

ARIZONA STATE UNIVERSITY

HONORS THESIS

**D.I.Y. Smartcube:
Tracking the Face Turns of a Rubik's Cube
using Embedded Speakers**

Author:
Joseph HALE

Thesis Director:
Dr. Robert HEINRICHES

Second Reader:
Dr. Baoxin LI

*A thesis submitted in partial fulfillment of the
requirements for an honors distinction
in the*

Fulton Schools of Engineering
Barrett, The Honors College

April 7, 2022

COPYRIGHT NOTICE

'DIY Smartcube' by Joseph Hale is licensed under the terms of the Mozilla Public License, v 2.0, which are available at <https://mozilla.org/MPL/2.0/>.

You can download the source code for 'DIY Smartcube' for free from:

<https://github.com/jhale1805/DIY-Smartcube>.

Declaration of Authorship

I, Joseph HALE, declare that this thesis titled, "D.I.Y. Smartcube: Tracking the Face Turns of a Rubik's Cube using Embedded Speakers" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: JOSEPH HALE

Date: 16 MARCH 2022

*If you are curious, you'll find the puzzles around you.
If you are determined, you will solve them.*

Erno Rubik

ARIZONA STATE UNIVERSITY

Abstract

Dr. Robert Heinrichs

Barrett, The Honors College

Honors Thesis in Software Engineering

D.I.Y. Smartcube:

Tracking the Face Turns of a Rubik's Cube using Embedded Speakers

by Joseph HALE

Speedsolving, the art of solving twisty puzzles like the Rubik's Cube as fast as possible, has recently benefitted from the arrival of smartcubes which have special hardware for tracking the cube's face turns and transmitting them via Bluetooth. However, due to their embedded electronics, existing smartcubes cannot be used in competition, reducing their utility in personal speedcubing practice.

This thesis proposes a sound-based design for tracking the face turns of a standard, non-smart speedcube consisting of an audio processing receiver in software and a small physical speaker configured as a transmitter. Special attention has been given to ensuring that installing the transmitter requires only a reversible centercap replacement on the original cube. This allows the cube to benefit from smartcube features during practice, while still maintaining compliance with competition regulations.

Within a controlled test environment, the software receiver perfectly detected a variety of transmitted move sequences. Furthermore, all components required for the physical transmitter were demonstrated to fit within the centercap of a Gans 356 speedcube.

Acknowledgements

The first order of gratitude goes to my wonderful wife. Though this project has consumed much of my attention throughout our entire friendship, courtship, and marriage, she has been my greatest source of encouragement and support.

Additional gratitude goes to my parents, particularly my father-in-law who spent hours with me at the workbench guiding me as I pieced together transmitter prototypes.

Finally, gratitude goes to Dr. Heinrichs for his patience through weekly meetings for a two-year project originally scheduled for only one year. His guidance has been instrumental throughout the entire project. Dr. Li's infusion of confidence from the start of the project was also inspiring.

Contents

Declaration of Authorship	iii
Abstract	vii
Acknowledgements	ix
1 Introduction	1
1.1 The Advent of Smartcubes	1
1.2 Obstacles to Adoption	1
1.3 Purpose of this Thesis	2
1.4 Thesis Overview	2
2 Background	3
2.1 A Brief History of the Rubik's Cube	3
2.2 The Anatomy of a Rubik's Cube	3
2.2.1 Move Sequence Notation	5
2.3 A Brief History of Speedsolving	6
2.3.1 The World Cube Association	6
2.3.2 The Rise of Smartcubes	7
3 State of the Art	9
3.1 Introduction	9
3.2 Commercial Smartcubes	9
3.2.1 Giiker Cube	9
3.2.2 Go Cube	10
3.2.3 Gans 356i	10
3.3 Academia	11
3.3.1 Computer Vision	11
Sticker Color Classification	11
Measuring a Face's Angle of Rotation	11
Classification of Single Moves and Entire Move Sequences	12
3.3.2 Magnetic Resonance	12
3.3.3 Muscle-Tracking Armband	13
3.4 Other Relevant Research	13
3.4.1 Sound	13
3.4.2 Radio Frequency Identification (RFID)	13
3.4.3 Off-Axis Magnetic Angle Sensors	14
3.5 Research Questions	14

4 The Protocol	17
4.1 Introduction	17
4.2 Requirements	17
4.2.1 Signal-to-Noise Ratio	17
How to Read a Spectrogram	18
Key Observations from the Spectrogram	18
4.2.2 Tone Distinctiveness	18
4.2.3 Frequency Response Range of Consumer Hardware	21
4.2.4 Human Auditory Range	22
4.3 Specification	22
4.3.1 Representing the Cube's Current State	22
4.3.2 Tracking Face Turns	23
4.3.3 Conveying State Through Sound	23
4.4 Alternative Protocol	24
4.5 Summary	25
5 The Receiver	27
5.1 Introduction	27
5.2 Creating Synthetic Audio	27
5.2.1 Representing the Audio Protocol	28
5.2.2 Representing the Rubik's Cube	29
5.2.3 Creating Synthetic Audio for an Arbitrary Algorithm	29
5.3 Decoding the Synthetic Audio	31
5.3.1 Computing the Spectrogram	31
The Brief Overview of the Math behind the Spectrogram	32
5.3.2 Extracting the Dominant Component Frequencies	33
5.3.3 Translating Component Frequencies to Centerpiece States	34
5.3.4 Extracting Move Sequences from Centerpiece State Sequences	36
5.3.5 Full Example: Extracting Moves from Synthetic Audio	37
5.4 Creating Realistic Audio	38
5.5 Decoding Realistic Audio	38
5.5.1 Fine-Tuning the Threshold	39
5.5.2 Filtering through Similar Peak Frequencies	41
5.5.3 Ignoring Noise when Extracting Move Sequences	42
5.5.4 Optimizing Algorithm Parameters	43
5.6 Summary	43
6 The Transmitter	45
6.1 Introduction	45
6.2 Requirements	45
6.2.1 Precision of Tone Generation	45
6.2.2 Responsiveness to Face Turns	46
6.2.3 Signal-to-Noise Ratio	46
6.2.4 Prospects of Miniaturization	46
6.3 Design	47
6.3.1 The 555 Timer	47
6.3.2 Creating Audio	48

6.3.3 Choosing the values of R and C	48
6.4 Prototyping	50
6.5 Minimizing Sound Obstruction	52
6.6 Miniaturization	52
6.7 Summary	53
7 Evaluation	55
7.1 Move Tracking Accuracy	55
7.1.1 Influence of Cube Noisiness	56
7.1.2 Influence of the Standard Deviation Threshold	58
7.1.3 Influence of Minimum Threshold	58
7.1.4 Influence of Window Size	58
7.2 Compatibility with Standard Speedcubes	62
7.3 Move Tracking Granularity	62
7.4 Competition Legality	63
7.5 Summary	63
8 Conclusion	65
8.1 Answering Research Questions	65
8.1.1 Feasibility on Consumer Hardware	65
8.1.2 Move Tracking Accuracy	66
8.1.3 Compatibility with Standard Speedcubes	66
8.1.4 Move Tracking Granularity	67
8.1.5 Competition Legality	67
8.2 Limitations	67
8.3 Outlook	68
A Longer Code Snippets	71
A.1 Spectrogram Generation for Figure 5.5	71
A.2 Spectrogram Generation for Figure 5.7	72
Bibliography	73

List of Figures

2.1	The Anatomy of the Rubik's Cube [12]	4
2.2	Rubik's Cube Notation	5
2.3	The move sequence R U2 R'	6
3.1	The internal components of the Giiker Cube	10
3.2	The internal components of the Go Cube [22]	10
3.3	Gans 356i Teardown	11
3.4	The IM3D technology as used in the Cube Harmonic	12
3.5	Off-Angle Magnetic Sensor Calculations	14
4.1	Background Noise of a Quiet Room while Speedsolving	19
4.2	Distinctiveness of Two Increasingly Similar Tones	21
4.3	"A typical frequency response curve for a microphone." [40]	22
4.4	The four possible rotations of a face of a Rubik's Cube [43]	23
5.1	Centerpiece State to Frequency Mapping	28
5.2	A simple abstraction of a Rubik's Cube	29
5.3	Generating audio for any Rubik's Cube algorithm	30
5.4	Example Audio Generation for a Rubik's Cube Algorithm	30
5.5	Spectrogram of the synthetic audio created in Section 5.2.3	31
5.6	Function to compute the spectrogram of a .wav file	32
5.7	Spectrogram from Figure 5.5 with a threshold at 85% of the maximum strength of the component frequencies.	33
5.8	Extracting dominant frequencies from one time step of audio	33
5.9	Example peak frequency decoding at a specific time step	34
5.10	Converting peak frequencies to centerpiece states	34
5.11	Example conversion of peak frequencies to states	35
5.12	Function to list the cube's state at each time step	35
5.13	Example listing of states over time	35
5.14	Code to extract move sequences from state sequences	36
5.15	Example move sequence extraction	37
5.16	Full Example: Audio generation and move extraction	37
5.17	Spectrogram of a realistic audio sample	38
5.18	An 85% threshold applied to realistic audio	39
5.19	Dominance of background noise between moves	39
5.20	Updated version of threshold computation in Figure 5.8a	40
5.21	Fine-Tuned Threshold	40
5.22	Updated version of state extraction in Figure 5.8b	41
5.23	Example: Refined conversion of peak frequencies to states	41

5.24 Updated version of move detection in Figure 5.14b	42
5.25 Example: Refined move sequence extraction	43
6.1 Internal pieces of a standard speedcube (Gans 356)	46
6.2 Standard 555 Timer Circuit for Astable Operation [49]	47
6.3 Centerpiece State Transmitter Circuit	48
6.4 Output Frequencies of Common R and C Values	49
6.5 Output Frequencies of Common R and C Values	50
6.6 Breadboard Prototype	51
6.7 Effects of Various Sound Obstructions on Signal Intensity	51
6.8 Transmitter Circuit Inside a Custom Gans 356 Centercap	53
7.1 Perfect Detections by Cube	57
7.2 Influence of Standard Deviation Threshold on Accuracy	59
7.3 Perfect Detections by Minimum Threshold	60
7.4 Perfect Detections by Window Size	61

List of Tables

4.1 Sample frequencies for encoding a Rubik's Cube's state	24
--	----

List of Abbreviations

PCB	Printed Circuit Board
TPS	Turns Per Second
WCA	World Cube Association

For the benefit of the general speedcubing community...

Chapter 1

Introduction

1.1 The Advent of Smartcubes

Speedsolving, the sport of solving twisty puzzles like the Rubik's Cube as fast as possible, has seen a resurgence of popularity since the early 2000s. [1] Over the past two decades many advances in cube technology have produced ever higher performing puzzles.

Recently, the speedcubing community has seen the entrance of smart cubes, special versions of a Rubik's Cube built around hardware that can connect to a mobile device over Bluetooth. These smartcubes have sparked a wave of excitement with the vast opportunities they offer for automatic turn tracking, performance analysis, personalized improvement feedback, and networked competition.

1.2 Obstacles to Adoption

While a revolutionary innovation, smartcubes still face several obstacles to widespread adoption.

- *Cost:* Smartcubes generally cost two to eight times as much as a comparable, non-smart speedcube.¹
- *Performance:* Existing smartcubes turn slower than comparable non-smartcubes, making them less reflective of a speedcuber's true skill. [6]
- *Reliability:* Many smartcube owners report inability to connect the smartcube to a mobile device and missed/inaccurate turn tracking. [6]
- *Regulation:* Current competition rules ban the use of electronics during timed solves, thus banning the use of smart cubes [7]. There is no foreseeable change to this rule. [6]

As a result of these obstacles, many speedcubers refrain from purchasing a smartcube, despite expressing significant interest in the opportunities smartcubes offer.

¹For example, one popular budget speedcube, the Yuxin Little Magic, costs only \$5 [2], while the cheapest smartcube, the Giiker Cube, starts at \$42 [3]. On the higher end, a premium speedcube, like the Gans 356 XS, retails for \$53 [4] while a premium smartcube, like the GoCube, retails just under \$100. [5]

Furthermore, all current smartcubes have been specifically built for the primary purpose of providing move-tracking functionality. There is no existing way to automatically track the moves of a standard, "non-smart" speedcube.

1.3 Purpose of this Thesis

The primary goal of this thesis is to create a proof-of-concept for a smartcube design that can enable speedcubers to obtain the same metrics and data that are offered by existing commercial smartcubes from their main, non-smart speedcube.

In other words, this thesis will explore the following question in depth:

Is it possible to track the face turns of a standard, "non-smart" speedcube in a non-destructive, competition-legal way?

1.4 Thesis Overview

This exploration begins in Chapter 2 with an introduction of several foundational topics about the history and structure of a Rubik's Cube along with an overview of speedcubing competitions. Chapter 3 reviews relevant efforts in the field of smartcube construction and outline the detailed research questions for this thesis. Chapter 4 then defines and analyzes the various factors to consider when building a sound-based smartcube. Chapter 5 begins the proof-of-concept design by proposing a software receiver for a sound-based move tracking protocol. Chapter 6 completes the proof-of-concept by proposing a design for the complementary physical transmitter to embed into a speedcube. Chapter 7 then evaluates the performance of the receiver and transmitter against a variety of key benchmarks. Finally, Chapter 8 synthesizes answers to the specific research questions from Chapter 3 and discusses the outlook for future research in the domain of sound-based Rubik's Cube turn tracking.

Chapter 2

Background

This chapter provides important background information about cube that will facilitate understanding the topics presented in later chapters. Section 2.1 will describe the historical origin of the cube. Then, Section 2.2 will detail the structure of the cube and the notation used to transcribe move sequences. Finally, Section 2.3 will outline the culture and regulations surrounding competitive speedsolving.

2.1 A Brief History of the Rubik's Cube

In 1974, Erno Rubik, a Hungarian professor of architecture, sought to help his students visualize space in three dimensions. To that end, he created a special cube whose faces could independently rotate around all three physical axes [8]. When he added colored stickers to further aid in visualizing the movements, Mr. Rubik realized he had created a new puzzle. He patented his cube in 1975, [9] and since then over 450 million units have been sold [10], allowing an estimated 1 in 7 humans on earth to try their hand at solving it [11].

2.2 The Anatomy of a Rubik's Cube

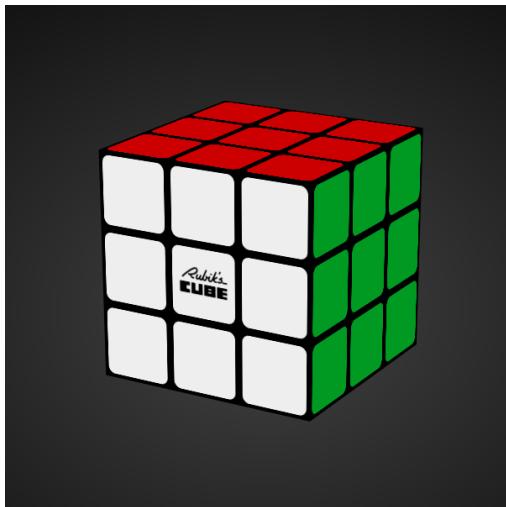
The Rubik's Cube, like a standard geometric cube, has six square faces positioned at right angles to each other. When solved, each of these faces has a single, unique color.

The Rubik's Cube is further subdivided into a 3x3x3 arrangement of smaller "cubies" such that each face consists of nine individual colored stickers/tiles (see Figure 2.1a). There are three different types of cubies - centers, edges, and corners - each which is distinguished by the number of unique colors it binds together into a single physical unit. The six center cubies have one color each (see Figure 2.1b), the twelve edges have two colors each (see Figure 2.1c), and the eight corners have three colors each (see Figure 2.1d).

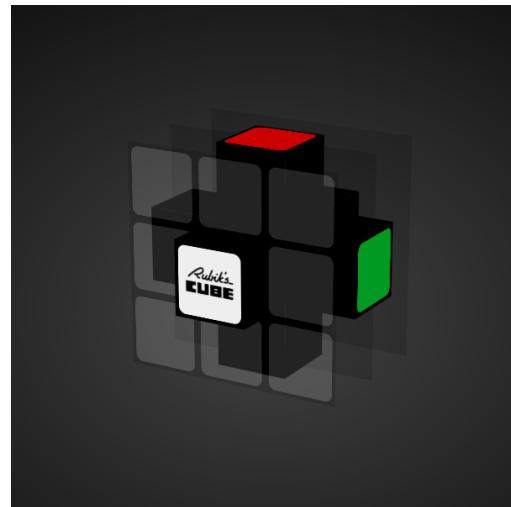
Each of the six center cubies are also attached to a common core which allows them to rotate freely, but fixes their position relative to each other. As such, the single color of each center cubie is also the color shared by the corresponding face when the entire cube is solved.

FIGURE 2.1: The Anatomy of the Rubik's Cube [12]

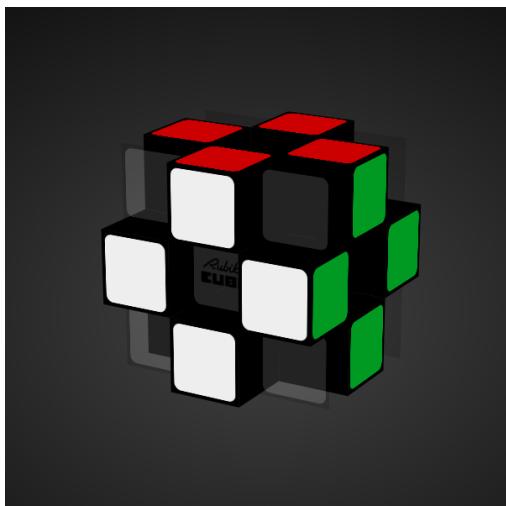
(A) Full Cube



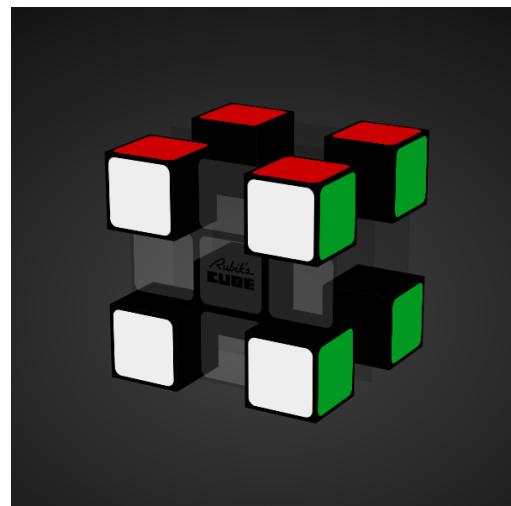
(B) Centerpieces



(C) Edges



(D) Corners



2.2.1 Move Sequence Notation

Shortly after the cube's invention, David Singmaster, a British mathematician, proposed a beginner-friendly, layer-by-layer approach to solving the cube. To facilitate communication of his solution, he outlined a notation for documenting the face turns required for each phase. This notation, now known as the Singmaster Notation, has become the standard notation used for documenting move sequences for the cube. [13]

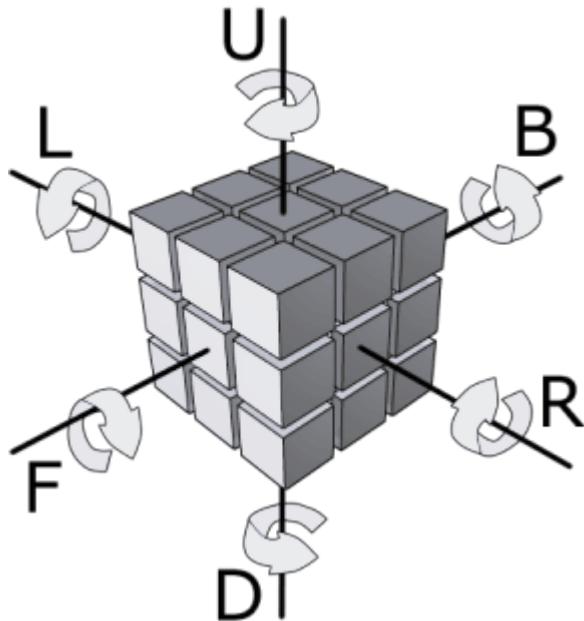
The notation consists of two components: (1) A label for each face of the cube and (2) a superscript to indicate a direction of rotation.

The labels for each face are as follows:

- *U* - "Up": The top face
- *D* - "Down": The bottom face
- *R* - "Right": The "Right" side face
- *L* - "Left": The "Left" side face
- *F* - "Front": The face directly in "Front" of the solver
- *B* - "Back": The face on the "Back" side of the cube facing away from the solver

Notice that the labels for each face are assigned based on its relative orientation, not the color of its centerpiece. For example, the top face is always labelled *U* whether the top facing centerpiece is green, yellow, or another color. This is demonstrated by the color-neutral image used to depict these face labels in Figure 2.2.

FIGURE 2.2: Notational names for each face of the Rubik's Cube [14]

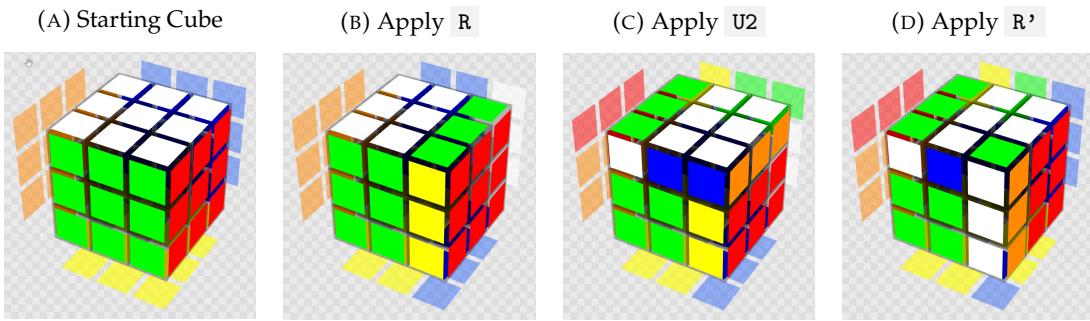


The superscripts are as follows:

- X^1 : 90° Clockwise rotation. Generally abbreviated as X
- X^{-1} : 90° Counter-clockwise rotation. Generally abbreviated as X'
- X^2 : 180° Rotation. Generally abbreviated as $X2$

For example, the move sequence $R U2 R'$ denotes turning the Right layer 90° clockwise, followed by a 180° rotation of the Up face, and finalized by a 90° counter-clockwise rotation of the Right face (see Figure 2.3)¹.

FIGURE 2.3: The move sequence $R U2 R'$



2.3 A Brief History of Speedsolving

As more people learned to solve the cube, there began to be increased interest in solving the cube as quickly as possible. In 1982, the first Rubik's Cube World Championship was held in Budapest with a winning solution time of 22.95 seconds by Minh Thai. However, due to falling interest, no further competitions were organized in the following two decades.

Interest in speedcubing picked back up in the late 1990s as the consumer internet facilitated sharing of strategies and techniques. In 2003, the second competition, the World Rubik's Games Championship 2003, was held in Toronto, Canada with 89 competitors and a winning average time of 20.00 seconds. Additional competitions were organized later that year and into 2004. [15]

2.3.1 The World Cube Association

To provide a set of standard regulations for the increasing number of speedcubing competitions, the World Cube Association (WCA) was founded in August 2004 by Tyson Mao and Ron van Bruchem. Today, the WCA governs all official speedcubing competitions by creating and enforcing a consistent set of regulations, providing tools for creating fair scrambles, and hosting the official platform for sharing competitions results. [15]

¹Images generated with the Twizzle Editor found at alpha.twizzle.net

2.3.2 The Rise of Smartcubes

In September 2018, the Xiaomi released the Giiker Cube, the first cube with built-in sensors to track its face turns during use. [3] For the first time, a solution could be analyzed without first requiring the tedious process of transcribing each individual move from a video recording! Additional "smartcubes" were released by a variety of manufacturers in the following years. [16] [17] [18]

However, despite their usefulness during personal practice, smartcubes cannot be used in official competitions. According to WCA regulation 2i, "While competing, competitors must not use electronics or audio equipment (e.g. cell phones, MP3 players, dictaphones, additional lighting) apart from the Stackmat timer or stopwatch." [7] This protects against competitors gaining unfair inspection time by previewing the scrambled cube from a connected device before it is brought to the competition table.

Chapter 3

State of the Art

3.1 Introduction

This chapter seeks to provide a comprehensive summary of the existing approaches to tracking the face turns of a Rubik's Cube. The most widely used solutions to date are found in Commercial Smartcubes (3.2), but significant research has also been carried out into Computer Vision based solutions (3.3.1). Other researchers have also explored the use of magnetic resonance and a muscle-tracking armband (3.4)

This chapter will also explore a selection of wireless communication techniques that at the time of writing have not been applied to the challenge of tracking the moves of a Rubik's Cube. Specifically, this chapter will review the potential usage of sound (3.4.1), Radio Frequency Identification (RFID) (3.4.2), and Off-Axis Magnetic Angle Sensors (3.4.3).

Finally, this chapter will close by detailing the specific research questions this thesis will seek to answer (3.5).

3.2 Commercial Smartcubes

Commercial Smartcubes are special Rubik's Cubes built around sensors that can detect face turns and transmit that information over Bluetooth. Some models can also measure and transmit data about the cube's orientation.

At the time of writing, there are four major smartcubes on the market: the Giiker Cube, the Go Cube, the Rubik's Connected (which is powered by GoCube technology) and the Gans 356i. This section will discuss the internal components of each of these cubes that provide this move-tracking functionality.

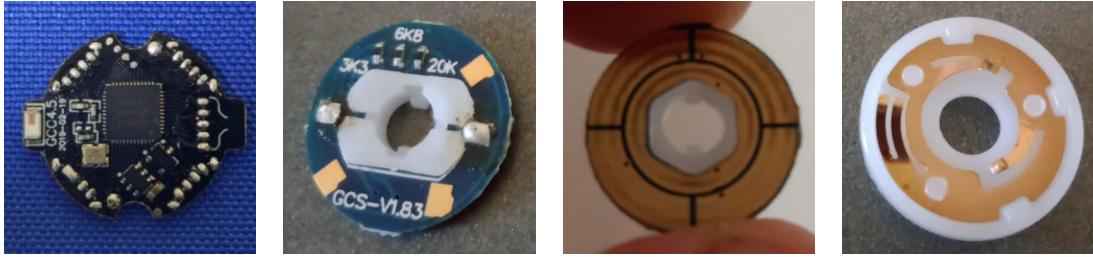
3.2.1 Giiker Cube

The Xiaomi Giiker Cube was released in September 2018 making it the first commercial smartcube on the market [3]. This "Supercube" as it was branded, used a relatively simple system for tracking the cube's movements. The core of the cube is built around a small circuit board with a microcontroller that measures the cube's movements and

a Bluetooth antenna that transmits those moves wirelessly (Figure 3.1a). The microcontroller detects each face turn by reading the voltage drop across a small circuit embedded within each center cap (Figure 3.1b). The center cap circuit controls its output voltage by using a copper brush to change between four separate electric paths, three different resistors and ground, as each face rotates (Figure 3.1d). [19]

FIGURE 3.1: The internal components of the Giiker Cube

(A) Internal Board [19] (B) Center Cap Board [20] (C) Center Cap Pads [20] (D) Center Cap Brush [20]



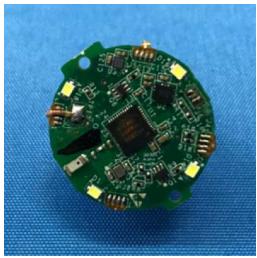
3.2.2 Go Cube

Announced on Kickstarter in June 2018, then shipped a year later [21], Patricula's GoCube was the first smartcube to include a gyroscope that would track a Rubik's Cube's orientation in addition to the face turns applied to it. [16] Like the Giiker Cube before it, the GoCube's core contains a small circuit board with the main electronics including a microcontroller, Bluetooth antenna, and the added gyroscope (Figure 3.2a). Though the teardown pictures from the Go Cube's FCC filing aren't particularly clear, it appears that the cube registers face turns similarly to the Giiker Cube: by producing a voltage drop via changing which one of the four resistors shown across the bottom of the center cap board in Figure 3.2b is in series with the circuit.

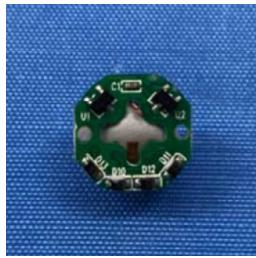
GoCube also serves as the underlying technology for the Rubik's Connected, the official smartcube from The Rubik's Company. [18]

FIGURE 3.2: The internal components of the Go Cube [22]

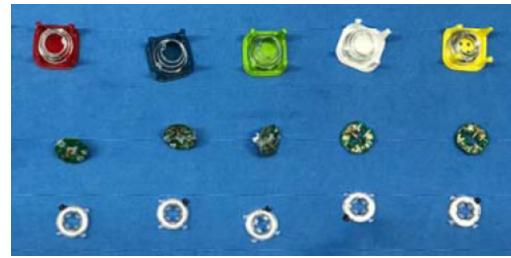
(A) Internal Board



(B) Center Cap Board



(C) Full Center Cap Teardown



3.2.3 Gans 356i

Released in July 2019, the Gans 356i was the first commercial smartcube produced by a traditional speedcube manufacturer. [17] While the Gans 356i also uses a microcontroller to process the face turns and Bluetooth to transmit the move data, it tracks moves

not through changing resistors in and out of a circuit, but via six plastic rods that connect the outer center caps to internal rotary encoders (Figure 3.3).

FIGURE 3.3: The Gans 356i Cube's internal components. [23]



3.3 Academia

In addition to the various commercial smartcubes, many academic research projects have involved some element of tracking the state/face turns of a Rubik's Cube.

This section summarizes the current state of academic research into using computer vision, magnetic resonance, and a muscle-tracking armband to track the state of a Rubik's Cube.

3.3.1 Computer Vision

Computer Vision refers to the "field of Artificial Intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos, and other visual inputs." [24] Since human manipulation of a Rubik's Cube is a physical, observable process, Computer Vision algorithms could be developed to extract face turn information from videos of Rubik's Cube solutions.

This section summarizes some of the relevant research in this area, including computer vision algorithms capable of extracting individual sticker colors from video, measuring the angle of rotation of a specific face, and detection of entire face turns and face turn sequences.

Sticker Color Classification

In 2015, Jay Hack, a graduate student studying Computer Science at Stanford developed a neural network capable of recognizing the colors of a Rubik's Cube face from video in various lighting conditions. His algorithm could classify frames within 7 milliseconds with 92% accuracy. [25]

Measuring a Face's Angle of Rotation

In 2019, OpenAI et al. published a viral video of a robot hand that had taught itself to solve a Rubik's Cube. While the final, most successful version of the robot hand's

software used a Giiker Cube to obtain the current rotational state of the cube, OpenAI et al. also researched the viability of tracking a Rubik's Cube's position using only computer vision. Their most successful vision-only algorithm measured only the rotation angle of the top-most face on the Rubik's Cube and assumed significant hardware requirements: a modified sticker set for the Rubik's Cube, a well-lit environment, three strategically positioned RGB Basler cameras, and a neural network trained on "a pool of optimizer nodes, each of which uses 8 NVIDIA V100 GPUs and 64 CPU cores". At peak performance, their vision-only algorithm's average error (the difference between the predicted face angle and the actual face angle) was 15.92° , nearly three times the 5.90° average error of the hardware-based face angle measurement. [26]

Classification of Single Moves and Entire Move Sequences

In 2020, Junshen Kevin Chen, Wanze Xie, and Zhouheng Sun, graduate Computer Science students at Stanford created the DeepCube dataset consisting of over 20,000 videos of Rubik's Cube face turns with consistent lighting and backgrounds. They also built a neural network to classify the videos with the face turn they contained. Their best performing model only made "one mistake every 15 moves" which corresponds to a 93.3% accuracy. [27]

3.3.2 Magnetic Resonance

In 2018, Maria Mannone et al. used the IM3D magnetic 3D motion tracking technology introduced by Huang et al. [28] to track the state of a Rubik's Cube across various movements for the purpose of generating a sequence of musical chords. This approach to turn tracking requires a special array of magnetic coils as shown in Figure 3.4a and the installation of "multiple small, light-weight, wireless markers (LC coils) with unique IDs" (a process that requires permanent modifications to the cube as evidenced by the damaged plastic in Figure 3.4b). Mannone et al. reported no issues with mistakes in this move tracking technology. [29]

FIGURE 3.4: The IM3D technology as used in the Cube Harmonic

(A) IM3D System Architecture [28]



(B) IM3D trackers in a Rubik's Cube [30]



3.3.3 Muscle-Tracking Armband

In 2017, Richard Polfreman and Benjamin Oliver researched ways to use the face turns of a Rubik's Cube as controls for a music synthesizer. They explored the use of a muscle-tracking armband (specifically the Myo Armband) to track the human solver's finger movements while manipulating the cube. However, since "the Myo moved a little when 'cubing'", they ultimately found greater success with a computer vision based tracking solution similar to those discussed in 3.3.1. [31]

3.4 Other Relevant Research

Finally, there are a number of research papers/commercial products that seek to transmit data in highly-constrained environments that are potentially relevant to the challenge of tracking the turns of a Rubik's Cube.

This section discusses some of these potential alternate move tracking mediums, specifically sound, RFID, and off-angle magnetic rotation sensors.

3.4.1 Sound

Sound is another communication medium that could be leveraged to track the moves of a Rubik's Cube. In 2015, Jonas Michel, a researcher at The University of Texas at Austin documented his exploration of the viability of creating an "acoustic modem" to transmit an arbitrary sequence of bits using sound. He observed that "as commercial off-the-shelf (COTS) smartphones become more powerful, it is worthwhile to revisit the use of sound as a medium for aerial digital device-to-device communications." [32]

Indeed, since many speedcubers practice by timing solves on a microphone-equipped smartphone or laptop¹, sound is a promising alternative communication medium to existing Bluetooth-based smartcubes.

3.4.2 Radio Frequency Identification (RFID)

Radio Frequency Identification (RFID) is a wireless technology often used in supply-chain systems [35] based on individual tags that can transmit a fixed set of information to a nearby reader. [36]. In 2018, Genovesi et al. proposed a rotary encoder based on RFID that produced angle measurements accurate to within 3° after a calibration that "must only be done once (when the sensor is put in place) if the distance between the reader and the tag does not change." [35]

While this approach may not be particularly useful for tracking the moves of human speedcubers since their movement of the cube would require constant re-calibration, it could be used in robotics based applications desiring to track the ongoing state of the cube.

¹One of the highest rated cubing timers on Android, Twisty Timer, has over 100,000 downloads on the Google Play Store [33]. By comparison, SpeedSolving.com, the central forum for speedcubing related discussion, has about 43,000 members [34]

3.4.3 Off-Axis Magnetic Angle Sensors

Another unique type of rotary encoder is an off-angle magnetic rotational sensor like the ones produced by NVE Corporation. [37] In the context of a Rubik's Cube, a properly sized diametric ring magnet could be fastened to the inner screw below the center cap of each face of the cube and the resulting optimal position for the magnetic sensor would stay within the walls of the center cap as shown by the the z_0 and R_0 values in Figure 3.5 which are smaller than the 8 mm distance between the center screw and the inner wall of the center cap in a standard-sized cube like the Gans 356.

FIGURE 3.5: Off-Angle Magnetic Sensor Calculations [38]



3.5 Research Questions

While there are many effective solutions for tracking the moves of a Rubik's Cube if the cube is built for that purpose, there does not yet exist a comparably effective technique for tracking the moves of a standard, "non-smart" Rubik's Cube. As shown above, significant effort has been spent researching and developing solutions that can leverage the Bluetooth transmitters and camera sensors of consumer grade smartphones and laptops with varying degrees of success. However, little to no research has been carried out exploring the viability of using the microphones readily available on the same devices for this purpose.

Thus, this thesis will seek to answer the overarching question from Chapter 1 "*Is it possible to track the face turns of a standard, "non-smart" speedcube in a non-destructive, competition-legal way?*" by detailing a proof-of-concept for a sound-based smartcube design created in response to the following questions:

1. **Feasibility on Consumer Hardware:** *What are the constraints for a sound-based smartcube design compatible with consumer-grade microphones like those found in common smartphones and laptops?*
2. **Move Tracking Accuracy:** *How could such a sound-based smartcube design track the face turns of a Rubik's Cube with high accuracy?*
3. **Compatibility with Standard Speedcubes:** *How could such a sound-based smartcube design be deployed within a standard, "non-smart" speedcube without requiring permanent modifications to the original cube?*
4. **Move Tracking Granularity:** *How could such a sound-based smartcube design record the time spent executing each individual face turn of a Rubik's Cube?*
5. **Competition Legality:** *How could such a sound-based smartcube design comply with competition regulations prohibiting the use of electronics while performing a competitive solve?*

Chapter 4

The Protocol

4.1 Introduction

Designing a sound-based protocol for tracking the moves of a Rubik's Cube requires great care. Lots of things produce sound that could interfere with the protocol: people talking, machines operating, nature stirring, and so forth. Furthermore, the structure of the Rubik's Cube itself imposes stringent physical constraints on the size of any components used to produce the sounds used in the protocol.

This chapter first seeks to clearly detail the specific constraints that must be considered when designing a sound-based protocol for tracking the moves of a Rubik's Cube in Section 4.2. From there, a specific protocol that meets these constraints will be proposed in Section 4.3. This proposed protocol will then be contrasted with an alternative option in Section 4.4. Finally, an overview of the plan to implement the proposed protocol in a proof-of-concept will be given in Section 4.5

4.2 Requirements

This section will detail the constraints required for designing a protocol for tracking the moves of a Rubik's Cube.

These constraints include a sufficiently strong signal-to-noise ratio (Section 4.2.1), sufficient distinctiveness between tones (Section 4.2.2), the frequency response range of consumer hardware (Section 4.2.3), and the human auditory range (Section 4.2.4).

4.2.1 Signal-to-Noise Ratio

Sound is an easily accessible, and therefore noisy, medium of communication. As a result, any data transmission protocol based on sound must be resilient to the presence of additional noise unrelated to the signal being transmitted. For a sound-based protocol to be effective, its tones must be easily distinguishable from background noise.

For the purpose of this thesis, "background noise" will be considered the ambient noise present in a quiet room when a speedcuber is actively solving a Rubik's Cube. Measurement of the background noise was carried out by using the Android app Spectroid [39] to create live visualizations of the background noise in a household bedroom (a typical place for a speedcuber to practice solving the cube) while solving three types of

speedcubes selected based on their "noisiness". A fourth visualization of just the ambient noise in the bedroom (i.e. while no cubes were being solved) was also recorded as a control. These visualizations are shown in Figure 4.1.

How to Read a Spectrogram

These visualizations are formally known as an audio "spectrogram", and show both the specific frequencies present in each audio sample and their intensity.

The graph on top shows the current and maximum strengths of each frequency present in the background audio with the yellow and red lines respectively. The x-axis measures the specific frequencies in Hz, while the y-axis shows the intensity (aka loudness) of those frequencies in dB.

The lower graph is the actual spectrogram of the background audio. It shares the same x-axis as the top graph, but its y-axis instead measures time with time = 0 at the bottom of the graph and time = now at the top. Here, the brighter colors represent a more intense (aka louder) frequency with the exact key shown on the left side of the graph. As such, this lower graph can be considered a "bird's eye view" of the top graph over time.

Key Observations from the Spectrogram

As shown in Figure 4.1a, the predominant background noises in a quiet bedroom are the tones between 80Hz - 500Hz which reach a max strength of -54dB.

In contrast, Figures 4.1b and 4.1c show that solving a speedcube creates additional noise in the frequency ranges from 1000Hz - 20000Hz, with Figure 4.1d showing particular strength from the noisy QiYi cube with the frequencies in the 1000Hz - 10000Hz range reaching up to -30dB.

Thus, a sound-based move tracking protocol must account for the following items in order to achieve an adequate signal-to-noise ratio:

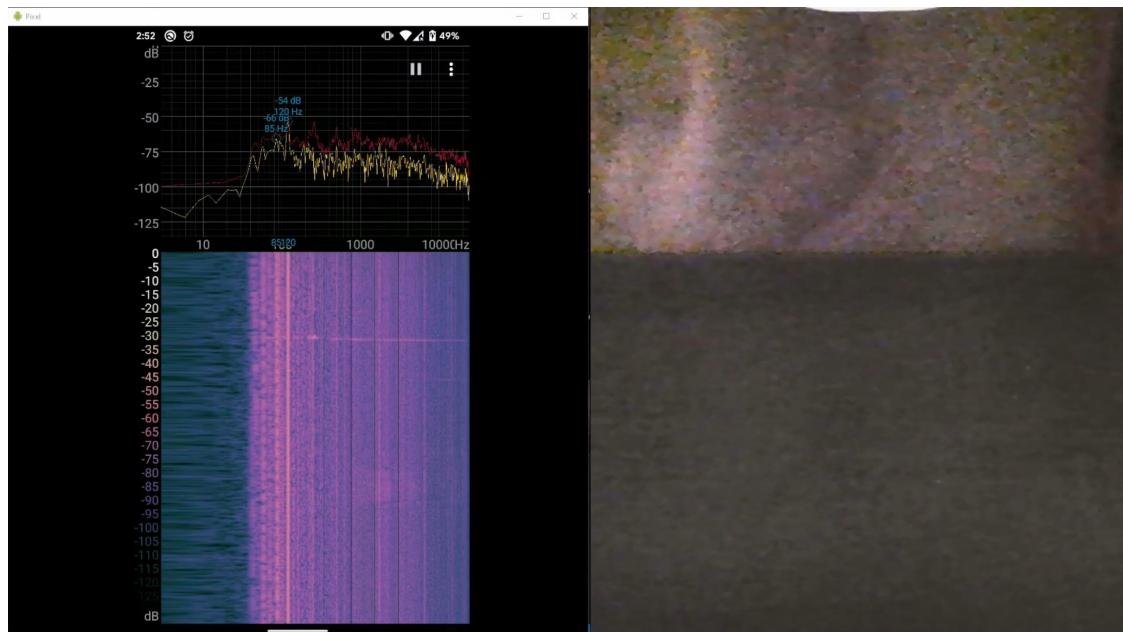
- Tones between 500Hz and 1000Hz are the easiest to detect while solving a speedcube since they have the least competition from other background sounds while solving a speedcube.
- Tones between 1000Hz and 20000Hz must be significantly louder than -30dB or they risk being indistinguishable from the background noise created by a noisy speedcube like the QiYi Qimeng.

4.2.2 Tone Distinctiveness

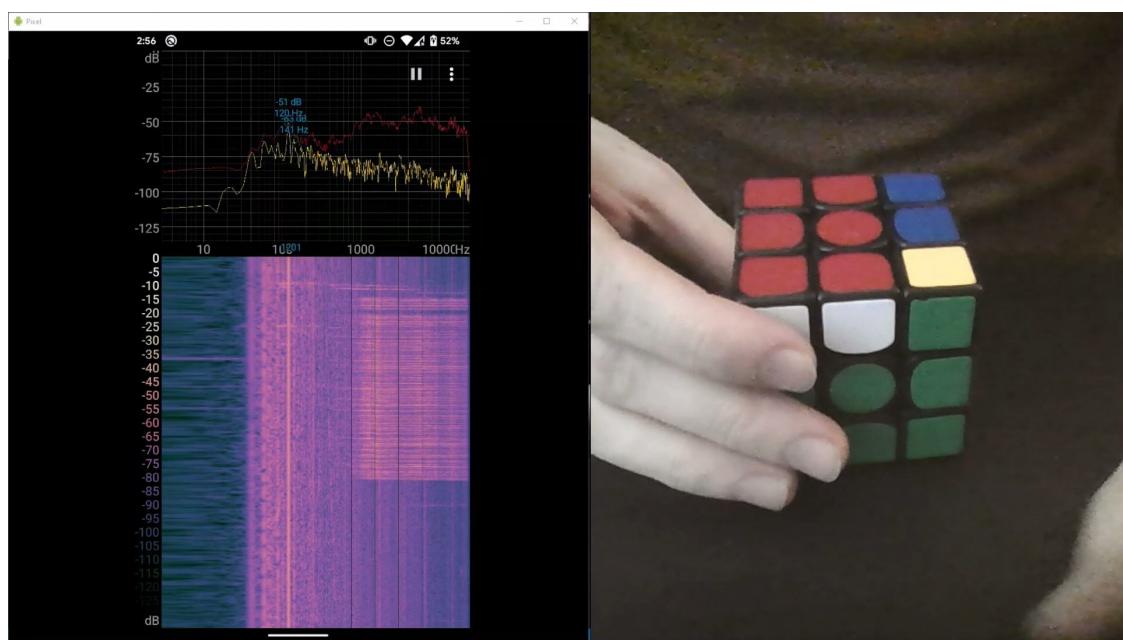
If more than one tone is required for the protocol, each tone must be unique enough to be easily distinguished from each other tone. Since a standard smartphone or laptop microphone will be used on the listening end of this protocol, the definition of "easily distinguishable" must be based on an assessment of how clearly smartphone and laptop-grade microphones can distinguish similar frequencies.

FIGURE 4.1: Background Noise of a Quiet Room while Speedsolving

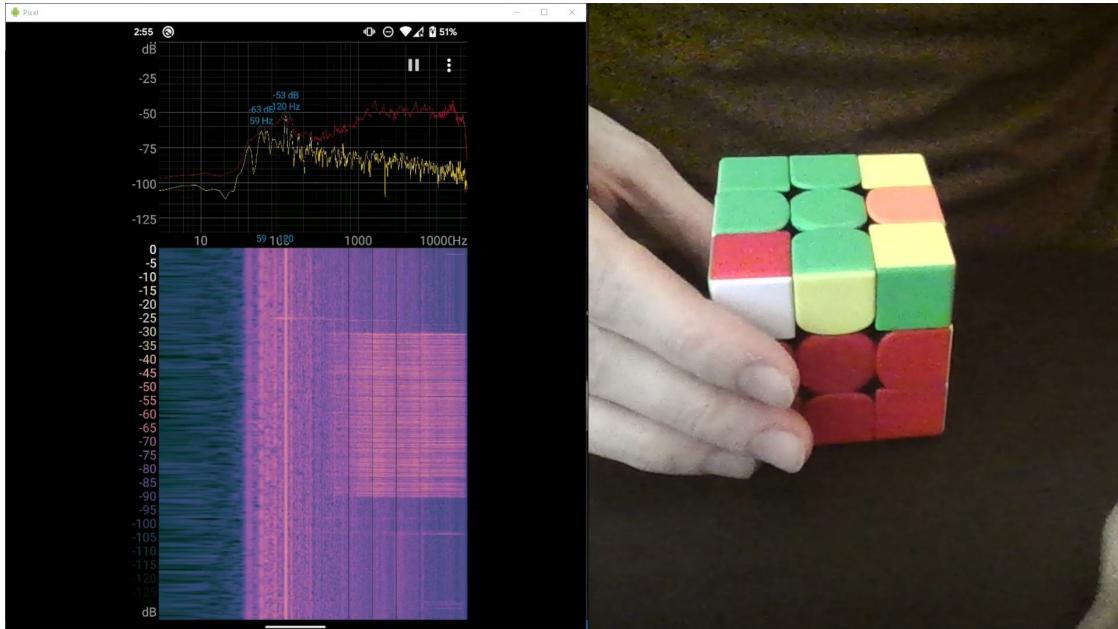
(A) No Speedsolving (Quiet Room)



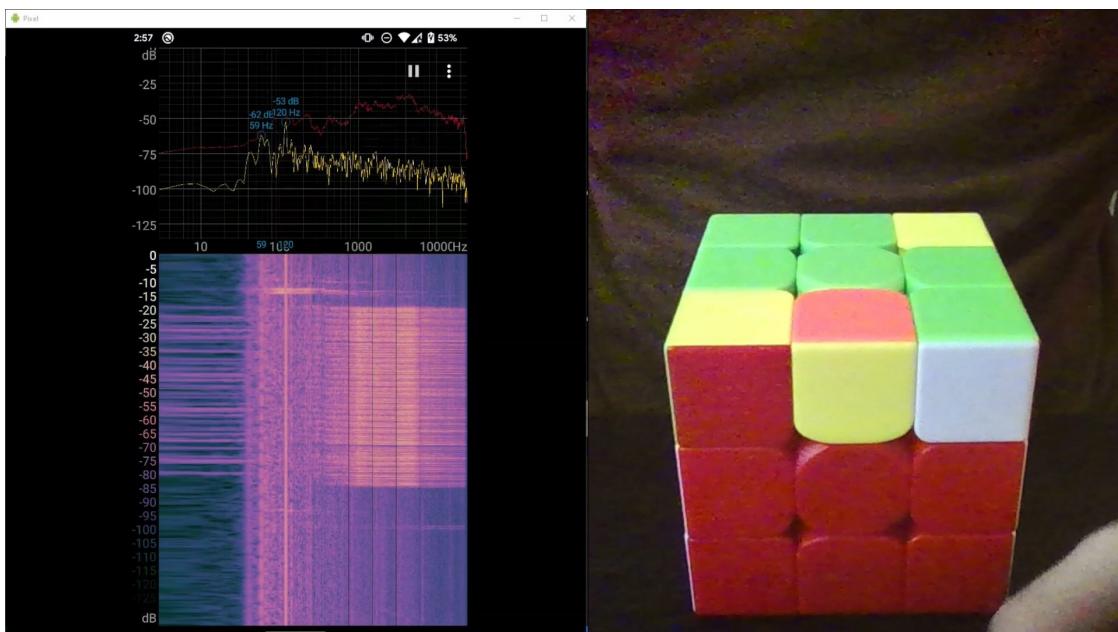
(B) Quiet Solving (Gans 356)



(C) Normal Solving (Gans XS)



(D) Loud Solving (QiYi Qimeng)



To measure the sensitivity of a standard smartphone microphone (in this case a Google Pixel 1), a recording was taken of two distinct tones that started 500 Hz apart and stepped closer together in 10 Hz increments every 0.5 seconds until their frequencies were identical.

As shown in Figure 4.2, the two tones are clearly distinguishable from 500Hz apart to 150Hz apart. However, the distinction began waning once the tones came within 100Hz of each other, and by 80Hz they became entirely indistinguishable.

FIGURE 4.2: Distinctiveness of Two Increasingly Similar Tones



Thus, a sound-based move tracking protocol must account for the following items in order to achieve adequate tone distinctiveness:

- Tones should be separated by at least 100Hz in order to be clearly distinguished from other tones in the protocol.

4.2.3 Frequency Response Range of Consumer Hardware

Since a standard smartphone or laptop microphone will be used on the listening end of this protocol, any tones used in the protocol must be within the range of tones that a smartphone or laptop-grade microphones can pick up, otherwise a more sophisticated, external microphone will need to be attached to the device. This range is formally known as the "frequency response range" of the microphone.

According to a Stanford research paper [40] that analyzed over 10,000 mobile devices, the typical smartphone microphone has a frequency response range of 20Hz to 20kHz as shown in Figure 4.3.

FIGURE 4.3: "A typical frequency response curve for a microphone." [40]



Thus, a sound-based move tracking protocol must account for the following restrictions on which tones can be used for the protocol:

- Tones used in the protocol must fall on the range of 20Hz-20kHz in order to be detectable by a typical smartphone or laptop microphone.

4.2.4 Human Auditory Range

An audible protocol could be distracting to a speedcuber. The human ear can detect audible frequencies from 20Hz to 20kHz, though this usually degrades with age with many people unable to notice sounds above 16kHz. [41]. However, not all tones are pleasant to listen to, particularly higher frequency tones. Tones within the frequency range of a piano (27.5Hz - 4kHz) are generally considered acceptable, while tones above the piano's upper range are often irritating. [42]

Thus, a sound-based move tracking protocol should be considerate of the human ear's sensitivity to various frequencies:

- The most acceptable tones for human speedsolvers are in the musical range of up to 4kHz or above the standard audible range of 16kHz

4.3 Specification

Given the above constraints, this section will detail a sound-based protocol for tracking the moves of a Rubik's Cube by continuously transmitting the current state of the cube. In this protocol, changes to the cube's state cause a change in the transmitted tones which can be recorded and analyzed to determine the face turn applied.

4.3.1 Representing the Cube's Current State

As mentioned in Section 2.2, a Rubik's Cube has six centerpieces that are fixed relative to each other, but can each rotate freely through four possible rotational positions as shown in Figure 4.4.

FIGURE 4.4: The four possible rotations of a face of a Rubik's Cube [43]

(A) Full Alignment (B) 90° Misalignment (C) 180° Misalignment (D) 270° Misalignment



The state of a centerpiece is defined as its current rotational position. Implicit in this definition is the fact that a centerpiece is guaranteed to occupy one and only one of its four possible states at any given time.

Each of the six centerpieces has an independent set of four possible states, yielding a total of 24 different centerpiece states for the Rubik's Cube, of which there will always be exactly six active at any given time.

It's important to note that a knowledge of the current state of each centerpiece does not imply knowledge of the exact state of each of edge and corner cubie. For example, after applying the algorithm R U R' U' all centerpieces have the same state they occupied prior to the algorithm's execution, while most of the edges and corners in the R and U layers will have been moved or rotated. However, in the reverse case, knowledge of the applied move sequence is sufficient information to determine the exact state of all cubies. Section 4.3.2 will explain how extract full move sequences using only the state information of the centers.

4.3.2 Tracking Face Turns

The only way to change a centerpiece's state is by applying a face turn. As such, when a centerpiece is rotated to a new position, its state changes to that of the new position. From there, a simple comparison of the new state to the previous state reveals the exact face turn applied to the cube.

Thus, in order to determine the sequences of moves applied to the Rubik's Cube, one simply needs to track how the state of its centerpieces changes over time.

4.3.3 Conveying State Through Sound

Conveying the current state of the cube's centerpieces through sound can be done by simply associating each of the 24 possible centerpiece states with a specific tone as shown in Table 4.1 and broadcasting a signal composed of the six tones corresponding to the active states of the cube's centerpieces.

Whenever a face turn is applied, the tone associated with the rotated centerpiece would change to the tone representing the piece's new state. From there, a microphone equipped device can record the broadcast frequencies, measure the changes in the frequencies

over time, convert the frequency changes to state changes, and finally extract any move sequence applied to the cube.

Since there will only ever be a single tone broadcast for a specific face, the four tones used to represent each of its possible states can be chosen closer to the minimum separation specified in Section 4.2.2. However, the gap between tones for different faces should be larger to help prevent incorrectly classifying a tone as conveying a state of a different face. In the sample frequencies shown in Table 4.1, the frequency gap between faces is 200Hz, a value which performed well in testing.

TABLE 4.1: Sample frequencies for encoding a Rubik's Cube's state

Alignment	White	Yellow	Red	Orange	Green	Blue
Full	800 Hz	1300 Hz	1800 Hz	2300 Hz	2800 Hz	3300 Hz
90°	900 Hz	1400 Hz	1900 Hz	2400 Hz	2900 Hz	3400 Hz
180°	1000 Hz	1500 Hz	2000 Hz	2500 Hz	3000 Hz	3500 Hz
270°	1100 Hz	1600 Hz	2100 Hz	2600 Hz	3100 Hz	3600 Hz

4.4 Alternative Protocol

Instead of transmitting the current state of the cube's centerpieces, an alternative protocol design could seek to directly transmit the face turns applied to the cube.

This perspective focuses on the fact that all move sequences can be broken down into a series of 90° face turns. Since the cube consists of 6 faces, and each face can be turned either clockwise or counterclockwise, one could design a two-tone protocol using only 8 discrete audio frequencies to build the smartcube. The first tone would come from one of six predefined audio bands, one for each face of the cube. The second tone would come from one of two separately predefined audio bands, one for each possible direction of rotation. From this, an audio processing model could be designed to process a sequence of these two-tone pairs and reconstruct the sequence of face rotations by recording the rotated face followed by its direction of rotation.

However, while this model minimizes the number of discrete frequencies required to communicate changes in the cube's state, it carries many challenges. Consider the example of a speedcuber averaging 5-8 turns per second (TPS) with bursts up to 20 TPS (common for a speedcuber that averages under 15 seconds per solution). The burst TPS would require the successful transmission of 40 sequential tones within a single second - only 25ms per tone, all in the midst of additional noise from the cube's pieces hitting each other harder at the higher turn speed. Adding to the difficulty, since each tone is only ever transmitted once, the audio detection model must achieve 100% tone recognition to be able to accurately reconstruct the originating move sequence. As a result, this model fails to support any sort of error correction that would make it resistant to the common challenges to data transmission through sound.

4.5 Summary

In summary, a sound based protocol for tracking the moves of a Rubik's Cube must be considerate of the following constraints:

1. Tones should be transmitted with a strength greater than -30dB to achieve a high signal-to-noise ratio. (Section 4.2.1)
2. Tones should be separated by at least 100Hz in order to be clearly distinguished from other tones in the protocol. (Section 4.2.2)
3. Tones must fall on the range of 20Hz-20kHz in order to be detectable by a typical smartphone or laptop microphone. (Section 4.2.3)
4. Tones should fall on the range 0Hz-4kHz or 16kHz-20kHz to minimize annoyance to the human speedsolver. (Section 4.2.4).

The protocol that will be used in subsequent chapters consists of 24 unique tones, one for each possible rotational state of each centerpiece. A transmitter embedded in the Rubik's Cube will continuously broadcast the six tones corresponding to the current rotational state of each centerpiece adjust those tones as face turns are applied (Details in Chapter 6). A software receiver will then listen to the transmitted audio and measure the changes in the transmitted frequencies over time to reconstruct the originally applied sequence of face turns (Details in Chapter 5).

Chapter 5

The Receiver

5.1 Introduction

The receiver for the sound-based move tracking protocol is in charge of decoding a sequence of face turns from the transmitted audio.

To do this, the receiver must go through the following steps:

1. Record the transmitted audio.
2. Measure the audible frequencies at each time step.
3. Convert each time step's audible frequencies to the cube's state at that moment.
4. Decode the applied face turns from the sequence of cube states.

This chapter will describe the development of a software algorithm in Python that can serve as a receiver for this sound-based move tracking protocol. This development began with the creation of synthetic audio recordings representing the tones that would be emitted by a Rubik's Cube equipped with an ideal transmitter (Section 5.2) followed by the implementation of an algorithm capable of decoding that ideal synthetic audio (Section 5.3). The algorithm was then made more robust by adding realistic noise to the synthetic audio to better simulate a real speedcubing environment (Section 5.4) followed by enhancing the previously designed algorithm to continue to decode the applied move sequence in the midst of the added noise (Section: 5.5).¹

5.2 Creating Synthetic Audio

The first step of designing this receiver is to synthesize an audio signal representative of the output of the ideal transmitter.

This audio synthesis starts with encoding the frequency corresponding to each centerpiece state (Section 5.2.1), then involves creating a virtual Rubik's Cube (Section 5.2.2), and finally culminates in generating the audio signal from the virtual Rubik's Cube's state (Section 5.2.3).

¹The contents of this chapter have been specially written so that the reader can copy them into a Jupyter Notebook running a Python 3.9 kernel and see the same results for himself/herself.

5.2.1 Representing the Audio Protocol

For the synthetic audio generator to produce a realistic signal, it needs to know which frequencies to transmit for each centerpiece state. This is easily accomplished by creating a dictionary to map each possible state of each face to its corresponding frequency.

For this design, the frequency assignments from Table 4.1 are converted to the dictionary shown in Figure 5.1a with two small changes. First, by assuming no cube rotations, the face color (e.g. "White") can be converted to a layer name (e.g. "U") which simplifies the code for both encoding and decoding a move sequence from audio. Second, the 0°, 90°, 180°, and 270° alignments from Figure 4.4 are divided by 90 to be represented by the simple sequence of consecutive integers 0, 1, 2, and 3 that are easier to manipulate in code.

With this dictionary in place, the frequency to transmit for any particular centerpiece's current rotation can be determined by looking up the centerpiece and its rotation in the dictionary as shown in Figure 5.1b.

FIGURE 5.1: Centerpiece State to Frequency Mapping

(A) Frequency mapping dictionary

```

1 FREQUENCY_MAPPINGS = {
2     "U": {
3         0: 800, 1: 900, 2: 1000, 3: 1100
4     },
5     "D": {
6         0: 1300, 1: 1400, 2: 1500, 3: 1600
7     },
8     "R": {
9         0: 1800, 1: 1900, 2: 2000, 3: 2100
10    },
11    "L": {
12        0: 2300, 1: 2400, 2: 2500, 3: 2600
13    },
14    "F": {
15        0: 2800, 1: 2900, 2: 3000, 3: 3100
16    },
17    "B": {
18        0: 3300, 1: 3400, 2: 3500, 3: 3600
19    }
20 }
```

(B) Frequency lookup function

```

21 def frequency_of(centerpiece: str, rotation: int) -> float:
22     return FREQUENCY_MAPPINGS[centerpiece][rotation]
```

5.2.2 Representing the Rubik's Cube

While there are many implementations of a digital Rubik's Cube that can track every cubie and render an interactive 3D cube, all that's needed for the synthetic audio generator is a representation of a Rubik's Cube on which individual face turns can be virtually applied and the

resulting centerpiece state can be read out. To do this, a `RubiksCube` object is created that encapsulates a dictionary containing the same keys for each face as the `FREQUENCY_MAPPINGS` dictionary in Figure 5.1a above, each associated with a single integer representing the current rotational state of that face. A method called `apply_move` is also defined with a parameter for a valid move like `U` or `U'` that will update the `RubiksCube`'s state to reflect the rotation.

FIGURE 5.2: A simple abstraction of a Rubik's Cube

```

1 class RubiksCube:
2
3     CLOCKWISE = 1           # aka a 90 degree turn
4     COUNTERCLOCKWISE = 3    # aka a 270 degree turn which achieves the
5                             # same resulting state as a -90 degree turn
6
7     def __init__(self):
8         self.state = { "U": 0, "D": 0, "R": 0, "L": 0, "F": 0, "B": 0 }
9
10    def apply_move(self, move: str):
11        # Extract the face and direction from the 'move' string.
12        face = move[0]
13        if len(move) == 1:                                # e.g. U
14            direction = RubiksCube.CLOCKWISE
15        else:                                         # e.g. U'
16            direction = RubiksCube.COUNTERCLOCKWISE
17
18        # Update the state to apply the move.
19        # NOTE: The 'mod 4' exists to wrap back around to 0
20        # after a full 360 degrees of rotation.
21        self.state[face] = (self.state[face] + direction) % 4

```

5.2.3 Creating Synthetic Audio for an Arbitrary Algorithm

Now the synthetic audio can be generated for any valid algorithm. This is done using the `tones` library created by Erik Nyquist. [44]

First, a separate audio track is created to simulate the output of each speaker embedded into each centerpiece on the cube (Figure 5.3a).

Second, a new function is written to add the tones corresponding to the virtual cube's current state to the audio mixer using the `frequency_of` function from Figure 5.1b (Figure 5.3b). This function includes a `tps` parameter that controls the duration of the added tones so that they change to the next tone at the same speed as the turns on a speedcube being solved at that number of turns per second.

Then, those two functions are tied together by iterating through each move in a Rubik's Cube `alg` orithm and saving the audio of the resulting state changes into a .wav file at a given file path (Figure 5.3c). Additionally, the `tps` parameter is propagated to this function to enable generating audio for an algorithm at different turn speeds.

FIGURE 5.3: Generating audio for any Rubik's Cube algorithm

(A) Function to create an audio mixer for a virtual Rubik's Cube

```

1 from tones.mixer import Mixer # https://pypi.org/project/tones/
2 from tones import SINE_WAVE
3
4 def _create_mixer(rubiks_cube: RubiksCube) -> Mixer:
5     mixer: Mixer = Mixer(sample_rate=44100, amplitude=1)
6     # Add a separate audio track for each centerpiece.
7     for face, _ in rubiks_cube.state.items():
8         mixer.create_track(face, SINE_WAVE, attack=0, decay=0)
9     return mixer

```

(B) Function to update an audio mixer with a virtual Rubik's Cube's state

```

10 def _render_cube_state(mixer: Mixer, rubiks_cube: RubiksCube, tps: float):
11     for face, rotation in rubiks_cube.state.items():
12         mixer.add_tone(face, frequency_of(face, rotation), duration=1 / tps)

```

(C) Function to create a .wav file of any Rubik's Cube algorithm

```

13 def render_audible_alg(alg: str, wav_path: str=None, tps: float=4):
14     # Create the virtual Rubik's Cube.
15     rubiks_cube: RubiksCube = RubiksCube()
16     # Create the audio mixer used to create the synthesized audio.
17     mixer = _create_mixer(rubiks_cube)
18     # Add the initial cube state to the mixer.
19     _render_cube_state(mixer, rubiks_cube, tps)
20     # Iterate over the alg, updating the mixer after each move.
21     moves = alg.split(" ")
22     for move in moves:
23         rubiks_cube.apply_move(move)
24         _render_cube_state(mixer, rubiks_cube, tps)
25     # Save the final audio to a .wav file.
26     mixer.write_wav(wav_path if wav_path else f"{alg}.wav")

```

With these functions in place, synthetic audio can be easily created for any valid Rubik's Cube algorithm. For example, generating synthetic audio that sweeps through every possible centerpiece state can be done with the two lines of code in Figure 5.4.

FIGURE 5.4: Example Audio Generation for a Rubik's Cube Algorithm

```

1 demo_alg = "U_U_U_U_D_D_D_R_R_R_L_L_L_F_F_F_B_B_B"
2 render_audible_alg(demo_alg, "demo_all_states.wav")

```

5.3 Decoding the Synthetic Audio

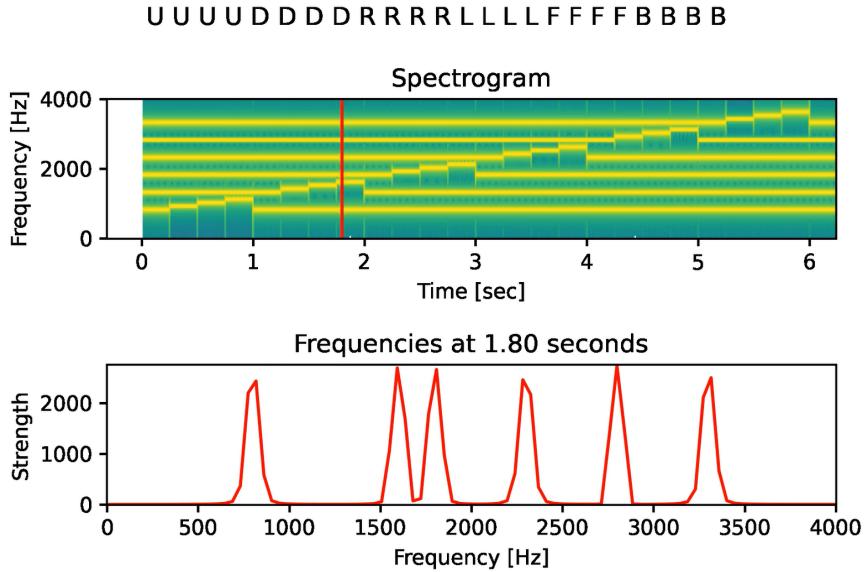
With synthetic audio now available for any valid Rubik's Cube algorithm, the next step is to create an initial software algorithm that can decode that audio back into the original move sequence.

Accomplishing this will require computing the synthetic audio's spectrogram (Section 5.3.1), followed by extracting the dominant frequencies present at each time step (Section 5.3.2), and converting those dominant frequencies into the corresponding Rubik's Cube centerpiece states (Section 5.3.3), all before finally recovering the originally applied move sequence (Section 5.3.4).

5.3.1 Computing the Spectrogram

The first step in decoding the synthetic audio is determining its component frequencies at any specific moment in time. These component frequencies can be easily visualized using a spectrogram, like the one in Figure 5.5 of the synthetic audio created in Section 5.2.3.²

FIGURE 5.5: Spectrogram of the synthetic audio created in Section 5.2.3



This spectrogram has the same pieces as the ones described in Section 4.2.1 with a few minor changes. First the order of the graphs has been reversed, so the top graph shows the actual spectrogram and the bottom graph shows the strength of the component frequencies at the specific point in time indicated by the vertical red line on the top graph. Additionally, the spectrogram has been rotated 90° so time is now on the x-axis and the component frequencies are shown on the y-axis. Finally, the color scheme is slightly different, with the presence of a strong component frequency indicated by a bright yellow color instead of bright purple or pink.

²The source code for Figure 5.5 is available in Appendix A.1.

The Brief Overview of the Math behind the Spectrogram

However, while Matplotlib's `specgram` plot [45] made it easy to create Figure 5.5, it didn't return the underlying data of the spectrogram required to decode the exact values and strengths of the component frequencies at each point in time. To compute that data directly, it is helpful to understand the basics of the math behind the spectrogram. Dr. Steve Brunton from the University of Washington created an excellent video series on this topic [46], of which the following few paragraphs are a short summary.

The foundational mathematical concept behind the spectrogram is the Fourier Transform, which is a technique for approximating any continuous function using only sine and cosine functions. In the case of an analog audio signal -which is inherently a composite of many sine functions (one for each component frequency)- applying a Fourier Transform would return data for a graph of all the frequencies present at any point in the entire signal and their overall strength throughout it.

But digital audio signals like .wav files are a series of discrete values, not continuous functions, which means the Fourier Transform cannot directly operate on them. Fortunately, there exists a variant of the Fourier Transform called the Discrete Fourier Transform which can operate on a list of discrete data points by assuming they sample a continuous function and then computing the Fourier Transform of that function.

However, when applied to a digital audio signal, the Discrete Fourier Transform still only computes which frequencies were present at *any* time during the duration of the signal, but not *when* they occurred. That data comes from another layer of computation called the Gabor Transform (i.e. the Short-Time Fourier Transform) which computes a Discrete Fourier Transform over a sliding window of the input samples to approximate the changes in the strength of each component frequency over time.

This knowledge, combined with the `numpy` [47] and `scipy` [48] libraries which can calculate these transforms, enables the creation of a function (Figure 5.6) to compute the spectrogram for the audio stored in a given .wav file. The return values `freq` and `time` are, respectively, lists of the .wav file's component frequencies and time steps, while `spectrogram` is a 2D array of the strengths of each component frequency at each time step. They are related by common indices, such that the strength of the frequency `freq[f_idx]` at the time `time[t_idx]` is found in `spectrogram[t_idx][f_idx]`.

FIGURE 5.6: Function to compute the spectrogram of a .wav file

```

1 import numpy as np
2 from scipy import signal
3 from scipy.io import wavfile
4
5 def compute_spectrogram(wav_path: str):
6     SAMPLES_PER_WINDOW = 1024 # Balances frequency/time accuracy
7     sample_rate, audio_samples = wavfile.read(wav_path)
8     freq, time, Zxx = signal.stft(audio_samples, fs=sample_rate,
9         nperseg=SAMPLES_PER_WINDOW, noverlap=(SAMPLES_PER_WINDOW // 4) * 3)
10    spectrogram = np.abs(Zxx).transpose()
11    return freq, time, spectrogram

```

5.3.2 Extracting the Dominant Component Frequencies

Looking back at the component frequency graph in Figure 5.5 it is clear that it has six distinct peaks: these are the transmitted frequencies representing the current state of the virtual Rubik's Cube's six centerpieces at that specific instant of time.

The exact frequencies of these peaks can be extracted by filtering out all frequencies whose strength is not above a specific threshold. In this case, a simple threshold of 85% of the maximum strength of any component frequency (see the green line in Figure 5.7³) isolates the six dominant component frequencies.

FIGURE 5.7: Spectrogram from Figure 5.5 with a threshold at 85% of the maximum strength of the component frequencies.



Using the actual spectrogram data, the exact values of these peaks can be computed. First, the threshold computation is broken into its own function (Figure 5.8a). From there, a second function (Figure 5.8b) can iterate through all the component frequencies of the computed spectrogram data at a specific time index to compile and return a list of the frequencies whose strength exceeds the threshold at that time step.

FIGURE 5.8: Extracting dominant frequencies from one time step of audio

(A) Function to calculate a simple threshold to isolate the most dominant frequencies

```
1 def compute_threshold(values: list):
2     return max(values) * 0.85
```

(B) Function to extract all frequencies stronger than a threshold at a given time step

```
3 def extract_important_freqs(freq, time, spectrogram, t_idx):
4     # Create a list to store the important frequencies
5     important_freqs = []
6     # Compute the threshold for this time step
7     threshold = compute_threshold(spectrogram[t_idx])
8     # Check the strength of each frequency at this time step
9     for f_idx in range(len(freq)):
10         # If the frequency's strength is above the threshold ...
11         if spectrogram[t_idx][f_idx] > threshold:
12             # ... save the frequency and its strength for later processing
13             important_freqs.append(dict(
14                 hz=freq[f_idx],
15                 power=spectrogram[t_idx][f_idx]
16             ))
17     return important_freqs
```

³The source code for Figure 5.7 is available in Appendix A.2.

With these functions, finding the actual frequencies of the peaks in Figure 5.7 can be done in just a few lines of code as shown in Figure 5.9.

FIGURE 5.9: Example peak frequency decoding at a specific time step

```

1 freq, time, spectrogram = compute_spectrogram(demo_wav_path)
2 important_freqs = extract_important_freqs(
3     freq, time, spectrogram, t_idx=310) # 1.80 seconds
4 print([f"{x['hz']:.0f}Hz" for x in important_freqs])

>> ["818Hz", "1593Hz", "1809Hz", "2283Hz", "2799Hz", "3316Hz"]

```

5.3.3 Translating Component Frequencies to Centerpiece States

With the specific frequencies of each detected peak, the original state of the Rubik's Cube at that moment in time can be computed by finding the states whose corresponding frequency is closest to each detected peak frequency.

This is done by first defining a function that will find the closest centerpiece state for one peak frequency by iterating through every centerpiece state in the FREQUENCY_MAPPINGS dictionary from Section 5.2.1, comparing the difference between the state's frequency and the given peak frequency, and returning the state with the smallest difference (Figure 5.10a). This function can then be called for all the detected peak frequencies to recover the state of the cube at that time step (Figure 5.10b).

FIGURE 5.10: Converting peak frequencies to centerpiece states

(A) Function to find the state most similar to the detected frequency

```

1 def _closest_state(detected_freq):
2     # Start with no assumptions regarding the closest state
3     closest_state = None
4     closest_difference = None
5     # Iterate through each possible state to find the closest one
6     for face, rotations in FREQUENCY_MAPPINGS.items():
7         for rotation, freq in rotations.items():
8             difference = abs(detected_freq - freq)
9             if closest_difference is None or difference < closest_difference:
10                 closest_difference = difference
11                 closest_state = dict(face=face, rotation=rotation)
12
13 return closest_state

```

(B) Function to determine the cube's state from a set of peak frequencies

```

13 def get_state_from_freqs(important_freqs: list) -> dict:
14     # Start with an unpopulated state
15     state = {}
16     # Record the state represented by each of the peak frequencies
17     for freq in important_freqs:
18         hz, power = freq.values()
19         closest_state = _closest_state(hz)
20         state[closest_state["face"]] = closest_state["rotation"]
21
22 return state

```

For example, Figure 5.11 shows how using these functions makes it easy to convert the peak frequencies extracted in Figure 5.9 to their corresponding centerpiece states.

FIGURE 5.11: Example conversion of peak frequencies to states

```

1 detected_state = get_state_from_freqs(important_freqs)
2 print(detected_state)

>> {"U": 0, "D": 3, "R": 0, "L": 0, "F": 0, "B": 0}

```

As such, obtaining a sequence of the Rubik's Cube's centerpiece states over the course of the recorded audio sequence only requires repeating this process for each time step in the spectrogram data. This can also be easily turned into another function as shown in Figure 5.12.

FIGURE 5.12: Function to list the cube's state at each time step

```

1 def get_state_over_time(freq, time, spectrogram):
2     # Start with an empty list of centerpiece states
3     state_over_time = []
4     # For each time step ...
5     for t_idx in range(len(time)):
6         # find the peak frequencies ,
7         important_freqs = extract_important_freqs(
8             freq, time, spectrogram, t_idx)
9         # convert the peak frequencies to the corresponding state ,
10        state = get_state_from_freqs(important_freqs)
11        # and save the state and time stamp for later processing .
12        state_over_time.append(dict(
13            time=time[t_idx],
14            state=state
15        ))
16    return state_over_time

```

And, for completeness, Figure 5.13 shows an example of using that function to get the full sequence of states for the synthetic audio generated in Section 5.2.

FIGURE 5.13: Example listing of states over time

```

1 state_over_time = get_state_over_time(freq, time, spectrogram)
2 print(state_over_time)

>> [{"time": 0.0, "state": {"U":0, "D":0, "R":0, "L":0, "F":0, "B":0}}, 
      {"time": 0.0058 "state": {"U":0, "D":0, "R":0, "L":0, "F":0, "B":0}}, 
      ... ]

```

5.3.4 Extracting Move Sequences from Centerpiece State Sequences

The final step to recover the original move sequence is to iterate over the sequence of cube states and register any change to the cube state as a move applied to the cube.

This starts with a function (Figure 5.14a) that, given a starting and ending rotational state for a specific face, can calculate which direction a face was turned and return the text notation of the applied face turn. A second function (Figure 5.14b) then iterates through the state sequence extracted in Section 5.3.3 checking each state in the sequence against the previous one to detect when the state changes. Upon detecting a change, this second function then calls the first to get the actual move that caused the state change and saves it to a list of `detected_moves` to be returned after iterating through all states. This list of `detected_moves` is the final move sequence that was extracted from the synthetic audio.

FIGURE 5.14: Code to extract move sequences from state sequences

(A) Function to determine the face turn between two centerpiece states

```

1 def _move_from(face, current_rotation, new_rotation):
2     direction = None
3     # Check whether a Clockwise or Counterclockwise turn occurred
4     # using the same logic as 'RubiksCube.apply_move'
5     if (current_rotation + RubiksCube.CLOCKWISE) % 4 == new_rotation:
6         direction = "" # Clockwise
7     elif (current_rotation + RubiksCube.COUNTERCLOCKWISE) % 4 == new_rotation:
8         direction = '/' # Counterclockwise
9     # Create and return the notation for the applied face turn
10    return f"{face}{direction}"

```

(B) Function to extract all moves from a centerpiece state sequence

```

11 def detect_moves(state_over_time):
12     detected_moves = []
13     current_state = {}
14     # Iterate through all of the extracted cube states
15     for idx, timed_state in enumerate(state_over_time):
16         time, state = timed_state.values()
17         for face, rotation in state.items():
18             # Save the first state as the initial state
19             if not (face in current_state):
20                 current_state[face] = rotation
21             # If the state of any face changes...
22             if rotation != current_state[face]:
23                 # ... get and save the applied move with its timestamp...
24                 detected_moves.append(dict(
25                     time=time,
26                     move=_move_from(face, current_state[face], rotation)
27                 ))
28                 # ... and update the tracked state based on the new move.
29                 current_state[face] = rotation
30     return detected_moves

```

Figure 5.15 shows an example usage of the `detect_moves` function. Additionally, the extracted Rubik's Cube algorithm is compared to the original one used to create the synthetic audio.

FIGURE 5.15: Example move sequence extraction

```

1 detected_moves = detect_moves(state_over_time)
2
3 pretty_moves = " ".join([i["move"] for i in detected_moves])
4 print(pretty_moves)
5 print(f"Matches_demo_alg? {demo_alg == pretty_moves}")

>> U U U U D D D D R R R R L L L L F F F F B B B B
>> Matches demo_alg? True

```

Clearly, the detected move sequence matches the demo algorithm used to create the synthetic audio in Section 5.2.3, which means this algorithm is a functional receiver for an ideal transmitter.

5.3.5 Full Example: Extracting Moves from Synthetic Audio

As a summary of this section on decoding synthetic audio, Figure 5.16 shows the full process for creating a synthetic audio clip and extracting the encoded move sequence from it. For this example, the chosen move sequence to transmit consists of every possible face turn, which will test that this strategy can correctly report back each face turn.

FIGURE 5.16: Full Example: Audio generation and move extraction

```

1 # Create the synthetic audio for a new algorithm
2 demo2_alg = "U'U'D'D'R'R'L'L'F'F'B'B"
3 demo2_wav_path = "demo_all_turns.wav"
4 render_audible_alg(demo2_alg, demo2_wav_path)
5
6 # Extract the move sequence from the new synthetic audio
7 freq2, time2, spectrogram2 = compute_spectrogram(demo2_wav_path)
8 state_over_time2 = get_state_over_time(freq2, time2, spectrogram2)
9 detected_moves2 = detect_moves(state_over_time2)
10
11 # Print out the results
12 pretty_moves2 = " ".join([i["move"] for i in detected_moves2])
13 print(pretty_moves2)
14 print(f"Matches_demo2_alg? {demo2_alg == pretty_moves2}")

>> U'U'D'D' R'R'L'L' F'F'B'B'
>> Matches demo2_alg? True

```

And again, the extracted move sequence does match the one used to generate the synthetic audio, further validating this strategy for decoding moves from audio.

5.4 Creating Realistic Audio

While the algorithm created in Section 5.3 can successfully recover a sequence of moves applied to a virtual Rubik's Cube with 100% accuracy, the synthetic audio that it operates on is not realistic. This is made obvious by a comparison between the spectrograms of the background audio in Figure 4.1 and the synthetic audio in Figure 5.5. In the latter, the bright yellow indicators of present frequencies are very clear and strong with no other frequencies present in the signal. In contrast, the former contains many areas with bright pink indicators of prominent frequencies (the color differences are due to the use of different applications to generate the diagrams). A more realistic signal would contain both the strong bands of the transmitted frequency along with the underlying background noise of whatever cube is actively being solved.

Adding this noise could be done in software by overlaying the synthetic audio with a recording of the background noise, but it was just as easy to play the synthetic audio from a laptop speaker and record it using a nearby smartphone while solving various Rubik's Cubes. An example of the spectrogram resulting from recording the audio of the demo algorithm from Section 5.2.3 on a Google Pixel smartphone while actively solving a Gans 356 speedcube can be seen in Figure 5.17.⁴

FIGURE 5.17: Spectrogram of a realistic audio sample



Notice how the horizontal bands representing the transmitted signal are dimmer in the realistic audio than in the purely synthetic audio. This reflects the loss of volume any tone experiences as it travels through the air. Additionally, the transmitted signal is also partially obscured by the additional audible noise of solving the speedcube.

5.5 Decoding Realistic Audio

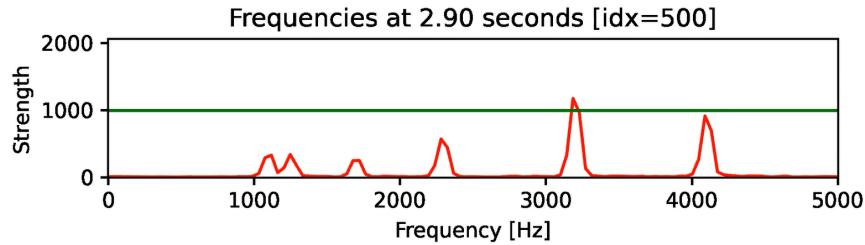
The loss of signal strength alongside the added noise creates several unique challenges that require more sophisticated analysis than that presented in Section 5.3. These challenges include added variation in the strength of each peak frequency (Section 5.5.1), the detection of conflicting centerpiece states (Section 5.5.2), and the erroneous reading of background noise as centerpiece states (Section 5.5.3).

⁴The astute reader will notice that the audio bands in Figure 5.17 fall on slightly different frequencies than in Figure 5.5. This is because the realistic audio samples used for analysis in Sections 5.4 and 5.5 were created before finalizing the frequency assignments presented Table 4.1. Ultimately, this change in frequencies shows that this algorithm works across a variety of encoding frequencies.

5.5.1 Fine-Tuning the Threshold

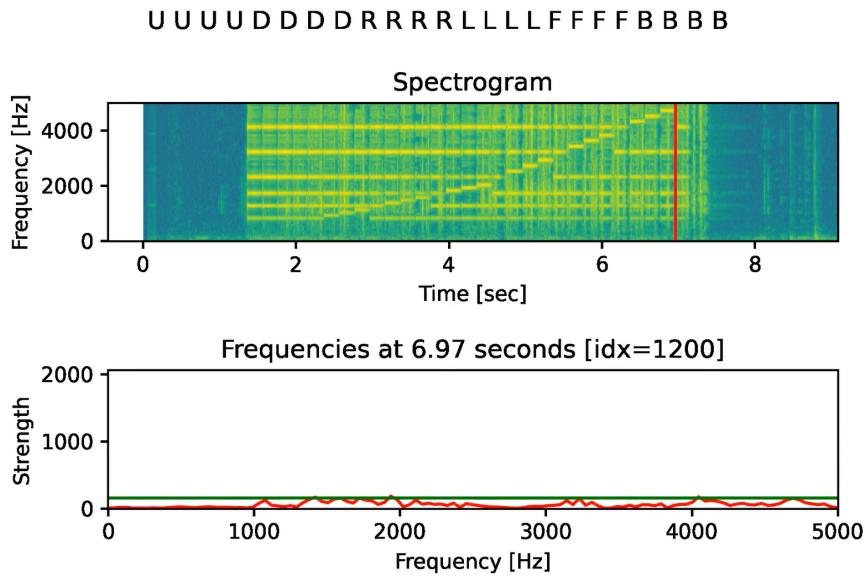
Because some audio frequencies are damped more than others as they travel through the air to the recording microphone, a static threshold like the one used in Section 5.3.2 fails to capture all the dominant frequencies that compose the audible signal. For example, Figure 5.18 shows how the 85% threshold from Section 5.3.2 misses five of the six signal peaks in the realistic audio.

FIGURE 5.18: An 85% threshold applied to realistic audio



Furthermore, the natural background noise can also cause false positives during the moments between moves while the signal is weaker as a result of changing from one state to the next. For example, Figure 5.19 shows a moment between face turns when the audio signal is nearly non-existent, and the background noise can falsely register as important peak frequencies for centerpiece state detection. In this case, between 1000Hz and 2000Hz alone there are six different points where the background noise reaches the 85% threshold compared to the three centerpiece states that get transmitted within that same band.

FIGURE 5.19: Dominance of background noise between moves



Resolving these issues starts by noticing that the power of each peak frequency bearing the audio signal is generally much higher than the power of all other component frequencies. As such, basing the threshold at a power equal to one standard deviation of the power of the component frequencies of a specific time step generally captures all the peak frequencies while still excluding the underlying noise of the cube/environment. Additionally, the use of a hard minimum value for the threshold helps reduce the number of false positives during the gaps in the signal between face turns.

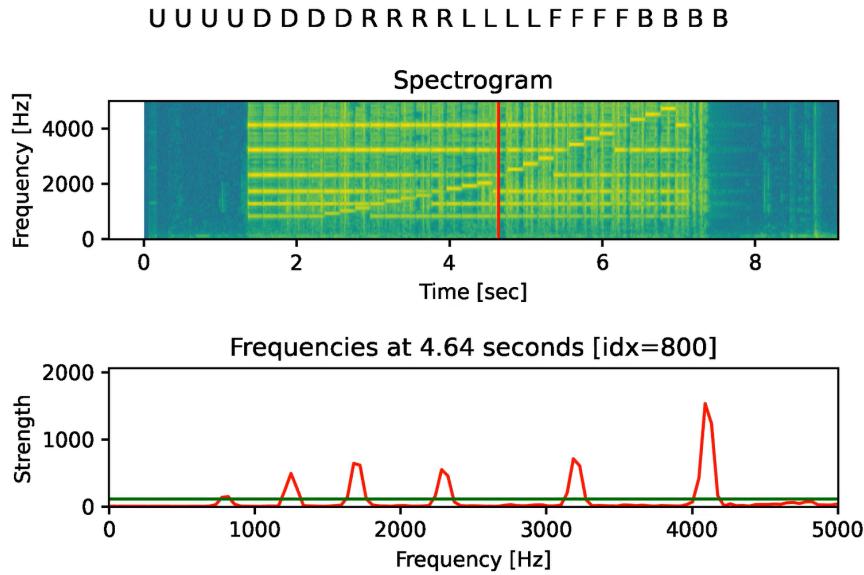
These two changes are encoded in an updated version of the `compute_threshold` function first defined in Section 5.3.2. This new version contains two new parameters: `stdv_pct`, which can be used to adjust the threshold to some multiple of the standard deviation for testing purposes, and `min_thresh` which is the "hard minimum value" discussed in the previous paragraph.

FIGURE 5.20: Updated version of threshold computation in Figure 5.8a

```
1 def compute_threshold(values: list, stdv_pct: float=1, min_thresh: int=50):
2     return max(min_thresh, np.std(values) * stdv_pct)
```

An example of the thresholds yielded by this new computation is shown in Figure 5.21, where the green line representing the threshold successfully captures all six peak frequencies of the audio signal while simultaneously staying just out of reach from the frequencies of the background noise.

FIGURE 5.21: Fine-Tuned Threshold



5.5.2 Filtering through Similar Peak Frequencies

While the new threshold calculation does properly capture all the peak frequencies, it also captures more samples than just the six tips of each peak. These extra samples can cause confusion when trying to extract each centerpiece's state if they correspond to different states for the same centerpiece. That said, as long as the power of each frequency is also saved, then any disagreements about a specific centerpiece's state can be resolved by accepting the one with the most intense frequency as the actual state.

In code, this is achieved by adding a temporary dictionary to the `get_state_by_freqs` method defined in Section 5.3.3 to track the strongest frequency associated with each face. Then, while iterating through the captured peak frequencies, only the strongest one for each face is used to determine the active centerpiece state at that time step.

FIGURE 5.22: Updated version of state extraction in Figure 5.8b

```

1 def get_state_from_freqs(important_freqs: list) -> dict:
2     state = {}
3     state_power = {"U": 0, "D": 0, "R": 0, "L": 0, "F": 0, "B": 0} # New
4     for freq in important_freqs:
5         hz, power = freq.values()
6         closest_state = _closest_state(hz)
7         if power > state_power[closest_state["face"]]: # New
8             state[closest_state["face"]] = closest_state["rotation"]
9             state_power[closest_state["face"]] = power # New
10    return state

```

Using this new version of `get_state_by_freqs` works just like it did in Figure 5.11. Here the extracted state is the one encoded by the frequencies depicted in Figure 5.21.

FIGURE 5.23: Example: Refined conversion of peak frequencies to states

```

1 transmitted_wav_path = "transmitted -356-5tps.wav"
2 freq, time, spectrogram = compute_spectrogram(transmitted_wav_path)
3 important_freqs = extract_important_freqs(freq, time, spectrogram,
4 t_idx=800) # 4.64 seconds
5
6 detected_state = get_state_from_freqs(important_freqs)
7 print(detected_state)

>> {"U": 0, "D": 0, "R": 0, "L": 0, "F": 0, "B": 0}

```

5.5.3 Ignoring Noise when Extracting Move Sequences

However, despite these noise filtering measures, the background noise occasionally causes the detection of an incorrect centerpiece state. Since the approach in Section 5.3.4 registers an applied face turn for *any* detected change in a centerpiece's state, *any* mis-detection would incorrectly register a face turn that never happened.

Fortunately, this issue can be mitigated by requiring that the state change persists over several time steps instead of blindly accepting any detected change in a centerpiece's state as a new face turn.

This is implemented by adding a sliding window to the `detect_moves` function first created in Section 5.3.4. The window implementation starts with a new `window_size` function parameter to control the number of consecutive time steps over which a state change has to persist before it is recorded as a move applied to the cube. Within the function, two new dictionaries are created to serve as a "staging area" for state changes: one to stage the new state and the second to record the index of the time step where the state first changed. As the function iterates through each time step, it updates these staging dictionaries each time a new state is detected. Then, it checks to see if the current index of iteration is `window_size` steps after the last reported state change. If so, the current state is updated to the value of the staged state, the change is recorded as a new move, and the window counter gets reset. Otherwise, the iteration continues until a new state is detected or the `window_size` is reached. The full code for this is shown in Figure 5.24.

FIGURE 5.24: Updated version of move detection in Figure 5.14b

```

1 def detect_moves(state_over_time , window_size=8):                      # Edit
2     detected_moves = []
3     current_state = {}
4     new_state = {"U": -1, "D": -1, "R": -1, "L": -1, "F": -1, "B": -1}    # New
5     new_state_idx = {"U": 0, "D": 0, "R": 0, "L": 0, "F": 0, "B": 0}      # New
6     for idx, timed_state in enumerate(state_over_time):
7         time, state = timed_state.values()
8         for face, rotation in state.items():
9             # Update new_state whenever the rotation changes           # Edit ↓
10            if rotation != new_state[face]:
11                new_state[face] = rotation
12                new_state_idx[face] = idx
13            # If the new state has persisted over window_size time steps ,
14            if idx - new_state_idx[face] == window_size:
15                if not (face in current_state):
16                    current_state[face] = new_state[face]
17                # and it is a different rotation than the current state ,
18                elif new_state[face] != current_state[face]:
19                    # then a face turn has been detected !
20                    detected_moves.append(dict(
21                        time=state_over_time[new_state_idx[face]]["time"],
22                        move=_move_from(face , current_state[face], rotation)
23                    ))
24                    current_state[face] = rotation
25                    new_state_idx[face] = 0
26    return detected_moves

```

And to validate that these changes work as expected, Figure 5.25 runs the realistic audio through this new `detect_moves` function and compares the returned move sequence with the one originally transmitted in the audio.

FIGURE 5.25: Example: Refined move sequence extraction

```

1 state_over_time = get_state_over_time(freq, time, spectrogram)
2 detected_moves = detect_moves(state_over_time)
3
4 pretty_moves = " ".join([i["move"] for i in detected_moves])
5 print(pretty_moves)
6 print(f"Matches demo_alg? {demo_alg == pretty_moves}")

>> U U U U D D D D R R R R L L L L F F F F B B B B
>> Matches demo_alg? True

```

And once again, the detected move sequence perfectly matches the transmitted signal despite both the reduced strength of the synthetic audio after being recorded through the air and the presence of the added noise of solving a speedcube, a result which further validates this approach as a viable proof-of-concept for tracking the moves of a Rubik's Cube using sound.

5.5.4 Optimizing Algorithm Parameters

Significant testing went into determining the best default values for each of the three new function parameters `stdv_pct` (Section 5.5.1), `min_thresh` (Section 5.5.1), and `window_size` (Section 5.5.3). While each cube responded differently to various combinations of settings, this testing discovered multiple combinations for each cube that would yield a perfect extraction of the original move sequence. A detailed exploration of this testing and its findings is discussed in Section 7.1.

5.6 Summary

In summary, this chapter demonstrated a proof of concept for a sound analysis algorithm that could detect the face turns of a speedcube equipped with the proper transmitter (the details of which are explored in Chapter 6). This proof of concept algorithm works by computing the spectrogram of the transmitted audio, extracting the most intense frequencies from each time step, converting those dominant frequencies to centerpiece states at that time step, then iterating through that list of states over time to detect changes caused by face turns. Several noise mitigation measures were employed to help mitigate the impact of loud cubes and other background tones. Ultimately, this algorithm design successfully decoded both a purely synthetic audio signal and a realistic audio signal created by re-recording the synthetic audio while solving a speedcube.

Chapter 6

The Transmitter

6.1 Introduction

The transmitter for the sound-based move tracking protocol is in charge of creating the tones representing each centerpiece's current state, and updating those tones each time a centerpiece changes state.

This chapter will detail a proof-of-concept design for a printed circuit board (PCB) capable of generating all the tones required for encoding one centerpiece's state.

This chapter assumes that the reader has a knowledge of basic circuit components like resistors and capacitors. It begins with a discussion of the requirements/constraints for the physical transmitter (Section 6.2). From there a formal design schematic for the transmitter will be presented (Section 6.3) and prototyped (Sections 6.4) followed by an analysis of the effects of the centercap enclosure on signal strength (Section 6.5). Finally, all the required components for the transmitter will be laid out inside a single speedcube centercap to demonstrate the viability of the presented design (Section 6.6).

6.2 Requirements

This section will detail the constraints within which the transmitter will be required to operate. These constraints include the precision of tone generation (Section 6.2.1), the reliability of the transmitter in changing tones to reflect a face turn (Section 6.2.2), the intensity of output audio that the transmitter can produce (Section 6.2.3), and the physical size of the transmitter (Section 6.2.4).

6.2.1 Precision of Tone Generation

While the receiver specified in Chapter 5 supports custom state to frequency mappings, it expects that the frequency corresponding to each centerpiece's state stays constant throughout the entire audio recording. As such, the chosen transmitter design can encode centerpiece states with any frequency (assuming the chosen frequencies work within the constraints specified in Section 4.2), but it must produce its chosen frequencies with high precision.

6.2.2 Responsiveness to Face Turns

The chosen transmitter design must respond to an applied face turn by changing the currently transmitted audio frequency to the frequency corresponding to the new centerpiece state. Given top speedcubers' burst turn speeds of 20 TPS (see Section 4.4) this frequency change must complete and stabilize long enough to be accurately detected within one twentieth of a second (50ms).

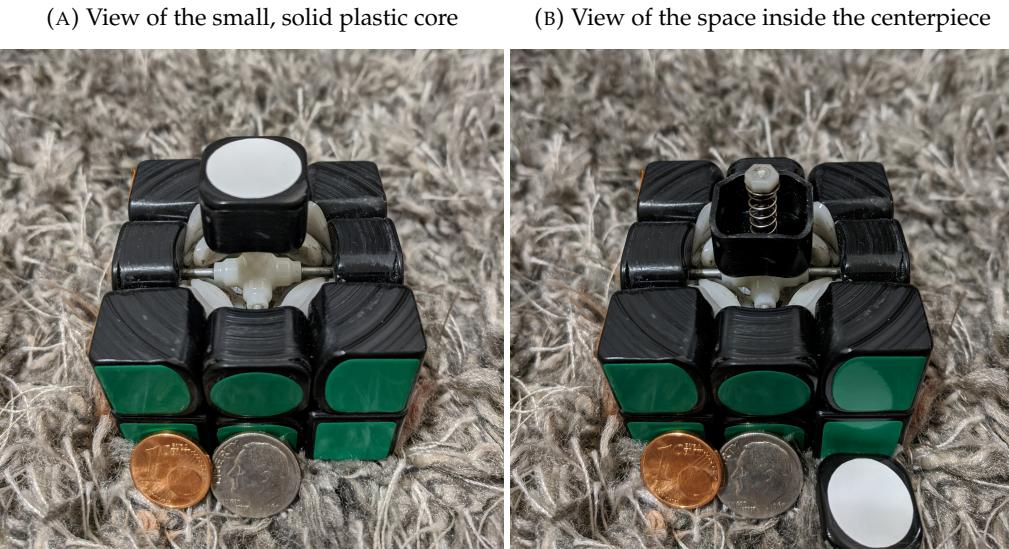
6.2.3 Signal-to-Noise Ratio

The transmitter must create tones loud enough to be easily distinguished from ambient noise, including the sound of the Rubik's Cube's own turns. In light of the above requirement for the transmitter to fit within a center cubie (6.2.4), this requirement will also require the transmitter design to consider how to overcome any audio dampening caused by such an enclosure.

6.2.4 Prospects of Miniaturization

The transmitter must be both removable and small enough to fit in the center cap of each face of a speedcube. This requirement stems from two sources. First, in contrast to all existing smartcubes, most non-smartcubes have small, solid cores (Figure 6.1a) that provide no extra space for the inclusion of any electronics, but do have a small amount of open space within their center cubies (Figure 6.1b). Second, the use of a cube with non-removable, embedded electronics violates the WCA competition regulation 2i [7] (See also Section 2.3.2).

FIGURE 6.1: Internal pieces of a standard speedcube (Gans 356)



6.3 Design

Given the above constraints, this section will detail a design for a printed circuit board capable of precisely generating four distinct tones (Requirements 6.2.1 and 6.2.2)

6.3.1 The 555 Timer

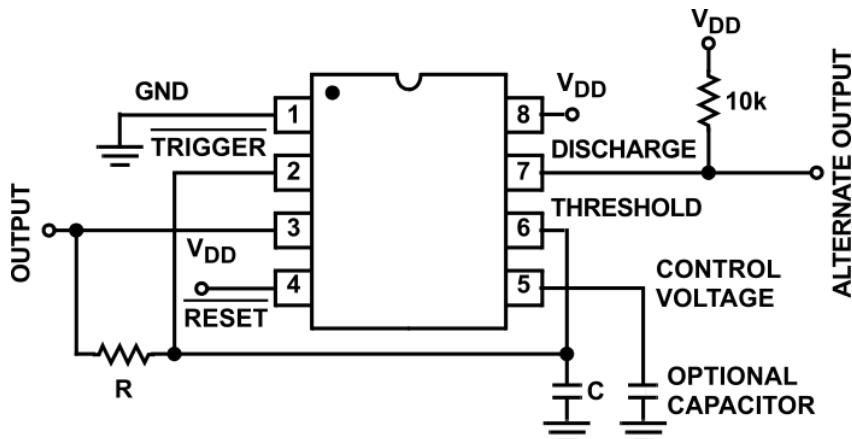
The core component in this PCB design is a 555 timer, which is a computer chip that facilitates the generation of many types of voltage frequencies in a circuit.

For this transmitter, the 555 timer will be configured to output a square wave voltage signal with a 50% duty cycle (i.e. "Astable Operation") [49]. This means the voltage on the output pin will alternate between low and high, spending equal amounts of time in each state.

The attentive reader will notice that the square wave signal proposed here differs from the sine wave used to create the synthetic audio in Figure 5.3. Valid concerns may even be raised about the fact that a square wave is actually a composite of a fundamental sine wave and infinitely many harmonics [50]. However, since the nearest harmonic in a square wave with a 50% duty cycle oscillates at a frequency three times as fast as the fundamental [51], then choosing a fundamental frequency of at least 6.67kHz places the nearest harmonics beyond the range of frequencies measurable by a typical smartphone or laptop microphone (see Section 4.2.3).

The standard schematic for creating this type of voltage signal is depicted in Figure 6.2.

FIGURE 6.2: Standard 555 Timer Circuit for Astable Operation [49]



The two key components are the resistor R and the capacitor C whose respective resistance and capacitance control the output frequency f via the relation shown in Equation 6.1 [49]:

$$f = \frac{1}{1.4RC} \quad (6.1)$$

6.3.2 Creating Audio

The 555 timer's varying voltage output can be easily converted to audible sound by attaching a voltage controlled speaker to the output wire coming from pin 3 in Figure 6.2. However, this will only produce one continuous tone, and each centerpiece will need to be able to switch between four distinct tones. As such, the resistor R will need to be replaced with four separate resistors (labeled R_1 , R_2 , R_3 , R_4 in Figure 6.3) and a switch. The switch will be connected to the cube so that a 90° rotation will change which resistor is in series with the circuit, thus changing the output audio frequency of the speaker.

FIGURE 6.3: Centerpiece State Transmitter Circuit



Alternatively, it is mathematically valid to instead switch between four different capacitors of different capacitance. However, for the reasons described in Section 6.3.3, opting to switch between resistors proved more practical.

6.3.3 Choosing the values of R and C

There are infinitely many combinations of R and C that will produce any single desired output frequency f ; however, the set of resistances and capacitances of commonly produced resistors and capacitors is finite. While multiple components can be combined to produce more specific resistances and capacitances, doing so would increase the overall component count for a circuit board that is constrained to a small physical size. As such, this proof-of-concept design will focus on producing frequencies attainable with a single capacitor and resistor pair.

Since each centerpiece transmitter needs to produce four distinct frequencies, four separate capacitor-resistor pairs are required. However, since the output frequency can be varied by adjustments to *either* the capacitance or resistance, one of those two options can be held constant to reduce the overall component count.

To chose which one to hold constant, the output frequencies of all possible pairings of capacitors and resistors from two cheap Amazon kits [52] [53] were calculated using

Equation 6.1. The resulting table was then color coded to highlight the usable frequencies in green (6.67kHz to 20kHz)¹ while leaving all other unusable frequencies in red. The result is shown in Figure 6.4 with the various resistances shown on the vertical axis and the various capacitances along the horizontal axis.

A close observation of the resulting sigmoid shape of usable frequencies reveals only one location where there are at least four usable frequencies associated with a fixed resistance (i.e. row of green cells), compared to ten locations where there are at least four usable frequencies associated with a fixed capacitance (i.e. column of green cells). As such, the most practical value to hold constant here is the capacitance.

One of the locations with a viable fixed capacitance is shown in Figure 6.5. This pairing of a 10nF capacitor with four resistors with respective resistances of 4.7kΩ, 5.1kΩ, 5.6kΩ, and 7.5kΩ will be used in the prototype created in Section 6.4.

FIGURE 6.5: Output Frequencies of Common R and C Values

	4.7E-09	1.0E-08	4.7E-08
3000	50658.56	23809.52	5065.86
4700	32335.25	15197.57	3233.53
5100	29799.15	14005.60	2979.92
5600	27138.51	12755.10	2713.85
7500	20263.42	9523.81	2026.34
8200	18533.62	8710.80	1853.36

6.4 Prototyping

The key step to produce a proof-of-concept for Requirements 6.2.1 and 6.2.2 is to put everything together on a breadboard and turn on the power while recording the output audio to make sure four distinct tones are produced.

This step begins by gathering all the components from the schematic in Figure 6.3 with the specific values of C, R1, R2, R3, and R4 shown in Figure 6.5. These components are then placed onto a physical breadboard in accordance with the schematic's defined layout. Connecting VDD and GND to power then causes the speaker to produce a tone. With power still connected, the switch can be moved to put any other resistor in series to change the resulting tone of the speaker.

The result is shown in Figure 6.6. The left section of the figure shows a spectrogram of the four distinct tones produced by moving the green wire (representing the rotary switch) through each of the four resistors on the breadboard in the right section of the figure.

FIGURE 6.6: Breadboard Prototype

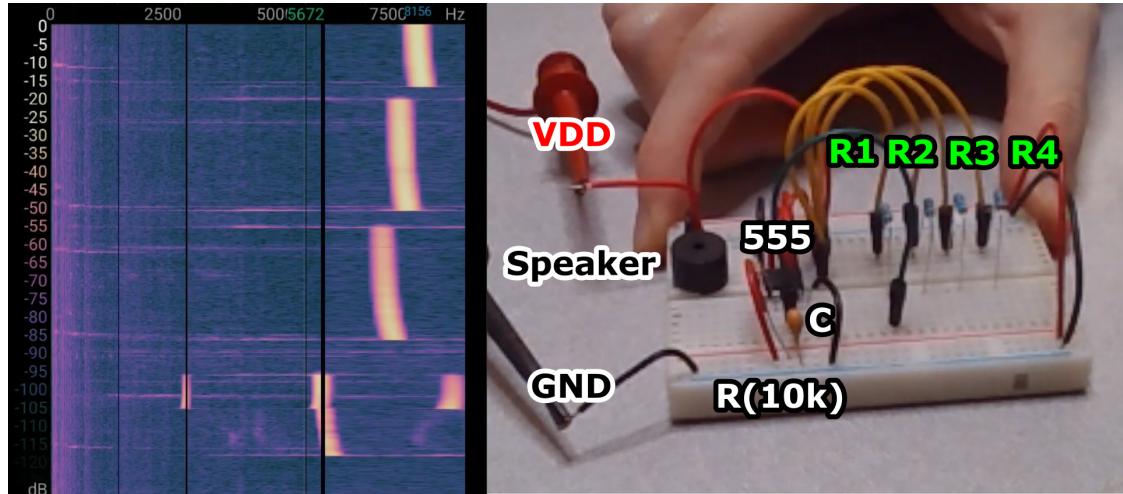
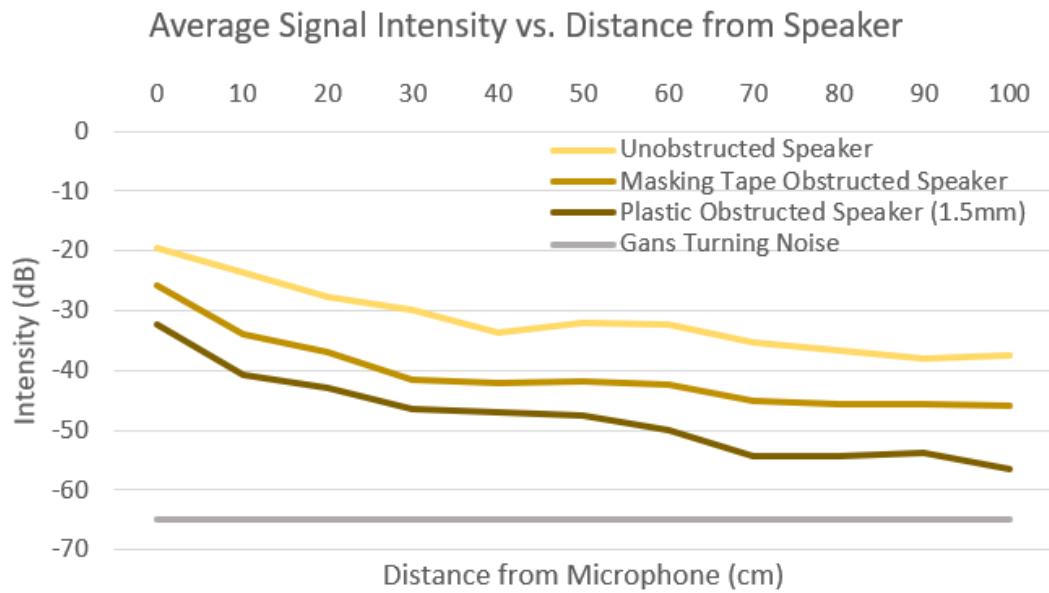


FIGURE 6.7: Effects of Various Sound Obstructions on Signal Intensity



6.5 Minimizing Sound Obstruction

In preparation for miniaturizing the prototype circuit from Section 6.4 to fit inside a speedcube's centercap, an experiment was performed to measure the effects of various physical barriers on the signal-to-noise ratio of a small speaker against background noise created by solving a Gans 356 speedcube (Requirement 6.2.3).

Specifically, three different types of obstructions were placed between a speaker similar to the type that would fit inside a speedcube centercap and a phone microphone: no obstruction, a piece of masking tape over the speaker (representative of a sticker on a speedcube), and a full plastic enclosure of the same thickness as a speedcube centercap. For each of these obstructions, the source speaker emitted eight different tones across the human audible spectrum and the average intensity of the signal was measured with the microphone at 10cm increments up to a full meter away from the speaker.

Figure 6.7 shows the resulting signal strength for each obstruction type (the top three yellow lines) compared with the background noise caused by solving a Gans 356 speedcube (the bottom gray line). As expected, the signal strength tapered off for all three obstruction types as the microphone moved farther away from the signal, and larger obstructions reduced the signal strength more than smaller obstructions.

Of most interest are the signal intensities between 10cm and 30cm from the microphone since that's the typical distance between a speedcuber's cube and his/her smartphone or laptop while practicing at home. By the far end of that range, only the completely unobstructed speaker's signal stayed above the -30dB limit discussed in Section 4.2.1. As such, it is clear that successfully miniaturizing the prototype circuit to fit inside a centercap will require a custom centercap design with space for sound to exit unobstructed and/or a speaker significantly louder than the ones used in these prototypes.

As a tangential note, the fact that the unobstructed speaker's signal strength drops to nearly -40dB at only 100cm away from the speaker means that the cube's moves cannot be reliably tracked while the cube is at a distance from the receiving microphone. This property may prove beneficial for competitive environments as a potential path towards allowing automated move tracking during competitive solves without incurring the risks of competitors using smartcubes as a way to spy on the scrambling process for an unfair advantage.

6.6 Miniaturization

Finally, to address Requirement 6.2.4, this section will demonstrate that all the required components can fit within the limited 16mm^2 of space inside a single center cap of a Gans 356 Speedcube.

Meeting this size constraint is best achieved by swapping out the large through-hole resistors, capacitors, and 555 timer used in the prototype for smaller surface mount (SMD) components. Additionally, the large black buzzer can be swapped out for a much thinner piezo buzzer that can lay on flat on top of a custom-designed replacement centercap.

¹6.67kHz is the lower bound derived in Section 6.3.1 and 20kHz the upper limit of the typical frequency response range discussed in Section 4.2.3

Finally, the DC power supply can be changed out for a button-cell battery oriented vertically.

An example of the final result could look like Figure 6.8. Notice that all required components easily fit inside of a custom designed center cap that fits perfectly into a Gans 356 speedcube without inhibiting the cube's freedom of rotation. Also notice the large hole in the center of the custom centercap visible in Figure 6.8a that provides an unobstructed path for the speaker's signal to exit the center cap.

The one required component not shown in Figure 6.8 is the rotary switch required to change the resistor actively in series with the circuit. This particular component could be modeled after the rotary switches placed on the reverse sides of the PCBs embedded within the centercaps of the commercial smartcubes shown in Section 3.2.

FIGURE 6.8: Transmitter Circuit Inside a Custom Gans 356 Centercap



6.7 Summary

In summary, this chapter demonstrated a proof of concept for a PCB capable of transmitting the current state of a Rubik's Cube via sound to the receiver described in Chapter 5. This proof-of-concept circuit works by using specific capacitors and resistors to produce a precise voltage frequency from a 555 timer that can be output into the audible spectrum by connecting a speaker. In addition to showing a working circuit on a breadboard, this chapter also described the results of an experiment showing the need for an unobstructed path between the transmitter and the receiver. Finally, this chapter showed how all the required components for the transmitter could reasonably fit inside a single centercap of a speedcube like the Gans 356.

Future researchers could build on the work of this chapter by engineering more refined prototypes to more completely validate if this proposed design is capable of meeting requirements 6.2.2, 6.2.3, and 6.2.4. Such refinements could start with a rotary encoder and/or fully miniaturizing the PCB and actuating the frequency change with the speedcube's rotation.

Chapter 7

Evaluation

This chapter seeks to measure the effectiveness of the design proposed in Chapters 4, 5, and 6. To do so, this chapter will use four key benchmarks derived directly from the research questions outlined in Section 3.5.

The benchmarks are as follows:

- *Move Tracking Accuracy*: How accurately can the design track the moves performed on the modified speedcube? (Section 7.1)
- *Compatibility with Standard Speedcubes*: How effectively can the design be deployed in a standard, non-smart speedcube without any permanent modifications or reducing the cube's performance? (Section 7.2)
- *Move Tracking Granularity*: What other metrics can each design provide (Section 7.3)?
- *Competition Legality*: To what extent is the design compliant with existing competition regulations? (Section 7.4)

7.1 Move Tracking Accuracy

For the design to be usable it must be able to track the moves of a traditional, non-smart speedcube. Furthermore, since untracked moves render the final move sequence invalid unless manually reviewed and corrected, it is critical that the design detect moves with high or perfect accuracy.

While Chapter 5 demonstrated a perfect detection of a transmitted move sequence, this section seeks to explore the limits of the software receiver. To do this, the receiver was subjected to a battery of tests consisting of running several noisy synthetic audio sequences (see Section 5.4) through the receiver algorithm with a variety of parameters and comparing the detected move sequence to the one used to generate the audio.

The noisy synthetic audio samples used in this testing were engineered to represent both the full range of face turns that could be performed on the cube and the turn speeds of speedcubers that average 12-30 seconds per solve. To that end, the "demo alg" from Figure 5.4 which sweeps through every possible centerpiece state was rendered at 2

TPS and 5 TPS¹ using the code shown in Figure 5.3. Then, those two audio samples were each augmented with the background noise of a quiet, normal volume, and noisy speedcube (respectively the Gans 356, Gans X, and QiYi QiMeng described in Section 4.2.1) using the techniques discussed in Section 5.4 for a total of 6 audio samples that serve as a basic cross-section of the most common variations in cubes and turn speeds.

With this representative sample set of audio signals in hand, the limits of the receiver could be tested. Each test comprised of setting the three parameters of the receiver to a specific combination of values (each chosen from a range of values determined empirically), running all six audio samples through the receiver, then comparing the detected move sequence with the actual move sequence for a percent similarity calculation. The standard deviation threshold parameter `stdv_pct` (Section 5.5.1) varied through from 0.25 to 3 in increments of 0.25 for a total of 12 values. The minimum threshold parameter `min_thresh` (Section 5.5.1) varied from 50 to 550 in increments of 50 for a total of 6 values. The window size parameter `window_size` (Section 5.5.3) varied from 1 to 10 time steps in increments of 1 for a total of 10 values. In total, the 12 standard deviation thresholds, 6 minimum thresholds, 10 window sizes, 2 turn speeds, and 3 cubes yielded 4,320 unique combinations with which to test the receiver algorithm.

The insights gained from analyzing the resulting data are discussed in the following subsections. Section 7.1.1 explores the influence of cube noisiness on detection accuracy. Section 7.1.2 does the same for the standard deviation threshold. Section 7.1.3 covers the minimum threshold. And Section 7.1.4 reviews the window size. The turn speed factor is discussed in all subsections as an additional angle for understanding each factor's unique influence on detection accuracy.

7.1.1 Influence of Cube Noisiness

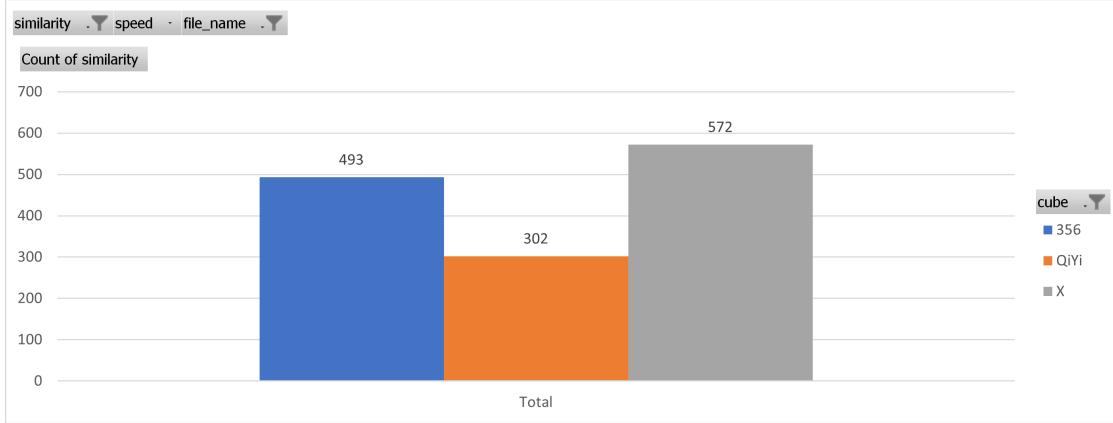
The three cubes used for this analysis - the Gans 356, Gans XS, and QiYi Qimeng - respectively represent quiet, normal, and loud speedcubes (see Section 4.2.1). The louder speedcubes were expected to cause more difficulties for decoding the transmitted move sequence since they create more noise that may conflict with the audible signal. Since all cubes produced more noise when turned at a higher speed and there is less time for detection to complete, the simulated turn speeds were also expected to experience inferior performance during move sequence detection.

Figure 7.1 shows the total number of perfect move sequence decodings broken down by cube type and simulated turn speed. As expected, there were far fewer perfect detections at the higher simulated turn speed compared to the lower simulated turn speed, with the loud QiYi even failing to register a single perfect detection at the higher turn speed. However, it was the normal noise Gans XS, not the quiet Gans 356 that achieved the most perfect detections. This is a peculiar result since Figure 4.1 shows that the Gans XS is louder across a wider frequency spectrum than the Gans 356.

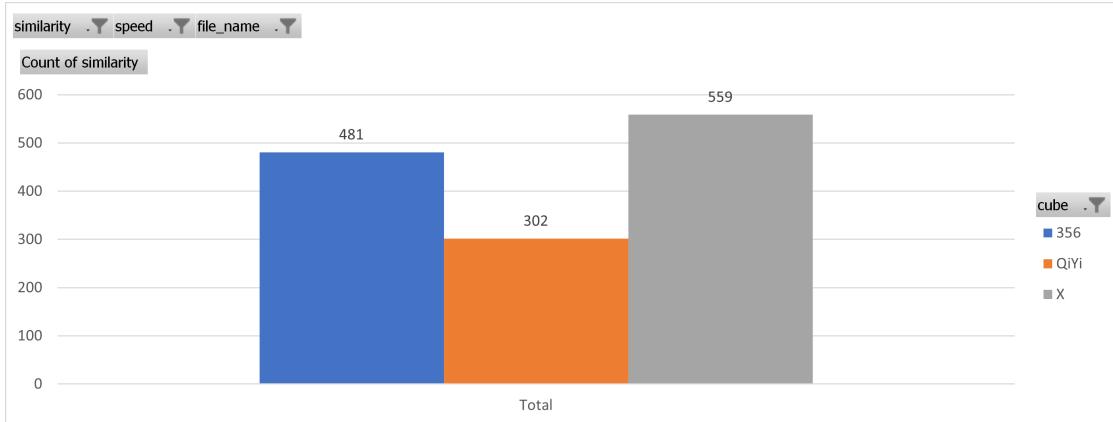
¹Given a that a typical speedsolver averages 60 moves per solve [54], 2TPS corresponds to an average solve time of 30 seconds and 5TPS corresponds to an average solve time of 12 seconds.

FIGURE 7.1: Perfect Detections by Cube

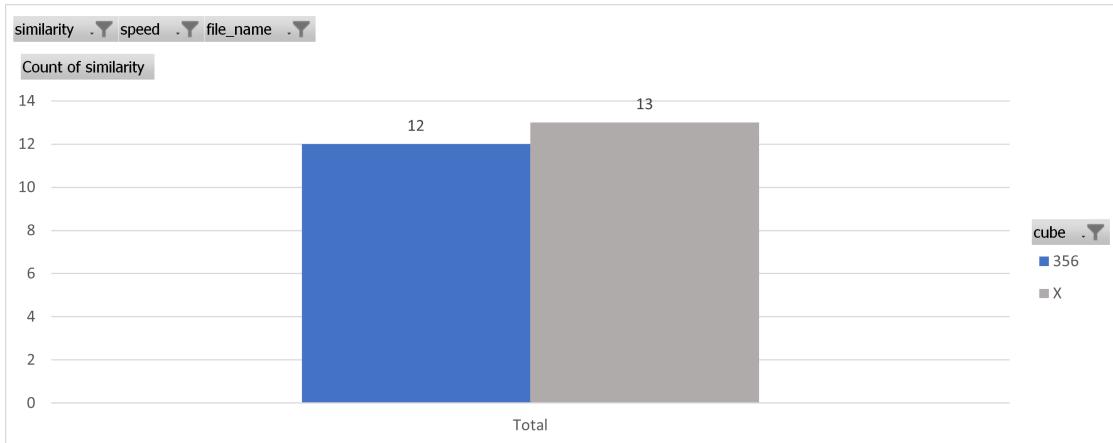
(A) Total number of perfect detections
(Aggregate of Figures 7.1b and 7.1c)



(B) Total number of perfect detections at 2TPS



(C) Total number of perfect detections at 5TPS



7.1.2 Influence of the Standard Deviation Threshold

The standard deviation threshold was one piece of the threshold calculation used to filter out background noise (see Section 5.5.1). Since frequency peaks tend to be statistical outliers, higher thresholds were expected to filter out more background noise, thus providing higher detection accuracy, so long as the thresholds were not so high as to exclude the peaks themselves.

Figure 7.2a shows the total number of perfect move sequence decodings broken down by the standard deviation threshold. Contrary to the expectations, the slightly sinusoidal results show that lower thresholds generally corresponded to slightly better chances of achieving a perfect detection than higher thresholds.

However, Figures 7.2b and 7.2c suggest that, overall, a higher standard deviation threshold typically corresponds to a higher average detection accuracy across all variations of the other parameters in this experiment. Of particular interest is the continuous improvement in detection accuracy experienced by the noisy QiYi Qimeng as the standard deviation threshold increased, particularly for the 5TPS audio sequences where it achieved a higher average accuracy than either of its less noisy counterparts.

7.1.3 Influence of Minimum Threshold

The minimum threshold was the second component in the threshold calculation discussed in Section 5.5.1, used primarily to filter out background noise during the quiet signal period between face turns. Similar to the standard deviation threshold, higher minimum thresholds were expected to yield higher detection accuracies, particularly for noisy cubes.

Interestingly, the minimum threshold performed similarly to the standard deviation threshold in regard to its correlation with perfect move sequence detections (Figure 7.3a). However, in contrast to the standard deviation threshold, the minimum threshold corresponded to far less variation in the average accuracy, though accuracy did trend upwards as the minimum threshold increased.

7.1.4 Influence of Window Size

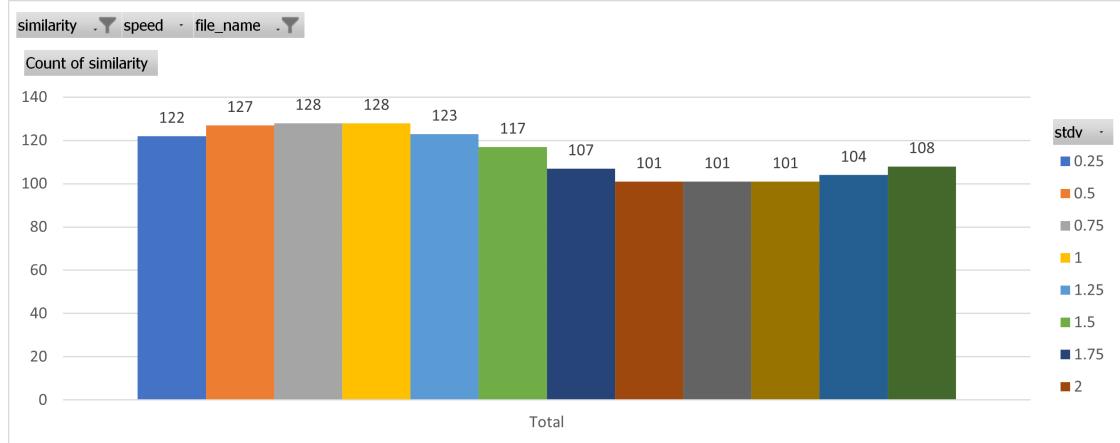
The final parameter tested in this experiment was the size of the sliding window used to ensure that detected face turns had stabilized before accepting them as real face turns. Since erroneous detections rarely last for more than a few time steps, a higher window size was expected to correspond to a higher detection accuracy.

Figure 7.4 clearly shows that a window size of five to ten time steps was consistently correlated with a higher accuracy of detection. In fact, at 2 TPS, a window size of five to ten corresponded to a nearly perfect *average* detection for all cubes, across the wide variation in the threshold parameters.

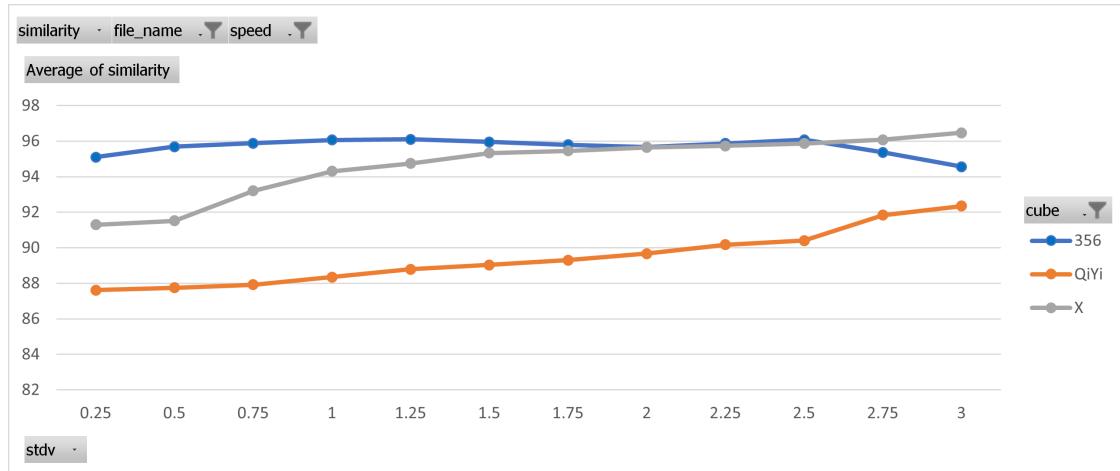
Furthermore, Figure 7.4c shows that a window size between five and eight time steps corresponded to >90% detection accuracy across all variations in the threshold parameters, three to ten percentage points higher than the best averages achieved by either threshold parameter.

FIGURE 7.2: Influence of Standard Deviation Threshold on Accuracy

(A) Perfect Detections by Standard Deviation Threshold



(B) Average Accuracy by Standard Deviation Threshold at 2 TPS



(C) Average Accuracy by Standard Deviation Threshold at 5 TPS

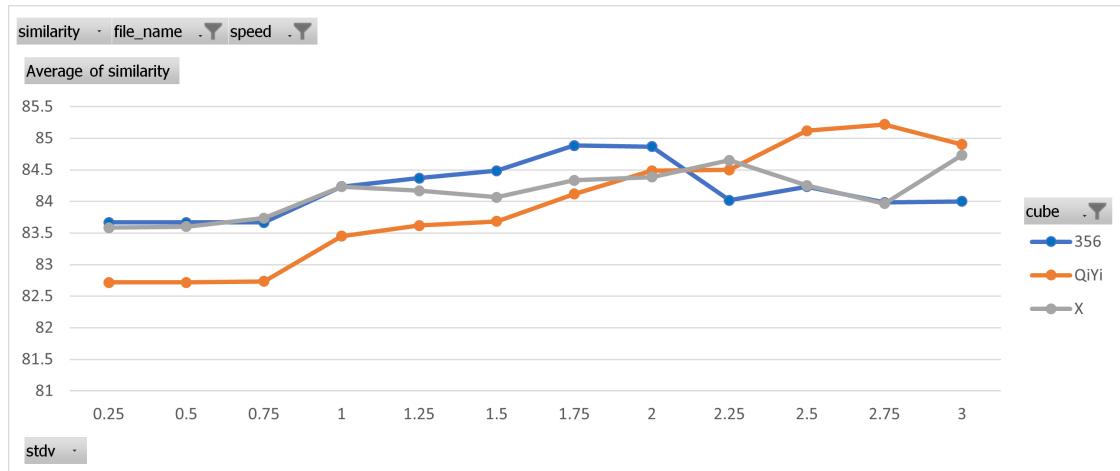
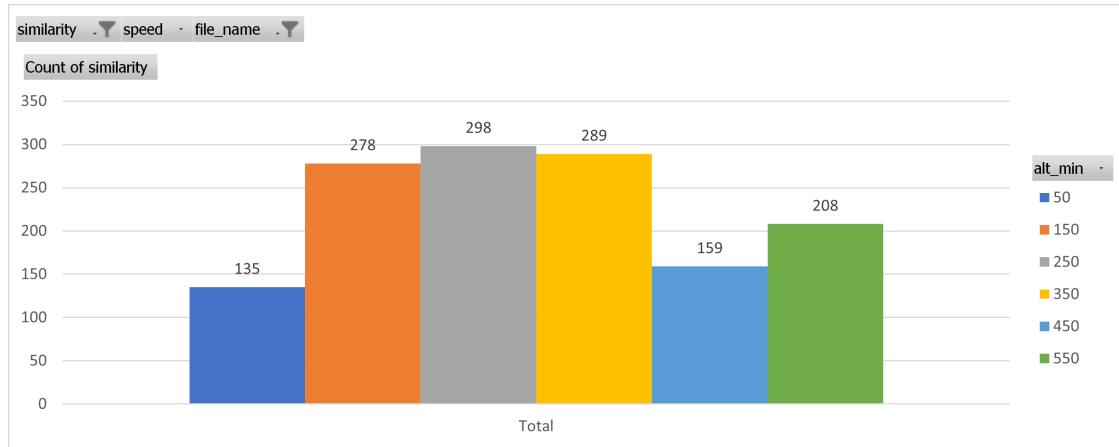
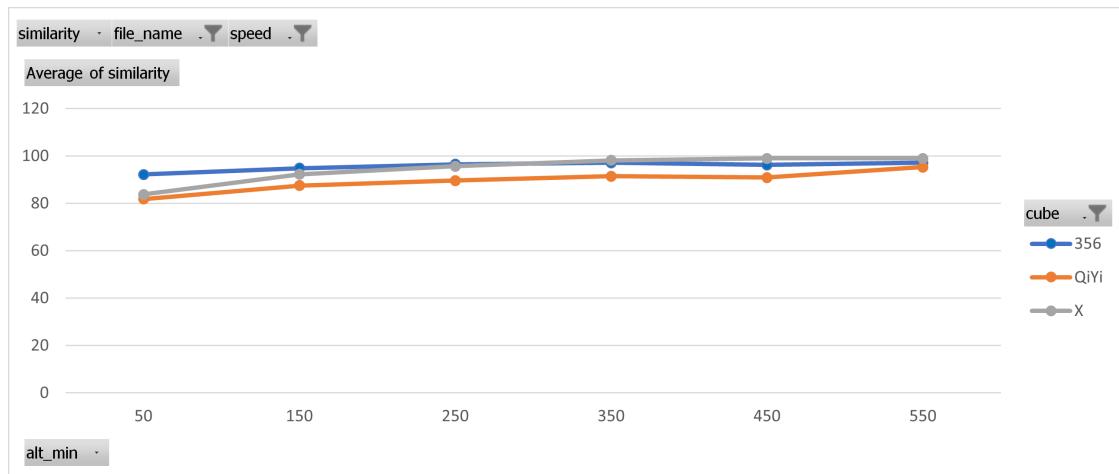


FIGURE 7.3: Perfect Detections by Minimum Threshold

(A) Total number of perfect detections



(B) Average Accuracy by Minimum Threshold at 2 TPS



(C) Average Accuracy by Minimum Threshold at 5 TPS

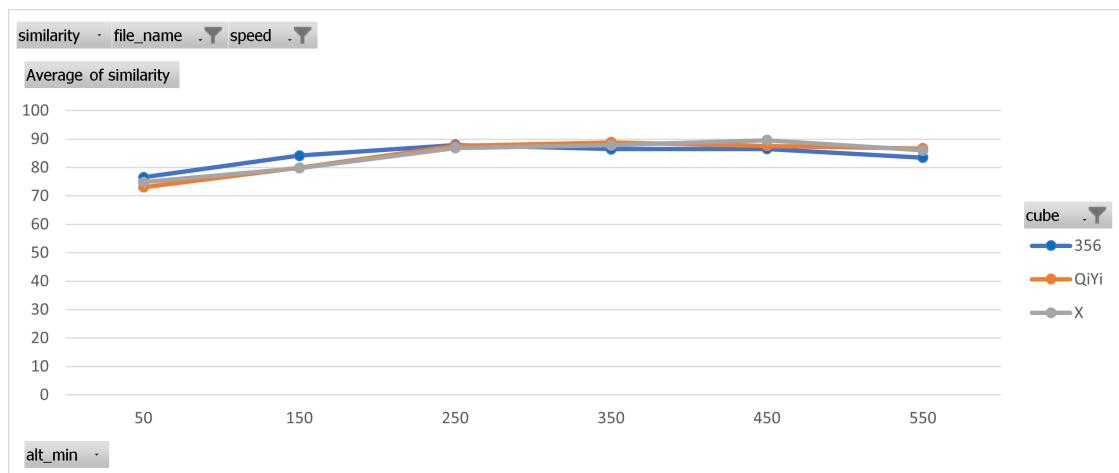
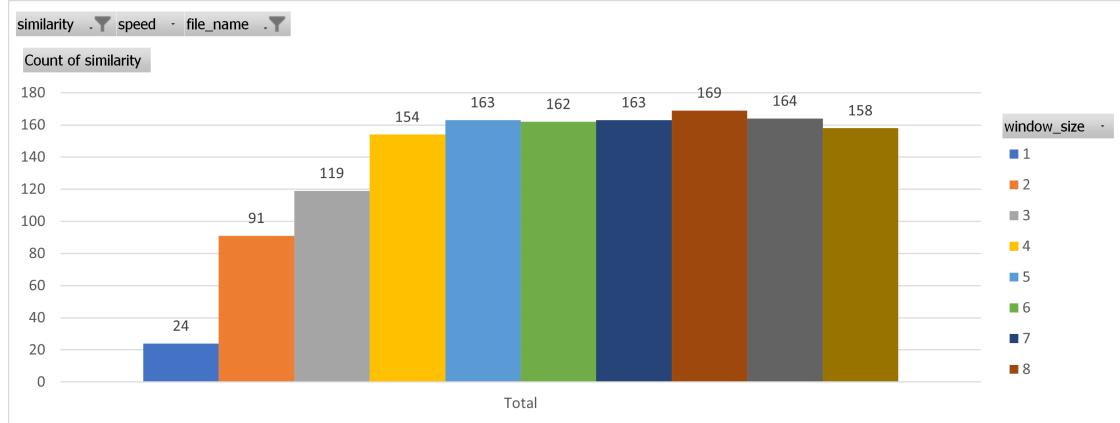
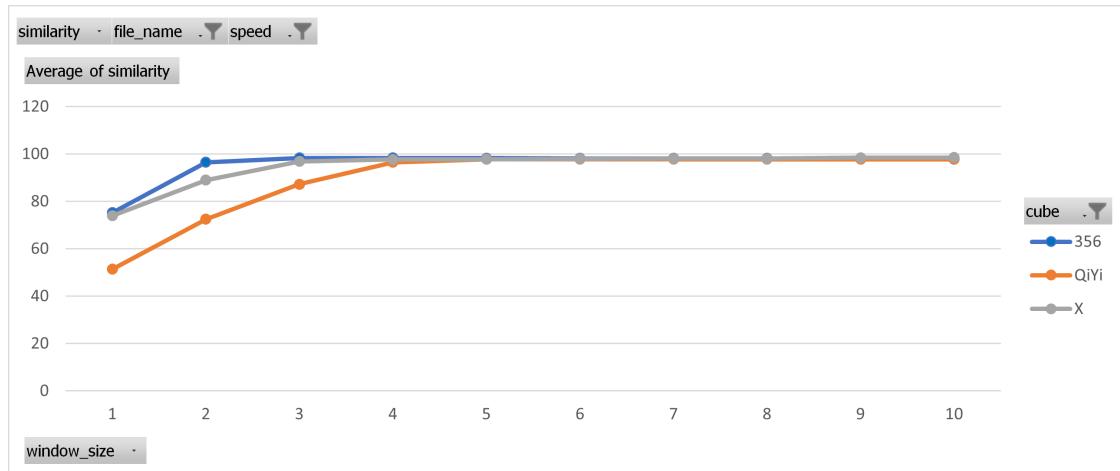


FIGURE 7.4: Perfect Detections by Window Size

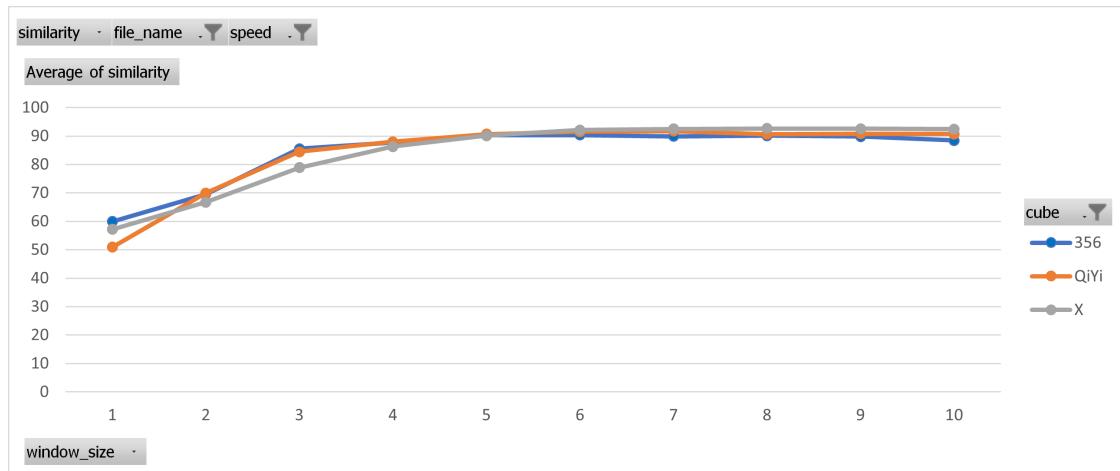
(A) Total number of perfect detections



(B) Average Accuracy by Window Size at 2 TPS



(C) Average Accuracy by Window Size at 5 TPS



7.2 Compatibility with Standard Speedcubes

Since a major goal of this research project is to allow speedcubers to practice with the same cube they compete with, it is imperative that the proposed design function without requiring permanent modifications to the original speedcube. For the same reason, the design must also retain the original performance of the speedcube as much as possible.

As shown in Figure 6.8, the proposed PCB transmitter can be deployed in a Gans 356 speedcube by simply changing out the original centercap for a custom-built one containing the transmitter. The modifications to the cube are purely temporary since the original cube can be perfectly restored by simply replacing the original centercaps.

The performance impact of the custom centercaps with the transmitters is also small since the little weight they add to the cube is placed close to the axes of rotation. On the Gans 356, a stock centercap has a mass of 0.6719 grams, while the custom centercap shown in Figure 6.6 has a mass of 1.6905 grams. Multiplying the approximately 1.2 gram difference per centercap by the six centercaps on the cube results in a 7.2 gram increase in mass for the entire speedcube for a total mass of 91.3104 grams. Since the stock Gans 356 has a mass of 84.1204 grams, this is only an 8.6% increase in mass, which is within the range of variation between different speedcubes.² This additional weight does increase the moment of interia of each face; however, this increase is small because the added weight placed close to the axes of rotation. As such, the corresponding performance penalty of the extra weight would be small.

7.3 Move Tracking Granularity

While tracking the moves performed on the speedcube constitutes the minimum viable functionality for a design, additional features make a design a more useful solution to the speedcubing community. Of most value is the ability to track the amount of time spent performing each face turn; from that single metric many other highly valuable metrics can be derived like TPS over time, and time spent completing each stage of the solution.

Though not shown in the pretty printed algorithm from Section 5.5.3, the proposed software receiver also records the time stamp at which it detected each face turn. By computing the difference between time stamps, one can calculate the amount of time spent performing each individual face turn. While not implemented in this design, a more sophisticated software receiver could provide even more precise measurements from the same source data by measuring the amount of time between stable signals for each centerpiece state.

However, since this design does not include a gyroscope, it cannot provide metrics about cube orientation over time like the highest-end commercial smartcubes. As such,

²For example, the acclaimed MoYu WeiLong GTS3 M has a mass of 92g [55]. More information about the variation in masses of different speedcubes can be found at TheCubicle.com.

this design cannot be used to directly report x, y, and z cube rotations, though sophisticated post-processing heuristics could predict the location of such rotations with decent accuracy.

7.4 Competition Legality

As discussed in Section 2.3.2, speedsolving competitors cannot use any electronics or audio equipment while solving the cube in competition. Since a major goal of this research project is to allow speedcubers to practice with the same cube they compete with, a design must either inherently comply with existing competition rules against the use of electronics or must result in a cube that can be easily modified to regain compliance.

While the design proposed in this thesis, certainly does not inherently comply with the existing competition regulations since it requires embedding electronics into the cube, it does result in a cube that can be easily modified to regain compliance. As mentioned in Section 7.2, the only modification to the cube is an interchangeable centercap that can be easily removed and replaced with the original centercap to produce a perfectly competition-legal cube.

Interestingly, the sound-based nature of this design provides a potential compromise on the use of smartcubes in competition. As mentioned at the end of Section 6.5, a sound-based transmitter only works at a short range and can easily be rendered inaccessible to a competitor by scrambling cubes several meters away from the competitors (especially in a different room), or playing audio at the same frequencies as the transmitter. As such, these properties of sound-based turn tracking could provide an avenue in which smartcubes could be safely allowed in competitions to aid in solve reconstructions without enabling competitors to spy on scrambles.

7.5 Summary

In summary, this chapter explored the effectiveness of the design proposed in Chapters 4, 5, and 6 against four benchmarks derived from the research questions outlined in Section 3.5.

The analysis of the receiver algorithm revealed many configurations that enabled it to achieve a perfect detection accuracy on the test samples. Most impactful in that detection was the usage of a high `window_size` which achieved accuracies of over 90% across all variations in the other parameters. Additionally beneficial is the receiver's ability to detect the time between face turns which opens the door to calculating useful metrics like TPS over time.

The analysis of the transmitter also found promising results. A rough prototype of the PCB was small enough to fit into a single centercap without requiring any permanent modifications to the cube. Furthermore, since the centercap was interchangeable, the electronic components impermissible in competition could easily be removed to return the cube to a competition-legal state. In addition, the added mass of the electrical components did not significantly increase the weight of the puzzle, nearly retaining its original performance.

Chapter 8

Conclusion

The D.I.Y. Smartcube seeks to enable speedcubers to obtain the same metrics and data from their main speedcube that are offered by existing commercial smartcubes. This chapter summarizes the progress made toward that goal (Section 8.1), the hurdles that remain (Section 8.2), and the next milestones that await future research (Section 8.3).

8.1 Answering Research Questions

Chapter 3 concluded with a detailed list of research questions that would be explored in this thesis (see Section 3.5). This section provides direct, summarized answers to each of those questions.

Section 8.1.1 summarizes the constraints for a sound-based move tracking protocol. Section 8.1.2 then reports on the accuracy achieved by the proof-of-concept proposed earlier in the thesis. Section 8.1.3 continues by reviewing the challenges associated with building a transmitter small enough to non-destructively embed into an existing speedcube. From there, Section 8.1.4 discusses additional data points that can be inferred from the transmitted audio. Finally, Section 8.1.5 addresses the ways in which a sound-based smartcube design can comply with competition regulations banning the usage of electronics.

8.1.1 Feasibility on Consumer Hardware

What are the constraints for a sound-based smartcube design compatible with consumer-grade microphones like those found in common smartphones and laptops?

As discussed in Chapter 4, there are four main factors that affect the viability of a sound-based smartcube design:

1. the achievable **signal-to-noise ratio** of the transmitter
2. the **distinctiveness of the tones** used for the protocol
3. the **range of tones audible** to the human ear
4. the **frequency response range of consumer hardware** used to implement the protocol

An analysis of these factors found that the tones used in the protocol should be (1) broadcast with an audio intensity of -30dB or greater (see Section 4.2.1) and (2) spaced at least 100Hz apart (see Section 4.2.2) in order to have a sufficiently strong signal-to-noise ratio for clear detection. Furthermore, the specific tones chosen for the protocol would ideally fit within the "musical range" of 500Hz and 4000Hz to simultaneously (3) avoid irritating human solvers (see Section 4.2.4) and (4) leverage the optimizations built into consumer audio recording hardware in smartphones and laptops (see Section 4.2.3).

8.1.2 Move Tracking Accuracy

How could a sound-based smartcube design track the face turns of a Rubik's Cube with high accuracy?

As discussed in Section 7.1, the receiver from Chapter 5 and the transmitter from Chapter 6 were able to transmit and decode a sequence of face turns with perfect accuracy across a variety of configuration parameters within a controlled environment.¹

Various factors aided in these results. First, each face position is associated with a signal present in a unique frequency band (See Section 4.2.2). This separation of frequencies makes it easier to distinguish changes in the state of independent faces on the cube. Second, the current position of each face is continuously broadcast, increasing the amount of time available for the receiver to detect and decode each position over time (See Section 4.4). Third, multiple layers of noise filters were applied within the receiver to better focus on the true signal frequencies and boost confidence in the accuracy of the decoded face turns (See Section 5.5).

8.1.3 Compatibility with Standard Speedcubes

How could a sound-based smartcube design be deployed within a standard, "non-smart" speedcube without requiring permanent modifications to the original cube?

The key challenges involved in creating a sound-based smartcube enhancement for an existing speedcube stem from the size constraints imposed by the structure of the cube. Since all modifications must be non-destructive, the required electronics must be embedded into the empty spaces within the cubies on the cube. Furthermore, since neither the edges nor the corners on the cube have access to the axles on which the cube rotates, the best place for these electronics is within the centercaps themselves. Unfortunately, these centercaps offer little space for add-ins. For example, as discussed in Sections 6.2.4 and 6.6, the relatively spacious Gans 356 only offers space for a $16mm^2$ PCB with rounded corners and a $6mm$ diameter hole in the middle.

¹Specifically, this controlled environment consisted of synthetically generated audio representative of the output of an ideal transmitter (see Sections 5.2 and 5.4) played from a Google Pixel phone and recorded by an HP Spectre laptop in a quiet room (see Section 4.2.1).

Despite these stringent size constraints, Chapter 6 specifies a promising transmitter design consisting of less than 10 discrete electrical components. With this limited component count, a PCB consisting of SMD components would be able to fit within a custom-printed centercap for a Gans 356 speedcube as shown in Figure 6.8.²

As discussed in Section 7.2, since the required electronics are contained within a replaceable, custom centercap, the modifications can be easily reversed by replacing the custom centercaps with the originals.

8.1.4 Move Tracking Granularity

How could a sound-based smartcube design record the time spent executing each individual face turn of a Rubik's Cube?

The receiver proposed in Chapter 5 decodes face turns by iterating over a time series of audio data. As a result, when a face turn is detected, the receiver also knows the time at which it occurred. By computing the difference between the times at which two successive turns occurred, the time spent executing an individual face turn can be computed. Thus, as long as the receiver can detect the face turns applied to the cube with high accuracy, then the time spent executing each one can be easily calculated (See also Section 7.3).

8.1.5 Competition Legality

How could a sound-based smartcube design comply with competition regulations prohibiting the use of electronics while performing a competitive solve?

Since the use of electronics while competitively solving a Rubik's Cube is banned by WCA regulation 2i [7], any move tracking solutions must not permanently embed electronics into the structure of the cube.

The solution proposed in Section 6.2.4 is to create a tiny sound transmitter that can be embedded into a custom centercap. Such a transmitter can be temporarily installed during personal practice then replaced with the original centercaps during competitions.

Alternatively, since the signal of a sound-based transmitter can be blocked by a physical wall or purposefully broadcasting interfering signals in the same frequency range, it's possible that, with proper negotiation with the WCA, a sound based move tracking solution could be deemed competition legal.

8.2 Limitations

As stated in Section 3.5, the overarching goal of this thesis is to present a proof-of-concept for a sound-based smart cube design. Like all proofs-of-concept, the design proposed in this thesis contains some inherent flaws and underlying assumptions that limit its ability to be successfully deployed in existing speedcubes.

²Note that an actual PCB was not constructed nor was a PCB layout attempted for any other cube than the Gans 356.

First, while Chapter 6 demonstrates that its proposed transmitter design appropriately considers all relevant requirements, a fully functional transmitter was never built and deployed within an existing speedcube. As such, the viability of this transmitter design cannot be fully verified until a transmitter is fully constructed, deployed, and tested.

Second, while great care was taken to extensively test the receiver, all tests were conducted with realistic, but synthetic audio samples since a physical transmitter was never built (See Section 5.4). As such, it is possible that further refinements to the receiver will be required if the signals from a real transmitter substantially differ from those synthesized for testing the receiver.

Third, though not a functional limitation, since the typical frequency response range of consumer hardware is entirely within the human auditory range (see Sections 4.2.3 and 4.2.4), all frequencies selected for transmitting the state of the cube will be human-audible. If the chosen frequencies are higher in pitch (as recommended by Section 6.3.3), the sounds produced by the transmitter may prove irritating to the speedcuber.

Fourth and finally, while the tests of the receiver achieved perfect accuracy on all test cases (see Section 8.1.2), the specific parameters used to achieve perfect accuracy differed between each one. Since there is not one set of parameters that worked for all test cases, additional experimentation and refinement will be necessary to determine the best parameters for decoding face turns in any given cube/environment combination.

8.3 Outlook

The proof-of-concept shown in this thesis demonstrates the potential of a sound-based move tracking system for realizing the goal of the D.I.Y. Smartcube: to enable speedcubers to obtain the same metrics and data from their main speedcube that are offered by existing commercial smartcubes.

To address the specific issues discussed in the Limitations section above (8.2), the following areas could be studied more thoroughly:

1. **Transmitter Construction:** Designing, manufacturing, and testing the actual PCB that would be embedded into a Rubik's Cube's centercap. This may include creating variations of the PCB for different types of speedcubes.
2. **Pairing the Receiver with a Physical Transmitter:** Once a physical transmitter is constructed, re-run the battery of tests described in Section 7.1 to assess the receiver's actual accuracy with a physical transmitter.
3. **Supersonic Frequencies:** Explore the viability of moving the frequency bands used for face turn transmission into the supersonic range ($>20\text{kHz}$) so that they cannot distract human speedsolvers. At a minimum, this would require connecting a supersonic microphone into a laptop or smartphone to detect the higher frequency tones.
4. **Refined Receiver Parameters:** Create a collection of optimal receiver parameters for a variety of combinations of cubes and environments. Additional research could also explore ways to determine these optimal parameters dynamically as a speedcuber gives feedback on the quality of the face turn decoding.

There are also a few areas other of potential research that would enhance the experience of using a sound-based smartcube.

- **Live Algorithm:** The current implementation of the receiver records the entire move sequence, then processes it all at once. A more interactive version would be able to continuously listen to the microphone and detect face turns in real time.
- **Support for Cube Rotations:** As mentioned in Section 5.2.1, the set of decoded face turns assumes no cube rotations. Cube rotations are extremely common during speedsolves and pose interesting challenges for move tracking systems since they do not rotate any axles but do change which color centerpiece is associated with each face (see Section 2.2). Potential solutions could involve the addition of a hardware gyroscope and/or a set of software heuristics based on speedcubers' preferences for R U L D moves.

Additional support for future research in this area can also be found by contacting the author and/or participating in speedsolving communities such as Speedsolving.com.

Appendix A

Longer Code Snippets

A.1 Spectrogram Generation for Figure 5.5

This code snippet actually generates a full animation of how the component frequencies change over time. Figure 5.5 is only one frame of this animation.

Make sure to run this code after running the code in Section 5.3.1.

```
# Heavily aided by https://matplotlib.org/stable/gallery/animation/simple_anim.html
import matplotlib.pyplot as plt
import matplotlib.animation as animation

def compute_threshold(values: list):
    return -500 # Out of sight of the chart

def animate(wav_path: str, alg: str):
    # Parse the audio sample into its spectrogram
    freq, time, spectrogram = compute_spectrogram(wav_path)

    # Generate the plots
    fig, (ax_spectrogram, ax_fourier) = plt.subplots(2)

    # Set the spectrogram
    ax_spectrogram.specgram(audio_samples, Fs=sample_rate,
                            NFFT=SAMPLES_PER_WINDOW,
                            nooverlap=SAMPLES_PER_WINDOW // 4 * 3) # 3/4 of the window size
    line_spectrogram, = ax_spectrogram.plot([], [], '-',
                                             color='red')
    time_tracker_spectrogram = ax_spectrogram.axvline(0, color='red')
    ax_spectrogram.set_xlim(0, 4000)
    ax_spectrogram.set_title(f"Spectrogram")
    ax_spectrogram.set_ylabel('Frequency [Hz]')
    ax_spectrogram.set_xlabel('Time [sec]')

    # Set the Fourier transform
    STEP = 10
    indices = np.arange(0, len(time) + 1, step=STEP)
    line_fourier, = ax_fourier.plot(freq, spectrogram[0], color='red')
```

```

threshold_fourier = ax_fourier.axhline(-500, color='green')
ax_fourier.set_title("")
ax_fourier.set_xlim(0, 4000) # Frequency
ax_fourier.set_xlabel("Frequency [Hz]")
ax_fourier.set_ylim(0, npamax(spectrogram)) # Strength
ax_fourier.set_ylabel("Strength")

# Space figures nicely
fig.suptitle(alg)
fig.tight_layout(h_pad=3)

def draw_frame_at(i):
    time_tracker_spectrogram.set_xdata(time[i])
    line_fourier.set_ydata(spectrogram[i])
    threshold_fourier.set_ydata(compute_threshold(spectrogram[i]))
    ax_fourier.set_title(f" Frequencies at {time[i]:.2f} seconds [idx={i}]")
    return line_spectrogram, line_fourier

ani = animation.FuncAnimation(fig, draw_frame_at,
    frames=indices, interval=1, blit=True, save_count=len(time)/STEP)
ani.save(f"plt/animation/{wav_path[:-4]}.mp4",
    fps=len(time)/STEP/5, dpi=300)
plt.show()

animate(demo_wav_path, demo_alg)

```

A.2 Spectrogram Generation for Figure 5.7

This code snippet actually generates a full animation of how the component frequencies change over time. Figure 5.7 is only one frame of this animation.

Make sure to run this code after running the code in Section 5.3.2 as the new *compute_threshold* function there is what adds the threshold line to this animation.

```
animate(demo_wav_path, demo_alg)
```

Bibliography

- [1] The World Cube Association. *Competitions*. URL: <https://www.worldcubeassociation.org/competitions>. (accessed: 21 Feb 2022).
- [2] The Cubicle. YUXIN LITTLE MAGIC 3X3. URL: <https://www.thecubicle.com/collections/3x3-speed-cubes/products/yuxin-little-magic-3x3>. (accessed: 21 Feb 2022).
- [3] The Cubicle. XiaoMi Giiker Cube. URL: <https://www.thecubicle.com/products/xiaomi-giiker-cube/>. (accessed: 25 Sep 2021).
- [4] The Cubicle. MAX GAN 356 XS 3X3. URL: <https://www.thecubicle.com/products/max-gan356-xs>. (accessed: 21 Feb 2022).
- [5] GoCube. GoCube Edge Full Pack. URL: <https://getgocube.com/products/gocube-full-pack/>. (accessed: 21 Feb 2022).
- [6] Speedsolving.com. *Should WCA replace Stackmat timers with smart cubes?* URL: <https://www.speedsolving.com/threads/should-wca-replace-stackmat-timers-with-smart-cubes.79133/>. (accessed: 21 Feb 2022).
- [7] The World Cube Association. *WCA Regulations*. May 2021. URL: <https://www.worldcubeassociation.org/regulations/#2i>. (accessed: 14 Oct 2021).
- [8] Rubik's Brand Ltd. *About Us*. URL: <https://www.rubiks.com/en-us/about>. (accessed: 10 August 2021).
- [9] Erno Rubik. *Rubik's Cube Patent*. URL: <https://www.hipo.gov.hu/hu/anim/pics/HU-170062.pdf>. (accessed: 10 August 2021).
- [10] Joan Verdon. *Rubik's Cube And Spin Master: A \$50 Million Deal With Endless Possibilities*. URL: <https://www.forbes.com/sites/joanverdon/2020/11/15/rubiks-cube-and-spin-master-a-50-million-deal-with-endless-possibilities/>. (accessed: 10 August 2021).
- [11] Rubik's Brand Ltd. *The History of the Rubik's Cube*. URL: <https://web.archive.org/web/20180908211659/https://www.rubiks.com/about/the-history-of-the-rubiks-cube>. (accessed: 10 August 2021).
- [12] Stewart Smith. *I Am The Cube*. URL: <https://iamthecu.be/>. (accessed: 02 Mar 2022).
- [13] David Singmaster. "Moral and mathematical lessons from a Rubik cube". In: *New Scientist* 96.1337-1338 (1982).
- [14] Tadeáš Miller. *Rubikova kostka 3x3x3*. URL: <https://hlavolam.maweb.eu/rubikova-kostka-3x3x3>. (accessed: 26 Feb 2022).
- [15] The World Cube Association. *History of the Speedcubing Community & WCA*. URL: <https://www.worldcubeassociation.org/speedcubing-history>. (accessed: 26 Feb 2022).
- [16] Go Cube. *Go Cube - Official Kickstarter Video*. URL: https://www.youtube.com/watch?v=tMtmzoC_WUY. (accessed: 26 Sep 2021).

- [17] The Cubicle. *Gan 356 I*. URL: <https://www.thecubicle.com/products/gan-356i>. (accessed: 26 Sep 2021).
- [18] Go Cube. *Rubik's Connected*. URL: <https://www.getgocube.com/products/rubiks-connected/>. (accessed: 27 Sep 2021).
- [19] Federal Communications Commission. FCC ID: 2ATHZ-SUPERCUBE - INT PHO. URL: https://apps.fcc.gov/oetcf/eas/reports/ViewExhibitReport.cfm?mode=Exhibits&calledFromFrame=Y&application_id=xbxQDjxzoHU1WywTbpvIAQ%3D%3D&fcc_id=2ATHZ-SUPERCUBE. (accessed: 24 Sep 2021).
- [20] Charlie Eggins and Geoff Eggins. *Giiker Cube Repair for Repeated Out of Sync Issues*. URL: <https://swiftcubing.com/2019/10/14/giiker-cube-repair-for-repeated-out-of-sync-issues/>. (accessed: 24 Sep 2021).
- [21] The GoCube Team. *Shipping Update*. URL: <https://www.kickstarter.com/projects/1928372437/gocube-the-classic-puzzle-reinvented/posts/2525348>. (accessed: 09 Mar 2022).
- [22] Federal Communications Commission. FCC ID: 2ASMEGC33 - Internal photos. URL: https://apps.fcc.gov/oetcf/eas/reports/ViewExhibitReport.cfm?mode=Exhibits&calledFromFrame=Y&application_id=tkd9t6CDQWEFnu4P%2B1WEbg%3D%3D&fcc_id=2ASMEGC33. (accessed: 24 Sep 2021).
- [23] Federal Communications Commission. FCC ID: 2AT27-GAN-3X3 - Internal photos. URL: https://apps.fcc.gov/oetcf/eas/reports/ViewExhibitReport.cfm?mode=Exhibits&calledFromFrame=Y&application_id=B75M7i7IVYNksPluyUFB1Q%3D%3D&fcc_id=2AT27-GAN-3X3. (accessed: 24 Sep 2021).
- [24] IBM. *What is computer vision?* URL: <https://www.ibm.com/topics/computer-vision>. (accessed: 23 Sep 2021).
- [25] Jay Hack and Kevin Shutzberg. *Rubik's Cube Localization, Face Detection, and Interactive Solving*.
- [26] OpenAI et al. "Solving Rubik's Cube with a Robot Hand". In: *arXiv preprint* (2019).
- [27] Junshen Kevin Chen, Wanze Xie, and Zhouheng Sun. "DeepCube: Transcribing Rubik's Cube Moves with Action Recognition". In: () .
- [28] Jiawei Huang et al. "IM3D: Magnetic Motion Tracking System for Dexterous 3D Interactions". In: *ACM SIGGRAPH 2014 Emerging Technologies*. SIGGRAPH '14. Vancouver, Canada: Association for Computing Machinery, 2014. ISBN: 9781450329613. DOI: 10.1145/2614066.2614084. URL: <https://doi-org.ezproxy1.lib.asu.edu/10.1145/2614066.2614084>.
- [29] Maria Mannone et al. "CubeHarmonic: A New Interface from a Magnetic 3D Motion Tracking System to Music Performance". In: *NIME Conference Proceedings*. 2018.
- [30] Maria Mannone et al. "CubeHarmonic: a new musical instrument based on Rubik's cube with embedded motion sensor". In: *ACM SIGGRAPH 2019 Posters* (2019).
- [31] Richard Polfreman and Benjamin Oliver. "Rubik's Cube: Music's Cube". In: *NIME Conference Proceedings*. Apr. 2017.
- [32] Jonas Michel. *mobile-acoustic-modems-in-action Wiki*. URL: <https://github.com/jonasrmichel/mobile-acoustic-modems-in-action/wiki/Wiki>. (accessed: 28 Sep 2021).
- [33] Ari Neto. *Twisty Timer*. URL: <https://play.google.com/store/apps/details?id=com.aricneto.twistytimer>. (accessed: 28 Sep 2021).

- [34] SpeedSolving.com. *Home Page*. URL: <https://www.speedsolving.com/>. (accessed: 28 Sep 2021).
- [35] Simone Genovesi et al. "Chipless Radio Frequency Identification (RFID) Sensor for Angular Rotation Monitoring". In: *Technologies* 6.3 (2018). ISSN: 2227-7080. DOI: 10.3390/technologies6030061. URL: <https://www.mdpi.com/2227-7080/6/3/61>.
- [36] The Food and Drug Administration. *Radio Frequency Identification (RFID)*. URL: <https://www.fda.gov/radiation-emitting-products/electromagnetic-compatibility-emc/radio-frequency-identification-rfid>. (accessed: 28 Sep 2021).
- [37] NVE Corporation. *Off-Axis Angle Sensing with NVE Angle Sensors*. URL: https://www.nve.com/Downloads/SB-SA-02_Off-Axis-Angle-Sensing.pdf. (accessed: 06 June 2020).
- [38] NVE Corporation. *Calculators/WebApps: Off-Axis Angle Sensing*. URL: <https://www.nve.com/spec/calculators#tabs-Off-Axis-Angle-Sensing>. (accessed: 06 June 2020).
- [39] Carl Reinke. *Spectroid*. URL: <https://play.google.com/store/apps/details?id=org.intoorbit.spectrum>. (accessed: 19 Oct 2021).
- [40] Hristo Bojinov et al. "Mobile Device Identification via Sensor Fingerprinting". In: (Aug. 2014).
- [41] Jeffrey Hass. 5. *What is Frequency | Page 2*. 2020. URL: https://cmtext.indiana.edu/acoustics/chapter1_frequency2.php. (accessed: 09 Oct 2021).
- [42] Carl Rod Nave. *The Piano*. 1999. URL: <http://hyperphysics.phy-astr.gsu.edu/hbase/Music/pianof.html>. (accessed: 09 Oct 2021).
- [43] J Perm. *Rubik's Cube Move Notation*. 2021. URL: <https://jperm.net/3x3/moves>. (accessed: 10 Oct 2021).
- [44] Erik Nyquist. *tones* 1.2.0. 2020. URL: <https://pypi.org/project/tones/>. (accessed: 29 Nov 2020).
- [45] J. D. Hunter. "Matplotlib: A 2D graphics environment". In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.
- [46] Steve Brunton. *Fourier Analysis*. Sept. 2021. URL: https://www.youtube.com/playlist?list=PLMrJAhkIeNNT_Xh30y0Y4LTj00xo8GqscC. (accessed: 20 Oct 2021).
- [47] Charles R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: <https://doi.org/10.1038/s41586-020-2649-2>.
- [48] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.
- [49] General Purpose Timers. ICM7555. Rev. 10.01. Renesas. 2020. URL: <https://www.mouser.com/datasheet/2/698/icm7555-56-1528913.pdf>.
- [50] Jeffrey Stolet. 11. *Harmonics*. 2009. URL: <https://pages.uoregon.edu/emi/14.php>. (accessed: 13 Nov 2021).
- [51] Jeffrey Stolet. 14. *Square and Rectangle Waves*. 2009. URL: <https://pages.uoregon.edu/emi/14.php>. (accessed: 13 Nov 2021).
- [52] DYWISHKEY. *DYWISHKEY 15Value 540PCS DIP Monolithic Multilayer Ceramic Chip Capacitors Assortment Kit, 10pF to 100nF Ceramic Capacitor Set*. 2021. URL: <https://www.amazon.com/gp/product/B07PSJ6V98/>. (accessed: 30 May 2021).

- [53] BOJACK. *BOJACK 50 Values 1350 Pcs Resistor Kit 0 Ohm-5.6M Ohm with 1% 1/4W Metal Film Resistors Assortment*. 2021. URL: <https://www.amazon.com/gp/product/B07P3MFG5D/>. (accessed: 30 May 2021).
- [54] Stefan Pochmann. "Analyzing Human Solving Methods for Rubik's Cube and similar Puzzles". MA thesis. Technische Universität Darmstadt, 2008.
- [55] The Cubicle. *MOYU WEILONG GTS3 M 3X3*. URL: <https://www.thecubicle.com/collections/3x3-speed-cubes/products/moyu-weilong-gts3-m>. (accessed: 02 Feb 2022).