SONIFYING SENTIMENT IN A REAL-TIME TWITTER STREAM

Purusottam Samal

Columbia College Chicago
600 S Michigan Avenue
Chicago, USA
psamal@colum.edu

David Worrall

Columbia College Chicago
600 S Michigan Avenue
Chicago, USA
dworrall@colum.edu

ABSTRACT

This paper describes the beginning of a research project involving sonification of the sentiment of a real-time stream of Twitter messages associated with a Twitter hashtag into a soundscape that affords a listener the ability to perceive the instantaneous sentiment, the average sentiment and the weighted impact of the stream. The goal of the sonification is to serve as a background feature that allows a subject to listen passively and still be in touch with the sentiment associated with a topic of interest. This paper will discuss the techniques used to create the sonification and the motivations behind them, including details of the sentiment analysis, mapping strategies, visual display and sound output.

1. INTRODUCTION

Sentiment analysis is the process of determining the emotional tone behind a series of words and is used to gain an understanding of the attitudes, opinions and emotions expressed by the words. This is important because emotions and attitudes towards a topic can become actionable pieces of information useful in numerous areas of business and research [1]. One such use-case is social media monitoring as it reveals an overview of the wider public opinion behind a certain topic. Sonification can be a useful tool to facilitate this process of monitoring [2].

The sonifications described here contain two sub-systems: (1) data collection, text preprocessing [1], and sentiment analysis system created in Python, and (2) the sound and visualization system created in Max/MSP [3]. The paper details the processes carried out by both the sub-systems and how they work together to form one cohesive unit.

2. DATA COLLECTION

The collection of data is written in Python. The Python library, Tweepy [4], is used to access the Twitter API [5] and receives the data for the sonification system. The data is updated every four seconds, with each update being acquired by the system in real time. The data is also filtered based on a specific hashtag [6] to ensure the tweets being received all address the same topic. The Twitter API encodes the data using JavaScript Object Notation (JSON) [7]. JSON is based on key-value pairs, with named attributes and associated values. These attributes, and their state are used to describe objects. The attributes accessed for the sonification system are {text} and {extended_text} from the Tweet object [8], which contains the actual UTF-8 text of the twitter message. Additionally, the {followers_count} attribute is obtained from the User object [9] due to it's correlation with a user's reach.

3. DATA AND SENTIMENT ANALYSIS

Text preprocessing is traditionally an important step for natural language processing (NLP)[2] tasks. It transforms text into a more digestible form so that machine learning algorithms can perform more effectively. Figure 1 demonstrates the series of preprocessing done on the raw text received from the {text} or {extended_text} attribute of a Tweet object.

A supervised machine learning model based on a Multinomial Naive Bayes classifier [3] was used to classify the text. The Twitter US Airline Sentiment data set [10] was used to provide training inputs, with each training input associated with its correct output. To test the effectiveness of a machine learning system, a data set is normally split into a training data set and a testing data set based on a 80/20 split. The model uses the training data set to analyze the relationship between the given inputs and outputs and finally predict with some accuracy the output; given a new input. The inputs are preprocessed as shown in Figure 1.

A model is generated to fit the training data using in Multinomial Naive Bayes and the quality of the model is evaluated in the testing data set. The accuracy of the model was measured at 60.25%, which is low and needed improvement.

4. SONIFICATION METHODS

This section discuss the sonification methods used and explain why and how the tweet features map to the parameters of the sound events.

The sentiment output and the user follower count of each tweet was coded in Open Sound Control(OSC) [11] messages and sent to the visualization and sonification modules in Max/MSP.

[1] Text preprocessing is a method to clean the text data and make it ready to feed data to a machine-learning model.

[2] NLP is a sub-field of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data.

[3] Classification is the process of predicting the class of given data points.

Raw text	RT @PuruSamal: 1234The dusk and dawn light in @DeathValleyNPS is amazing. I visit almost every year for #photography. VIDEO:https://www.youtube.com/watch?v=Kauv7MVPcsA #losangeles
Remove users	: 1234The dusk and dawn light in is amazing. I visit almost every year for #photography. VIDEO:https://www.youtube.com/watch?v=Kauv7MVPcsA #losangeles
Remove links	: 1234The dusk and dawn light in is amazing. I visit almost every year for #photography. VIDEO: #losangeles
Remove hashtags	: 1234The dusk and dawn light in is amazing. I visit almost every year for . VIDEO:
Remove AUDIO/VIDEO	: 1234The dusk and dawn light in is amazing. I visit almost every year for .
Lowercase	: 1234the dusk and dawn light in is amazing. i visit almost every year for .
Strip punctuation	1234the dusk and dawn light in is amazing i visit almost every year for
Remove double spacing	1234the dusk and dawn light in is amazing i visit almost every year for
Remove numbers	the dusk and dawn light in is amazing i visit almost every year for
Lemmatize and Tokenize	['dusk', 'dawn', 'light', 'amaze', 'visit', 'year']
List of words	'dusk dawn light amaze visit year'

Figure 1: Preprocessing applied to a Twitter message.

4.1. The Instantaneous Sentiment Stream

Sentiment values (valences) are received as real numbers in the range (-1,1), with valence 0 representing a neutral sentiment, negative valence representing negative sentiment, and positive valence representing positive sentiment. Four rectangular oscillators spread over a stereo space serve as the sound source for the instantaneous sentiment stream and are triggered each time a new twitter message is received. A pseudo-random number in the range 0-100ms is added to the trigger times to provide some affective listener variety [12].

A combination of harmonic content, vibrato, amplitude, event onset time and convolution is used to represent the instantaneous sentiment of a twitter message. The mapping to sonic features is as follows:

- **Harmonic Content:** This mapping is based on the ecological methodology described by Winters and Wanderley [13].The frequencies of each rectangular oscillator is specified by MIDI note numbers [14]. Positive sentiment is represented by a C Major chord (MIDI: 60 72 79 76). Negative sentiment is represented by a C Minor7 chord (MIDI: 60 75 67 70). As a result of this mapping, positive events sound brighter in both their harmonic and spectral content and negative events sound darker. Events with neutral sentiment use only one oscillator tuned to the root note C (MIDI: 60) making it easier to distinguish them from positive and negative events.
- **Vibrato:** The frequency of each rectangular oscillator is modulated by a sinusoidal oscillator causing slight deviations in the frequencies of the rectangular oscillators. The amount of deviation is mapped to the the absolute value of the sentiment, as follows:

$$|S(-1,\ldots,1)| \rightarrow f_m(4Hz,\ldots,8Hz), \quad (1)$$

where S is the sentiment value and f_m is the frequency of modulation. This mapping is subtle and present mainly to introduce sonic variety so as to not pall the listener's hearing.
- **Amplitude:** The amplitude value of each rectangular oscillator is mapped to the absolute value of the valence of a tweet. The amplitude of the first oscillator (MIDI:60) is fixed at maximum amplitude while the absolute value of sentiment $|S|$ affects the amplitude of oscillators 2-4. Amplitude of the second oscillator (MIDI:75 for S<0, MIDI:72 for S>0) reaches maximum amplitude at $|S| = 0.5$. Amplitude of the third oscillator (MIDI:67 for $S < 0$, MIDI:79 for $S > 0$) reaches maximum amplitude at $|S| = 0.75$. Amplitude of the fourth oscillator (MIDI:70 for $S < 0$, MIDI:76 for $S > 0$) reaches maximum amplitude at $|S| = 1$. Figure 4 depicts the amplitude mapping of the 4 rectangular oscillators with respect to $|S|$. This mapping makes the harmony 'richer' at higher $|S|$ values.

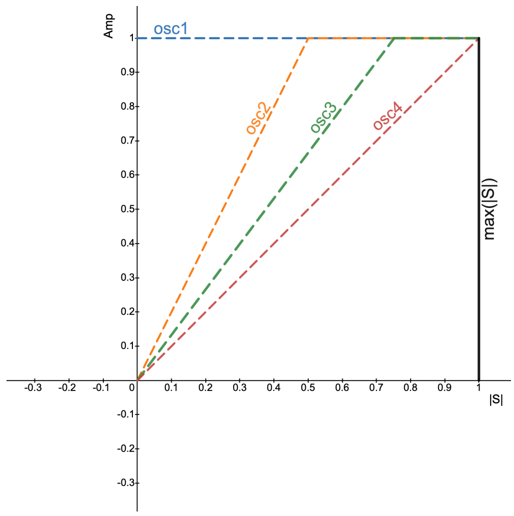
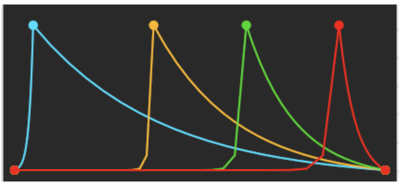
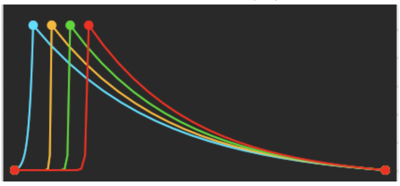


Figure 2: Amplitude vs $|S|$

- **Event onset time:** The oscillator outputs are further shaped by an envelope. The absolute value of sentiment $|S|$ is mapped to the attack times of each oscillator envelope and affects each oscillator differently. Attack times range from 50ms to 1000ms with the attack time of the first oscillator fixed at 50ms. Attack times are inversely proportional to valence [15]. Attack time of the second oscillator is 375ms at $|S| = 0$ and is 100ms at $|S| = 1$. Attack time of the third oscillator is 625ms at $|S| = 0$ and is 150ms at $|S| = 1$. Attack time of the fourth oscillator is 875ms at $|S| = 0$ and is 200ms at $|S| = 1$. This mapping is not linear and uses a cubic function and so that for lower $|S|$ values there is a clear separation between the oscillators and for higher $|S|$ values the oscillators are almost approaching unison. Figure 5 depicts attack times for each oscillator for $|S| = 0$ and $|S| = 1$ respectively.



(a) Attack times for $|S| = 0$



(b) Attack times for $|S| = 1$

Figure 3: Mapping of Event Onset Time

- **Cross Synthesis:** Cross-synthesis [16] is the technique of impressing the spectral envelope of one sound on the flattened spectrum of another. A Short-Time Fourier Transform (STFT) is performed on both the modulator and carrier signals. The spectral envelope of each time-frame is computed. The spectral frame is then multiplied by the envelope of the corresponding modulator frame, thereby replacing the carrier's envelope by the modulator's envelope. The four rectangular oscillators serve as the carrier. A speech 'Yay!' sound is the modulator in cases where the valence is positive. A speech 'No!' is the modulator in cases where the valence is negative. This mapping adds a very noticeable spectral imprint to the four rectangular oscillators further increasing the difference of perception between the positive and negative sentiments. Figure 6 illustrates this effect.

4.2. The Average Sentiment Stream

The average sentiment stream provides a context for the temporal evolution of the valence of the Twitter stream. It is based on the moving average of 8 valences. The moving average is updated every time a new twitter message is received. This value is mapped to a soundscape of 7 of wind chimes samples, each detuned slightly. The frequency of each sample is tuned as follows:

$$For\ S >= 0,$$
$$Midi\ Note : 36\ 40\ 43\ 48\ 52\ 55\ 60;$$
$$For\ S < 0,$$
$$Midi\ Note : 39\ 43\ 46\ 48\ 51\ 55\ 58;$$

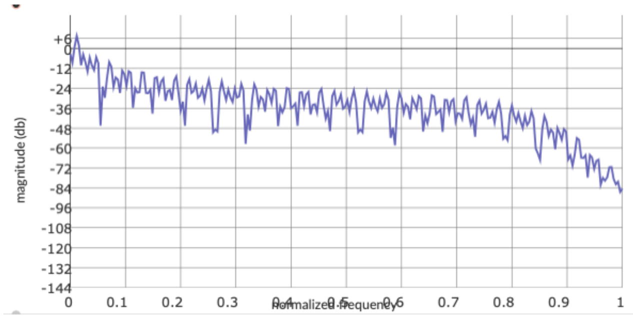
As a result of this mapping, the soundscape of wind chimes are tuned to a C Major chord for positive sentiment, and are tuned to a C Minor7 chord for negative sentiment.

Additionally, The octave at which the soundscape plays is mapped to moving average of valence as described in the Figure 7.

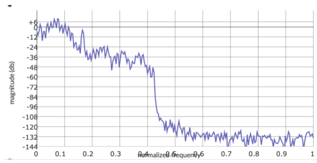
4.3. Impact of user-follower count

This sonification model accounts for the number of followers a Twitter user has as a measure of how much importance should be allocated to each event. The overall amplitude of the events of the instantaneous stream is mapped to the number of followers as described in the Figure 8(a).

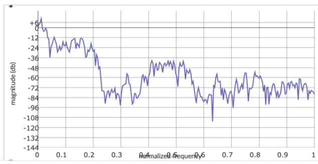
Additionally, the follower-count is mapped to another stream comprising of 7 slightly-detuned copies of keystrokes shaped by an envelope that fades in after the an event in the instantaneous sentiment stream dies out. This to preserve perception of certain events by compensating for the drop in amplitude of events in instantaneous stream where the follower count of a user associated with a twitter message is too low. The octave the soundscape is played back at is mapped to the number of followers as described in Figure 7(b).



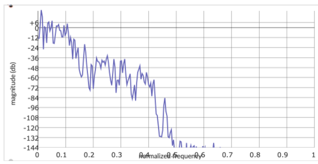
(a) Carrier.



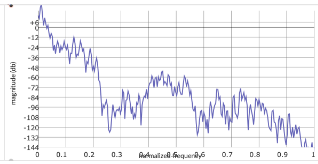
(b) 'Yay' Modulator ($|S| > 0$)



(c) 'No' Modulator ($|S| < 0$)



(d) Output, $|S| > 0$



(e) Output, $|S| < 0$

Figure 4: Cross-Synthesis

Sentiment Range	Octave Number
-1 <= S < -0.5	1(+12 semitones)
-0.5 <= S < 0	2(+24 semitones)
S = 0	0
0 < S <= 0.5	3(+36 semitones)
0.5 < S <= 1	4(+48 semitones)

Figure 5: Octave scaling for average sentiment stream

No. of followers	Amplitude Value
0 - 5,000	0.2
5,000 - 50,000	0.4
50,000 - 300,000	0.6
300,000 - 1,000,000	0.8
Greater than 1,000,000	1.0

(a) Mapping of instantaneous stream to number of followers.

No. of followers	Octave Number
0 - 5,000	0
5,000 - 50,000	1(+12 semitones)
50,000 - 300,000	2(+24 semitones)
300,000 - 1,000,000	3(+36 semitones)
Greater than 1,000,000	4(+48 semitones)

(b) Mapping of keystroke stream to number of followers.

Figure 6: Impact of a follower count

5. VISUALIZATION

Figure 9 shows a simple visualization interface that accompanies the sonification where one can see the hashtag being tracked, the contents of the original Twitter message, the number of followers of the user, the instantaneous sentiment, and the moving average sentiment.

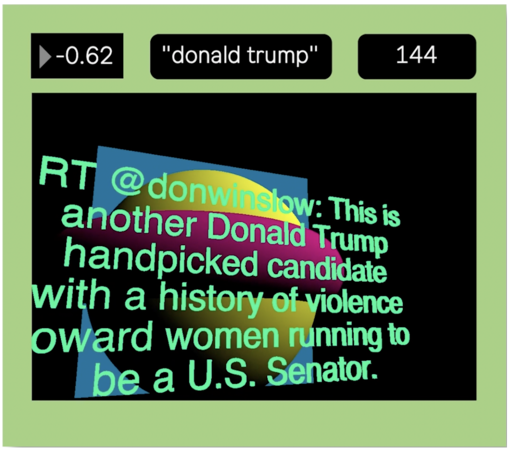


Figure 7: Simple visualization interface.

6. SUMMARY

The sentiment analysis process is based on inputting bag-of-words features (See Section 3.2) into a Multinomial Naive Bayes algorithm to classify the text as positive, negative or neutral. The accuracy of the system could be improved by using deep learning algorithms, however they require much larger datasets. Further, instead of using the bag-of-words approach which rely on frequencies of words under the unrealistic assumption that each word occurs independently of all others, use of word embeddings, which consider each word in its context, might produce better overall accuracy.

The visualization provides a minimal yet useful tool in monitoring the real-time stream of Twitter messages. However, there is room for further development of this aspect of the sonification. For instance, representing the features of the Twitter messages using colors, opacity and position in an X-Y space might be a useful addition.

Informal tests indicate that the sonification successfully allows the listener, with practice, to monitor multiple simultaneous features of the Twitter stream without requiring too much attention. The instantaneous sentiment stream provides a sequential sonification of the data which produces repeating patterns in which small variations are easy to detect, while the average sentiment stream allows the user to obtain a grasp on the bigger picture i.e. the moving average sentiment associated with the stream. But, while attack times and cross-synthesis serve as useful parameters to differentiate positive and negative valence, the mapping of sentiment to harmonic content is still the most prominent. This might cause some inter-section between the perception of a neutral sentiment and negative sentiment of lower valence values.

Overall, the sonification appears to achieve its aim of allowing users to determine the instantaneous sentiment, the average sentiment and the weighted impact of a real-time stream of Twitter messages in an aurally appealing way. Further experience with the system is required to fine-tune the sonification and visualisation algorithms, as well as to fully understand its possibilities and limitations.

The anonymity afforded by social media and limitations to which user interaction occurs (character limits) facilitates a form of human interaction that is more volatile than real life. Listening to hidden sequential patterns of sentiment in Twitter feeds might help us gain a better understanding of the user interaction on the platform regardless of the topic. It can help us potentially identify patterns in all forms of conversation and understand how the means through which people interact affect the flow of the discourse.

The demonstration of the sonification can be viewed here [17].

7. REFERENCES

[1] R. Feldman, "Techniques and applications for sentiment analysis," *Commun. ACM*, vol. 56, no. 4, p. 8289, apr 2013. [Online]. Available: https://doi.org/10.1145/2436256.2436274

[2] P. Vickers, "Sonification for process monitoring," in *The Sonification Handbook, T. Hermann, A. Hunt, and J. G. Neuhoff, Eds.* Berlin, Germany: Logos Publishing House, 2011, ch. 18, pp. 455–491.

[3] Cycling74, "Max/msp," https://cycling74.com/products/max, 2022, accessed: 2022-05-11.

[4] Tweepy, "Tweepy," https://www.tweepy.org, 2022, accessed: 2022-05-11.

[5] Twitter, "Twitter api," https://developer.twitter.com/en/docs/twitter-api, 2022, accessed: 2022-05-11.

[6] ——, "Twitter hashtag," https://help.twitter.com/en/using-twitter/how-to-use-hashtags, 2022, accessed: 2022-05-11.

[7] JSON, "Json," https://www.json.org/json-en.html, 2022, accessed: 2022-05-11.

[8] Twitter, "Tweet object," https://developer.twitter.com/en/docs/twitter-api/v1/data-dictionary/object-model/tweet, 2022, accessed: 2022-05-11.

[9] ——, "User object," https://developer.twitter.com/en/docs/twitter-api/data-dictionary/object-model/user, 2022, accessed: 2022-05-11.

[10] Kaggle, "Twitter airline sentiment," https://www.kaggle.com/datasets/crowdflower/twitter-airline-sentiment, 2022, accessed: 2022-05-11.

[11] M. Wright and A. Free, "Open sound control," https://opensoundcontrol.stanford.edu, 2022, accessed: 2022-05-11.

[12] A. L. Horner, A., "Contiguous-group wavetable synthesis of the french horn in csound," in *The Csound Book: Perspectives in Software Synthesis, Sound Design, Signal Processing, and Programming*, R. Boulanger, Ed. Cambridge, MA, USA: MIT Press, 2000, ch. 193, Figure 8.5, pp. 455–491.

[13] R. Winters and M. Wanderley, "Sonification of emotion: Strategies and results from the intersection with music," *Organised Sound*, vol. 19, pp. 60–69, 04 2014.

[14] D. Smith and C. Wood, "*MIDI*," https://www.midi.org, 2022, accessed: 2022-05-11.

[15] D. Worrall, *Sonification Design: From Data to Intelligible Soundfields*. Springer Publishing Company, Incorporated, 2019, ch. Parameter-Mapping Sonification of Tick-Data, pp. 240–241.

[16] J. O. Smith, *Spectral Audio Signal Processing*. W3K Publishing, 2011, ch. Cross-Synthesis.

[17] P. Samal, "Sonification sentiment in a real-time twitter stream," https://ufile.io/5hj28mp1, 2022, accessed: 2022-05-11.