fantail. It also allows for the observation of anomalies within a bird song. For example, a bird song that is visualised through mostly blue and green shades may contain a different colour within a certain segment. This would indicate a particular section of the bird song which does not strictly follow the structure of the other parts of the song.

As well as comparing the songs of different species of birds, the system can also compare different songs and calls from the same species. As showcased in the previous chapter, unexpected similarities can be observed between different species of birds. For example, a Kōkako song and Kea song do not appear similar in their most familiar structure, however, Figure 101 shows that a Kōkako alarm call and a Kea song appear very similar. This similarity can also be observed by listening to each bird song, which shows the system can be used to observe auditory differences or similarities through visual identification. Kōkako's most commonly heard song is tuneful which is reflected by the individual coloured forms being separated throughout the visualisations, indicating that the song contains a large variety of notes which occur in a eloquent sequence. The Kea song, as well as the Kōkako's alarm call is an incessant "squawk" sound which has been visualised by the system in a much more chaotic way. The similarity in colour and structure indicates that the songs have an extremely similar pitch and range of frequencies.

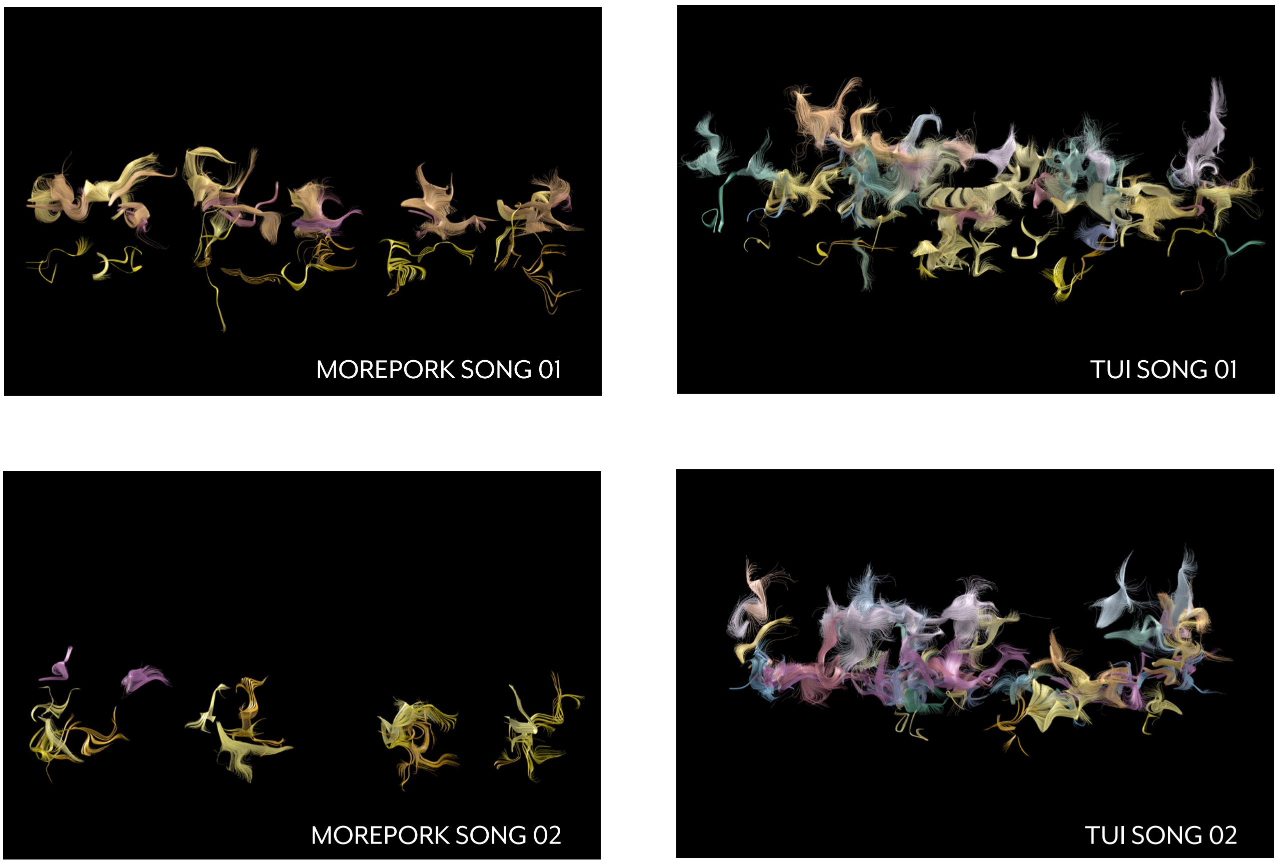


▲ **Figure 100.** *Comparison between the low frequency Morepork song, the mid range frequency Kea song, and the high frequency Fantail song. The variation in colour, form and position in space communicate these auditory differences.*



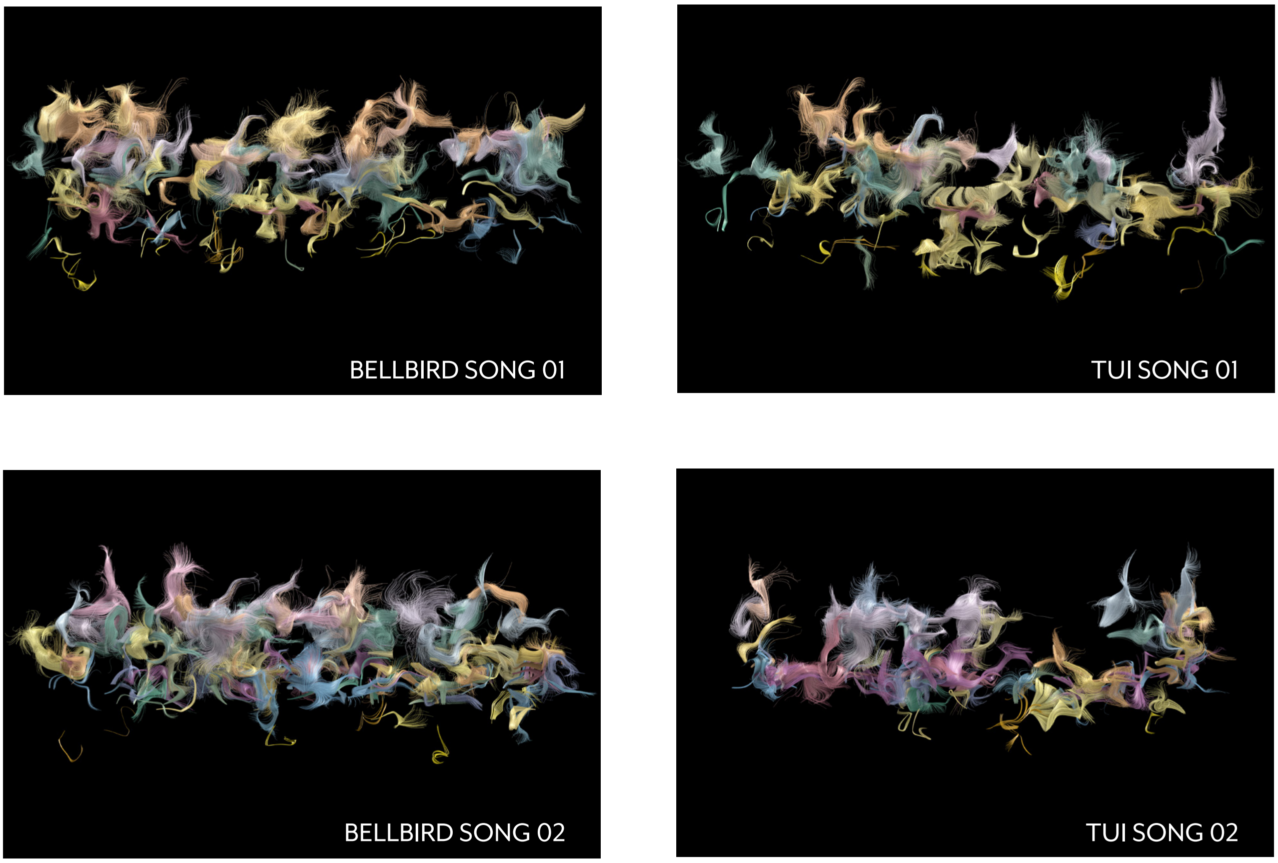
▲ **Figure 101.** *Comparison between Kea song, Kōkako song and a Kōkako alarm call. The more defined sequence of colours of the Kōkako song visualisation show that it is more tuneful when compared to the incessant "squawk" of the Kea song and Kōkako alarm call.*

The variation and complexity of various bird songs can also be observed when comparing multiple visualisations from the same species. As can be observed in Figure 102, variance can occur within the appearance of more complex songs such as the Tui, but a simpler song such as that of a Morepork always appears relatively similar. It can be visually observed that the morepork song has a much more condensed frequency range which consists of lower pitches.



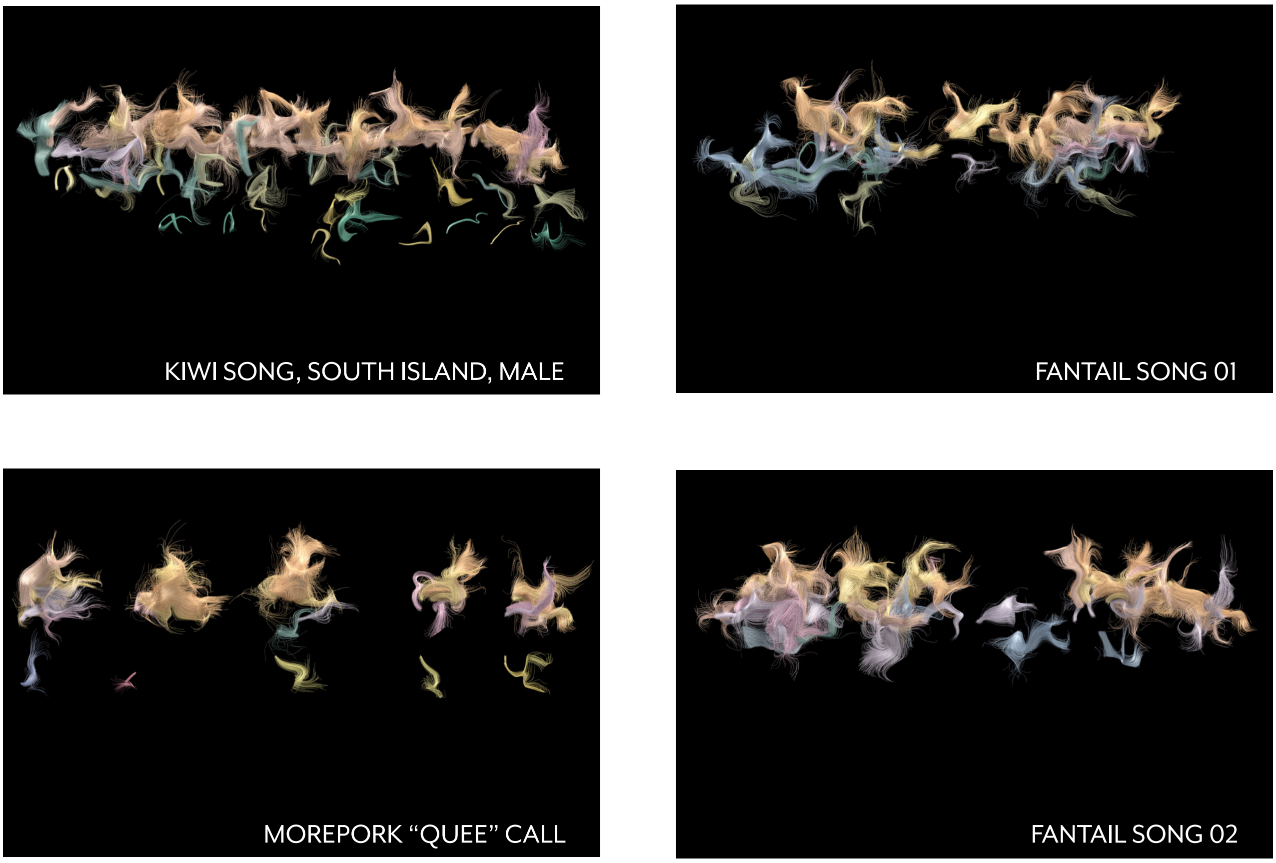
▲ **Figure 102.** *Comparison between two Morepork and two Tui songs. The larger range and variation of colours within the Tui song visualisations show that it is more chaotic compared to the Morepork song visualisations, which are in a consistent, predictable sequence.*

▼ **Figure 103.** *Comparison between two Bellbird and two Tui songs. The variety of colours within both visualisations show that they cover a large amount of the frequency range, and the forms illustrate that their structure is fairly similar.*



In comparison to this, Figure 103 shows that the song of the Tui and the Bellbird have an extremely similar structure. Both cover a large amount of space on the vertical axis, and have a large variety of colours and forms throughout the visualisation. This showcases the system working successfully, as the songs of the Tui and Bellbird are in fact extremely similar on an auditory level, and commonly get confused due to this similarity.

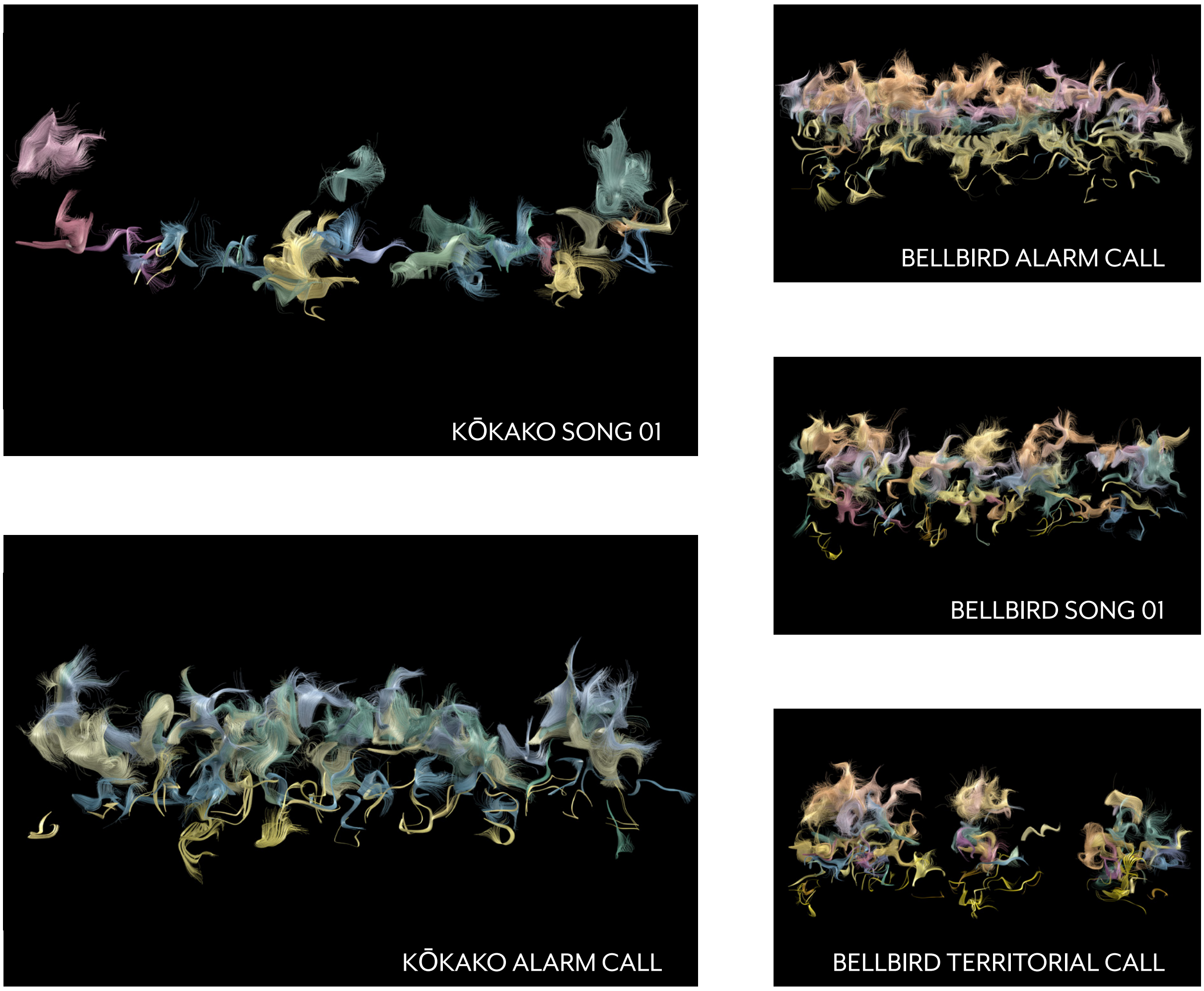
Figure 104 shows a morepork's less commonly heard "quee" call being compared to a male kiwi song, which it often gets confused with. As can be identified through visual comparison using the audio-visual design system, both calls have a very similar frequency range. As can be observed by the spaces in the form, the pattern of the morepork "quee" call has a sparser structure. As well as this, the male Kiwi song has a greater range of frequencies present throughout the call as can be identified throughout the greater variety of colour and forms. Through visual identification, it can also be observed that Fantail song also operates in a similar frequency range.



▲ **Figure 104.** *Comparison between a male Kiwi song, a Morepork "quee" call, and two Fantail songs. The similarity in colour and position in space illustrates the auditory similarity of each of these songs.*

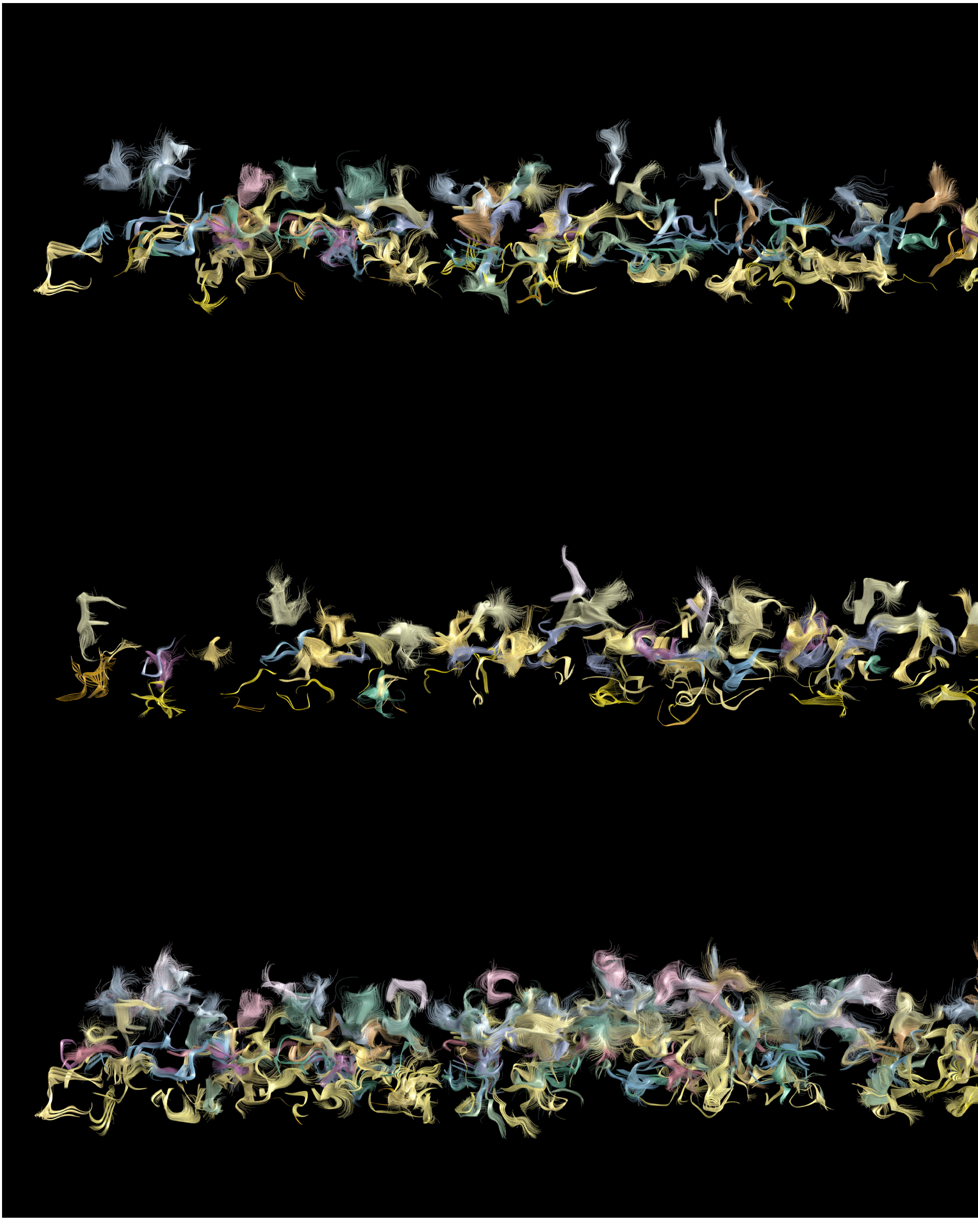
The system can also be used to identify differences between various songs and calls made by the same species. It is shown in Figure 105 that the structure of a Bellbird common song, alarm call, and territorial call are largely similar, with the territorial call being slightly sparser in it's pattern. However, the structure of a Kōkako common song is largely different to it's alarm call which has an entirely different frequency range shown by the difference in colours and form.

The audio-visual system has also been tested on recordings of a large variety of birds. Figure 106 shows a comparison of a dawn chorus recording consisting mainly of Tui vocalisations, a morning chorus consisting of Bellbird vocalisations, as well as a mixed avian choir. As can be visually identified in the images, the Tui and Bellbird choruses have much less variation than the mixed avian choir, indicating the greater variety of species present within that recording.

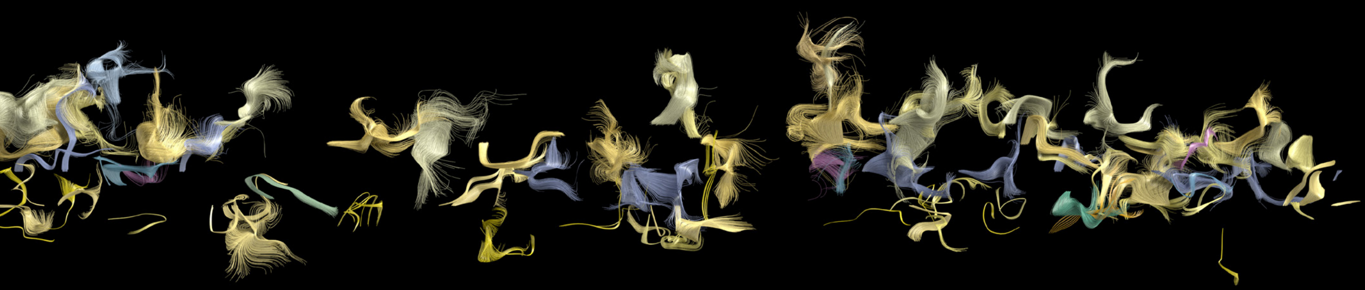


▲ **Figure 105.** *Comparison between a Bellbird song, alarm call and territorial call and a Kōkako song and alarm call. The similar colour and structure of each Bellbird visualisation illustrate that the species vocalisations are similar despite their function. In comparison to this, the Kōkako alarm call is a lot different to its most commonly heard song.*

▼ **Figure 106.** *Comparison between a dawn chorus with Tuis in the foreground, a Bellbird morning song, and a mixed avian choir. The greater variety of audible birds is illustrated within the avian choir visualisation, shown by the large range of colour and forms.*



DAWN CHORUS WITH TUIS IN FOREGROUND



BELLBIRD MORNING CHORUS

AVIAN CHOIR

IMPLICATIONS

The workflow developed throughout this research project provides a new method for visualising sound data. Using the audio-visual design system developed throughout this research, bird song sound data can be examined in an innovative, artistic visual format which allows distinct audio characteristics to be observed. The most communicable visualisations of sound in the past have concentrated on measuring volume levels over time, most commonly seen in scientific visualisations such as sonograms and spectrograms. This research project has the potential to transform the way we consider and visualise sound data through more artistic and communicable methods. This is intended to encourage further consideration towards how sound can be visually translated using a combination of scientific and artistic methods, while integrating human's phonaesthetic understandings of sound.

LIMITATIONS

Computing Power

There were technical limitations to this research, due to the significant amount of computer processing power required in order to read information from audio files and generate animated 3D geometry simultaneously using Houdini. Adjustments to the design system were required to optimise the iteration process. A considerable amount of the research process was dedicated to developing optimisation strategies so that the final output could be completed.

Audio Manipulation

The way the trigger system functions within Houdini meant that visuals could only be generated once the volume level reached a certain defined threshold. This meant that audio files had to be pre-processed before being used to generate visuals, to ensure that they all met appropriate volume levels. Background noise reduction was also applied to the files so that other auditory elements would not interfere with the signal of the bird song. Although audio normalisation can be achieved using Houdini's CHOP functions, it could not be relied upon to pre-process audio files in real-time as this would create the need for more computing power, as well as not being as consistent as programs solely intended for audio editing such as Audacity. As technology develops in the future, the capabilities of generative software for more specialised functions are likely to improve as audio-visualisation methods continue to develop within mainstream culture.

Subjectivity

Although this research project had a focus on the perceptual mapping of audio-visual elements which integrated people's phonaesthetic associations between visual and sonic elements, these associations can be largely subjective and dependent on the individual person. Although there is no definitive method for visually translating sound so that it is objectively understandable, the final design system uses a mathematical basis for comparing and examining sound data through using the raw input signal as a design tool. This allows for distinct elements of bird songs to be compared against each other, while remaining representative of the original data.

FUTURE RESEARCH

Artificial Intelligence

The workflow developed throughout this research project could be combined with artificial intelligence processes to further develop the area of audio categorisation and recognition.

For example, the system could be adapted to assign visual parameters to categorised bird calls and songs to differentiate between territorial calls, mating calls, and other types of communication between birds. This could be integrated into the development of a visual interface for artificial intelligence sound recognition systems. These are often limited in their visual output despite being incredibly useful for various environmental purposes, such as identifying the location and species of birds through analysing audio data.

Education

The integration of artificial intelligence could also lead towards the development of the design system to be used as a sustainable educational tool. It could allow people to learn about different bird species and what calls they make, as well as what is being communicated through that song through distinct visual association. Furthermore, the developed audio-visual system could be used for helping to annotate and describe audio-visual properties within the deaf community. There may need to be more development in the clarity of transcription to make the system suitable for this application.

Tourism

The visual appeal of this research project could also lend itself to a tourism campaign advertising the beauty of New Zealand bird song, which could encourage tourists to visit New Zealand and encourage the growth of the local economy. The audio-visual artworks could be exhibited in gallery and public spaces, either as prints, 3D sculptures, or video installations.

3D Printing

The workflow developed throughout this research could also be adapted for the generation of 3D forms and visual embellishments, using sound data as a design tool. The geometry files created within this research project could easily be exported from Houdini and used within a variety of other softwares for other applications such as 3D printing. This could be an interesting experiment as it would allow viewers to observe the 3D form of a sound through the observation of a tangible object.

CONCLUSION

This research project was guided by the question:

> ***How can sound data be used as a design tool to create an audio-visual design system capable of comparing New Zealand bird songs through an artistic format?***

Through the implementation of a generative audio-visual design system, sound data can be input and compared through artistic 3D audio-visualisations. These visualisations implement human's phonaesthetic understanding of audio-visual connections. The use of generative software such as Houdini and Touch Designer allowed for the exploration of how visual parameters could be modified to effectively represent sound through visual mapping parameters such as form, colour and movement.

The specific focus on New Zealand bird song allowed for the optimisation of the system so that it was able to visually compare a defined dataset. The final design output outlines that the overarching frequency range and pattern of various New Zealand bird songs can be visually identified through colour, form, and position in space. The use of generative design techniques throughout the design process means through the use of the final audio-visual system, audio files can easily be swapped between and visually compared.

This research project has the potential to transform the way we consider and visualise sound data through more artistic and communicable methods. Through using sound data as a design tool, this research project has resulted in a design system capable of generating artistic audio-visualisations which can also be considered a mathematical reflection of the data. This research has investigated interdisciplinary approaches to make data more accessible to non-specialists by integrating a variety of both scientific and artistic visualisation techniques throughout the design process. The use of these methods for sound analysis holds opportunities for future innovation and collaboration between disciplines. In conclusion, this research project has determined that sound data can be used as a design tool to develop a generative audio-visual design system which can effectively analyse and compare a defined dataset such as New Zealand bird song.

REFERENCE LIST

Agkathidis, A. (2015). Generative design methods implementing computational techniques in undergraduate architectural education. *ECAADE 2015: Real time - Extending the Reach of Computation, 2*, 47-55. https://doi.org/10.13140/RG.2.1.3767.1120

Anadol, R. (2020). Synaesthetic architecture: A building dreams. *Architectural Design, 90*(3), 76–85. https://doi.org/10.1002/ad.2572

Arfini, M. (2013). Abstract film as viewable music: Early experiments of Hans Richter, Walther Ruttmann and Oskar Fischinger. *Music in Art, 38*(1-2), 213-222. https://www.jstor.org/stable/musicinart.38.1-2.213

Ars Electronica. (2018). *Bird song diamond.* https://birdsongdiamond.com/ars-electronica/

Azar, J., & Bell, B. (2016). Acoustic features within a forest bird community of native and introduced species in New Zealand. *Emu - Austral Ornithology, 116*(1), 22-31. https://doi.org/10.1071/MU14095

Badyaev, A. V., & Leaf, E. S. (1997). Habitat associations of song characteristics in Phylloscopus and Hippolais Warblers. *AUK, 114,* 40–46. https://doi.org/10.2307/4089063

Barclay, A. (2020). *Making census data open, accessible and useful for real people.* Medium. https://medium.com/nightingale/making-census-data-open-accessible-and-useful-for-real-people-f3ffb167913a

Barnett, M., Vaughn, M. H., Strauss, E., & Cotter, L. (2011). Urban environmental education: Leveraging technology and ecology to engage students in studying the environment. *International Research in Geographical and Environmental Education, 20*(3), 199-214. https://doi.org/10.1080/10382046.2011.588501

Bassot, B. (2015). *The reflective practice guide: An interdisciplinary approach to critical reflection.* Routledge. https://doi.org/10.4324/9781315768298

Bertelli, S., & Tubaro, P. L. (2002). Body mass and habitat correlates of song structure in a primitive group of birds. *Biological Journal of the Linnean Society, 77,* 423–430. https://doi.org/10.1046/ j.1095-8312.2002.00112.x

Blake, J. B. (2016). Sound symbolism in English: Weighing the evidence. *Australian Journal of Linguistics, 37*(3), 286-313. https://doi.org/10.1080/07268602.2017.1298394

Brumm, H., & Slater, P. J. B. (2006). Animals can vary signal amplitude with receiver distance: Evidence from Zebra Finch song. *Animal Behaviour, 72*, 699–705. https://doi.org/10.1016/j.anbehav.2006.01.020

Card, S., Mackinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think.* Morgan Kaufmann Publishers Inc.

Carrier, S., Scott, C., & Hall, D. (2012). Soundsational science. *Science Activities, 49*(1), 1–6. https://doi.org/10.1080/00368121.2011.566643

Cody, M., Cohen D., Sasahara, K., & Taylor, C. (2012). Structural design principles of complex bird songs: A network-based approach. *PLOS ONE, 7*(9). https://doi.org/10.1371/journal.pone.0044436

Derivative. (n.d.). *Touch Designer.* https://derivative.ca/

Dewey, J. (1934). *Art as experience.* Minton, Balch & Company.

Farina, A., Pieretti, N., & Piccioli, L. (2011). The soundscape methodology for long-term bird monitoring: A Mediterranean Europe case-study. *Ecological Informatics, 6*. https://doi.org/10.1016/j.ecoinf.2011.07.004

Fischinger, O. (1932). Sounding ornaments. In W. Moritz, *Optical poetry: The life and work of Oskar Fischinger* (pp. 179–80). Indiana University Press.

Frankel, L., & Racine, M. (2010) The complex field of research: For design, through design, and about design. In Durling, D., Bousbaci, R., Chen, L, Gauthier, P., Poldma, T., Roworth-Stokes, S., & Stolterman, E. (2010, July 7-9). *Design and complexity* [Conference presentation]. DRS International Conference 2010. Montreal, Canada. https://dl.designresearchsociety.org/drs-conference-papers/drs2010/researchpapers/43

Gross, B., Bohnacker, H., Laub, J., & Lazzeroni, C. (2018). *Generative design revised: Visualize, program, and create with javascript in p5.js*. Princeton Architectural Press.

Hanington, B., & Martin, B. (2012). *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions.* Rockport Publishers.

Inal, A. (2021). *Alican Inal: Work*. https://alicaninal.com

Inal, A. (2020). *Sonic architecture: Exploring the relationship between spaces and sound. Populous.* https://populous.com/sonic-architecture-exploring-the-relationship-between-spaces-and-sound

Innes, J. (2013). North Island Kokako. In Miskelly, C. M. (ed.), *New Zealand birds online*. www.nzbirdsonline.org.nz

Kandinsky, V. (1935). *Succession* [Painting]. Philips Collection, Washington, DC, United States. https://www.wikiart.org/en/wassily-kandinsky/succession-1935

Kemp, J. (2013). Kea. In Miskelly, C. M. (ed.), *New Zealand birds online*. www.nzbirdsonline.org.nz

Keulemans, J., & Buller, W. (1888). *A history of New Zealand birds*. (n.p.). https://doi.org/10.5962/bhl.title.128503

Kirk, A. (2016). *Data visualisation : A handbook for data driven design*. Sage Publications.

Köhler, W. (1929). *Gestalt psychology.* Liveright.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development.* Prentice Hall.

Lee, I., & McDonald, K. (2014). Mother. *Leonardo, 47*(4), 410-411. https://www.muse.jhu.edu/article/552313

Levin, G. (2009, February). *Art that looks back at you* [Video]. TED Conferences. https://www.ted.com/talks/golan_levin_art_that_looks_back_at_you

Levin, G., Lieberman, Z., Blonk, J., & Barbara, J. L. (2003). *Performances of Messa di Voce.* https://www.flickr.com/photos/tmema/sets/72157594393980031/

Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees and humans. *Proceedings of the National Academy of Sciences of the United States of America, 108*(51), 20661–20665. https://doi.org/10.1073/pnas.1112605108

Maher, R. (2011). Automated analysis and interpretation of long-term soundscape audio recordings. *Journal of The Acoustical Society of America, 129*. https://doi.org/10.1121/1.3588483

McCarthy, B. (2014). The modern audio analyzer: Understanding dual-channel FFT audio analyzers. *Sound & Video Contractor, 32*(3), 58. https://go.gale.com/ps/i.do?p=ITOF&u=vuw&id=GALE%7CA370754054&v=2.1&it=r

McPherson, L. (n.d.). The birds [Audio recordings]. https://www.rnz.co.nz/collections/birds

Milone, F., & Camarda, D. (2017). Modeling knowledge in environmental analysis: A new approach to soundscape ecology. *Sustainability, 9*(4). https://doi.org/10.3390/su9040564

Milton, A., & Rodgers, P. (2013). *Research methods for product design.* Laurence King Publishing. https://ebookcentral-proquest-com.helicon.vuw.ac.nz/lib/vuw/detail.action?docID=5741494

Moorhouse, R. J. (2013). Kaka. In Miskelly, C. M. (ed.), *New Zealand birds online*. www.nzbirdsonline.org.nz

Moran, S., McCloy, D., & Wright, R. (2014). *PHOIBLE online.* http://phoible.org/

Morton, E. S. (1975). Ecological sources of selection on avian sounds. *American Naturalist, 109*, 17–34. https://doi.org/10.1086/282971

Nair, P. (2018). *The dummy's guide to MFCC.* Medium. https://medium.com/prathena/the-dummys-guide-to-mfcc-aceab2450fd

Obe, A. (2013). Data visualisation and beyond: A multi-disciplinary approach to promote user engagement with official statistics. *Statistical Journal of the IAOS, 29*(3), 173–185. https://doi.org/10.3233/SJI-130783

Ovington, J. D., & Pryor, L. D. (1983). Temperate broad-leaved evergreen forests of Australia. In Temperate broad-leaved evergreen forests. *Ecosystems of the World, 10*, 73-101.

Perron, J. (2004). *Golan Levin, Messa di Voce.* https://www.fondation-langlois.org/html/e/page.php?NumPage=364

Powlesland, R. G. (2013). New Zealand Fantail. In Miskelly, C. M. (ed.), *New Zealand birds online*. www.nzbirdsonline.org.nz

Processing. (n.d.). *Processing.* https://processing.org/

Ramachandran, V., & Hubbard, E. (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies, 8*(12), 3-34.

Ramírez, E. R. (2017). A postgraduate thesis model for research through design based on design criteria. *The International Journal of Designed Objects, 11*(4), 11-27. https://doi.org/10.18848/2325-1379/CGP/v11i04/11-27

Reas, C., McWilliams, C., & Barendse, J. (2010). *Form + code in design, art, and architecture.* Princeton Architectural Press.

Robertson, H. A. (2013). North Island Brown Kiwi. In Miskelly, C. M. (ed.), *New Zealand birds online*. www.nzbirdsonline.org.nz

Robertson, H. A. (2013). Tui. In Miskelly, C. M. (ed.), *New Zealand birds online*. www.nzbirdsonline.org.nz

Rogers, H. (2013). *Sounding the gallery: Video and the rise of art-music*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199861408.001.0001

Safe Software. (n.d.). *Maximize the value of data*. https://www.safe.com/

Sagar, P. M. (2013). Bellbird. In Miskelly, C. M. (ed.), *New Zealand birds online*. www.nzbirdsonline.org.nz

Sasahara, K., Cody, M., Cohen, D. & Taylor, C. (2012). Structural design principles of complex bird songs: A network-based approach. *PLOS ONE, 7*(9). https://doi.org/10.1371/journal.pone.0044436

Seaton, R., & Hyde, N. (2013). Morepork. In Miskelly, C.M. (ed.), *New Zealand birds online*. www.nzbirdsonline.org.nz

Shaffer, D. (2004). Pedagogical praxis: The professions as models for postindustrial education. *Teachers College Record, 106*, 1401-1421. https://doi.org/10.1111/j.1467-9620.2004.00383.x

SharkD. (2008). *Colour solid comparison* [Image]. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Color_solid_comparison_hsl_hsv_rgb_cone_sphere_cube_cylinder.png

SideFX. (n.d.). *Houdini - 3D modeling, animation, VFX, look development, lighting and rendering.* https://www.sidefx.com/

Silva, M. (2017). *How colours matter to philosophy.* Springer. https://doi.org/10.1007/978-3-319-67398-1

Slabbekoorn, H., & Peet, M. (2003). Birds sing at a higher pitch in urban noise. *Nature, 424,* 267. https://doi.org/10.1038/424267a

Stowell, D., & Plumbley, M. D. (2014). Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *PeerJ*. https://doi.org/10.7717/peerj.488

Styles, S., & Gawne, L. (2017). When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter. *i-Perception, 8*(4), 1-17. https://doi.org/10.1177/2041669517724807

The Department of Conservation. (n.d.). *Bird songs and calls* [Audio recordings]. https://www.doc.govt.nz/nature/native-animals/birds/bird-songs-and-calls/

Thibeault M. D. (2011). Learning from looking at sound: Using multimedia spectrograms to explore world music. *General Music Today, 25*(1), 50-55. https://doi.org/10.1177/1048371311414050

Truslow, H. (2014). *Bird songs vs calls.* https://blogs.massaudubon.org/yourgreatoutdoors/bird-songs-vs-calls/

Tubaro, P. L., & Lijtmaer, D. A. (2006). Environmental correlates of song structure in forest grosbeaks and saltators. *Condor, 108*, 120–129. https://doi.org/10.1650/0010-5422(2006)108[0120:ECOSSI]2.0.CO;2

Vesna, V. (2018). *Bird song diamond.* https://birdsongdiamond.com

Vesna, V., Campbell, B., & Samsel, F. (2019). Victoria Vesna: Inviting meaningful organic art-science collaboration. *IEEE Computer Graphics and Applications, 39*(4), 8-13. https://doi.org/10.1109/MCG.2019.2916962

Wang, X., Shi, X., Yang, D., & Zhou, Y. (2017, May 7-9). *Research on the application of 3D spectrogram in bird tweet and speech signals* [Conference presentation]. 2017 29th Chinese Control And Decision Conference (CCDC). Chongqing, China. https://doi.org/10.1109/CCDC.2017.7978597

Ward, J. (2013). Synesthesia. *Annual Review of Psychology, 64*(1), 49–75. https://doi.org/10.1146/annurev-psych-113011-143840

Zhang, Y., Pan, Y., & Zhou, J. (2018). Study on application of audio visualization in new media art. *Journal Of Physics: Conference Series, 1098*. https://doi.org/10.1088/1742-6596/1098/1/012003

Zimmerman, J., & Forlizzi, J. (2014). Research through design in HCI. In J. S. Olson & W. A. Kellogg, *Ways of knowing in HCI* (pp. 167–189). Springer New York. https://doi.org/10.1007/978-1-4939-0378-8_8

SYNAESTHETIC SANCTUARY

WILLIAM SHAW

2021