Tonale Farbtöne
crossmodal perception of (quasi) synthetic sounds

Gabriel Feller, Christoph Reuter

SInES, Musikwissenschaftliches Institut, Universität Wien, 1090 Wien, E-Mail: gabriel.feller@univie.ac.at

Hintergrund: Crossmodal Correspondences-Forschung mit Sinus-Wellen

Crossmodal Correspondences sind intersubjektive, assoziationsgesteuerte Verbindungen zwischen zwei verschiedenen Sinneseindrücken. Auch Klang- und Sehwahrnehmungen werden in solchen Crossmodal Correspondences verknüpft. Beispielsweise werden Sinuswellen rundere Formen zugeordnet und Rechteckwellen stattdessen kantigere Formen [1]. Auch bezüglich der Tonhöhe zeigen sich einige Crossmodal Correspondences: So werden höhere Töne eher mit helleren als dunkleren Farben [2], spitzeren statt stumpferen Winkeln [1] sowie kleineren statt größeren Objekten und Formen [1][3][4] assoziiert. Lange zeigte sich jedoch kein klarer Zusammenhang zwischen Tonhöhe und Farbton (Color Hue) [2][5][6]. Stattdessen kann davon ausgegangen werden, dass der assoziierte Color Hue abhängig von der Klangfarbe des Musikinstruments ist [7].

In einer vorangegangenen Pilotstudie [10] sollte geprüft werden, wie grundlegende synthetische Klänge crossmodal wahrgenommen werden. 44 Versuchspersonen wurden in einer Online-Umfrage Sinus-, Dreiecks-, Rechteck- und Sägezahnschwingungen auf jeweils 5 Oktaven von C vorgespielt. Zunächst sollten diese Klänge in Adjektiv-Paare [8] eingeordnet werden. Danach mussten die Teilnehmenden drei Farben aus dem 38 Farben-Schema des Berkeley Color Projects [9] auswählen, die sie als am passendsten zu dem jeweiligen Klang bewerteten. Die so insgesamt 132 ausgewählten Farben pro Klang wurden gemittelt und die Farbwerte auf Korrelationen mit den Audio-Features der Klänge geprüft. Bestätigt werden konnte der positive Zusammenhang zwischen ansteigender Tonhöhe und ansteigender Farbhelligkeit. Zusätzlich fiel auf, dass der gemittelte assoziierte Farbton von Sinus- und Dreieckwellen mit steigender Tonhöhe von Blau nach Rot verlief, der Farbton von Rechteck- und Sägezahnwellen hingegen tonhöhenunabhängig beim Rotton blieb. Die teiltonarmen Klänge wurden auch insgesamt eher mit Blau und die teiltonreichen Klänge eher mit Rot assoziiert [10]. Die damalige Kritik der Teilnehmenden über eine zu geringe Farbauswahl, schlecht hörbare und schwer erträgliche Klänge und zu lange Umfragedauer wird in der folgenden in diesem Beitrag vorgestellten neuen Studie zu Farbassoziationen bei synthetischen Klängen einbezogen.

Forschungsfrage

Als Reaktion auf die Kritik der zu langen Umfragedauer wird die Zuordnung zu Adjektivpaaren ausgelassen und sich nur auf Crossmodal Correspondences zwischen Farben und synthetischen Klängen auf einem kleineren, angenehmer zu hörenden Tonraum konzentriert. Aus diesem Vorsatz ergibt sich die Forschungsfrage „Nach welchen Regeln werden Farben zu synthetischen Klängen assoziiert?“

Methodik

24 Personen (11 ♂, 12 ♀, 1 ☿, Alter: 19-79, x̃ 36) nahmen bereits an der neuen Online-Studie teil. 10 verschiedene synthetische Wellenformen wurden im Tritonusabstand im Tonbereich C3-Fis5 erstellt. Die Klänge bestanden wieder aus reinen Sinus-, Dreiecks-, Rechteck- und Sägezahnwellen sowie Mischformen aus Sinus- und den anderen drei Wellenformen. Zusätzlich wurden – aufgrund ihrer Teiltonarmut – sechs Klangbeispiele einer Querflöte eine Oktave höher als die Synthesizer-Klänge hinzugenommen. Die Klangbeispiele wurden in randomisierter Reihenfolge vorgespielt. Zu jedem Klangbeispiel sollten die Versuchspersonen in einem für diese Studie erstellten 60 Farben-Schema jene Farbe anklicken, die sie am ehesten mit dem jeweiligen Klang assoziieren.

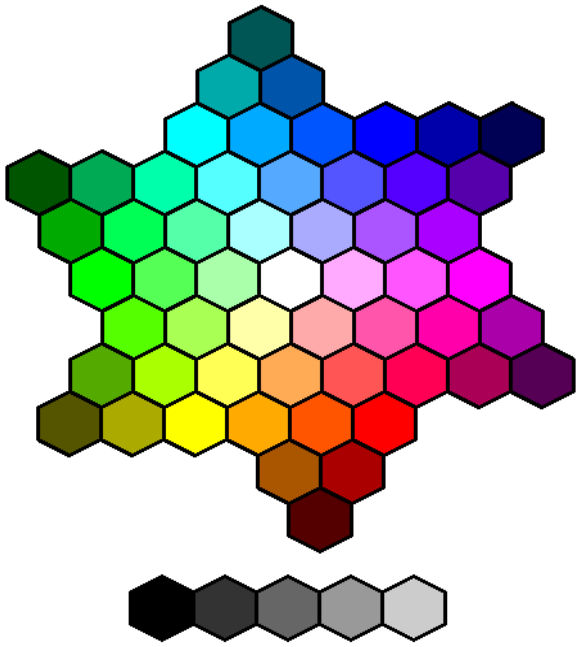

Abbildung 1: Das verwendete 60 Farben-Schema bildet die Farbtöne des HSV-Raums in einem 20° Color Hue-Abstand ab mit geringerer Sättigung zum Inneren der Figur und geringerem Farbwert nach außen hin sowie mit einer zusätzlichen Auswahl an Grau-Tönen.

Welche Farbe im Bild ausgewählt wurde, wurde in Form von Koordinaten gespeichert und die Farbwerte für die Auswertung anhand dieser aus dem Bild gezogen. Die ausgewählten Farben wurden für jedes Klangbeispiel anhand der RGB-Werte gemittelt. Zusätzlich wurden die 60 auswählbaren Farben in 13 übergeordnete Farbkategorien

eingeteilt, deren Auswahlhäufigkeit gezählt werden kann (Weiß, Hellblau, Hellgrün, Grau, Gelb, Orange, Rosa, Rot, Braun, Dunkelgrün, Dunkelblau, Violett, Schwarz).

Aus den insgesamt 66 Klängen wurden mit den Audiosignalanalyse-Libraries Essentia [11] und AudioCommons Timbral Models [12] grundlegende spektrale Klangeigenschaften ermittelt.

Anhand von Korrelationsanalysen und t-Tests sollten Zusammenhänge zwischen den Audio Features und den Farbwerten bzw. den Auswahlhäufigkeiten der Farbkategorien festgestellt werden.

Ergebnisse

Während die tabellarische Aufstellung der gemittelten Farben in der Pilotstudie bereits erkennbar den überraschenden Farbverlauf bei den Sinus- und Dreieckswellen darstellte [9], scheinen die gemittelten Farben der aktuellen Untersuchung durchmischter und auch innerhalb einer Wellenform ungleichmäßiger (s. Tabelle 1).

Tabelle 1: Darstellung der in den RGB-Werten gemittelten Farben je Klangbeispiel

	C 3	F# 3	C4	F#4	C 5	F#5	C6	F# 6
Rechteck								
Rechteck + Sinus								
Rechteck – Sinus								
Sägezahn								
Sägezahn + Sinus								
Sägezahn – Sinus								
Dreieck								
Dreieck + Sinus								
Dreieck – Sinus								
Sinus								
Querflöte								

Die Daten zeigen einen Abfall des Color Hues mit steigender Tonhöhe bei der Sinuswelle und „Dreieck + Sinus“.

Tonhöhe und Farbdimensionen

Zuerst wurden die Farbwerte und die Auswahlhäufigkeiten der Farbkategorien für jede der vier Wellenformkategorien (Sinus, Dreieck, Rechteck, Sägezahn; letztere +/- Sinus) auf Korrelationen mit der Tonhöhe in Form des MIDI Pitch analysiert:

Bei Dreiecks-, Rechteck- und Sägezahnwellen zeigt sich die bekannte Crossmodal Correspondence [2] zwischen Tonhöhe und Helligkeit der Farbe (ausgedrückt im Color Value). Aber auch ein höherer Rotwert und eine geringere Auswahl von Farben aus der Kategorie „Dunkelblau“ bei steigender Tonhöhe zeigt sich bei diesen drei Wellenformen. Bei allen vier Wellenformen gibt es allerdings ein positives Korrelat zwischen Tonhöhe und der Auswahl gelber Farben.

Zwar ließen sich auch Korrelationen zwischen Tonhöhe und Color Hue im HSV-Raum bei Sinus, Dreieck und sogar Rechteck nachweisen, allerdings wird im weiteren Verlauf der Auswertung der Color Hue – aufgrund der auf seinem zyklischen Charakter basierenden Probleme bei der Korrelationsanalyse – vernachlässigt. Stattdessen werden auf die RGB-Werte zurückgegriffen. Ein klarer Zusammenhang zwischen Tonhöhe und Farbton wie in der Pilotstudie (starker Anstieg des Rot- (r=0,902) und Grünanteils (0,533), aber nicht des Blauanteils mit steigender Frequenz) [10] zeigt sich jedoch nicht.

Farbdimensionen und Wellenformen

Bei t-Tests, in denen die Unterschiede der Farbauswahl zwischen den Wellenformen (inkl. Querflöte) analysiert wurde, wiesen die assoziierten Farben von Dreiecks-, Sinuswellen und Querflöte einen signifikant höheren Blauanteil und eine häufigere „Hellblau“-Auswahl auf gegenüber den Farben der Rechteck- und Sägezahnwellen, für welche signifikant häufiger ein Rotton als passender empfunden wurde. Zwischen den teiltonarmen Klängen (Sinus, Dreieck, Querflöte) gab es wesentlich weniger Unterschiede (für Sinus wurde häufiger Grau gewählt als für Dreieck und die Flöte; für Dreieck häufiger „Rot“ als für die Flöte; häufigere Auswahl von „Hellgrün“ bei der Flöte im Gegensatz zu Sinus) als zu den teiltonreichen Klängen (Rechteck, Sägezahn). Unter letzteren gab es ebenfalls kaum signifikante Unterschiede, lediglich wurde für Rechteckwellen häufiger „Weiß“ gewählt. Korrelationsanalysen und t-Tests unter https://muwiserver.univie.ac.at/synthcolor/statistics.pdf

Farbdimensionen und Klangeigenschaften

Bei der Korrelationsanalyse zwischen den Farbwerten und den Audio Features zeigten sich für fünf dieser Audio Features vielfach Korrelationen zu Farbauswahlhäufigkeiten mit $p<0{,}05$: Timbral Roughness, Timbral Sharpness und Timbral Booming aus den AudioCommons Timbral Models sowie Spectral Complexity und Spectral Skewness aus Essentia (s. Tabelle 2).

Aus dem wiederholten Auftreten dieser Audio Features lässt sich schließen, dass diese fünf zentral sind für die Assoziation von Farben zu synthetischen Klängen.

Tabelle 2: Fünf wiederkehrende Audio Features mit ihren Korrelationen zu den Farbwerten und -auswahlhäufigkeiten; die Zahl gibt den r-Wert der Korrelation an, die Farbe wie stark dieser Wert von Null abweicht.

Five Features:	Timbral Roughness	Timbral Sharpness	Spectral Complexity	Spectral Skewness	Timbral Booming
White	-0,446	-0,256	-0,293		
Light Blue	-0,648	-0,617	-0,416	0,533	0,256
Light Green	0,312	0,519		-0,393	-0,629
Gray	-0,251	-0,39		0,419	0,355
Yellow		0,25	-0,354		-0,512
Orange					-0,336
Pink		0,282			-0,335
Red	0,67	0,588	0,424	-0,49	
Brown	0,298		0,636		0,343
Dark Green					
Dark Blue		-0,403		0,295	0,609
Violet		-0,291			0,467
Black					
R		0,48		-0,343	-0,678
G			-0,505		-0,672
B	-0,758	-0,587	-0,609	0,49	
H		-0,339			0,315
S		0,37	0,405	-0,335	
V		0,264	-0,397		-0,636

Aus den Korrelationen leiten sich einige Grundsätze ab: (1) Je „rauer“ ein Klang ist, desto geringer ist der Blauanteil der assoziierten Farbe (r=-0,758, p<0,001). (2) Je „boomender“ der Klang ist, desto geringer der Rot- (r=-0,678, p<0,001) und der Grünanteil (r=-0,672, p<0,001) in der assoziierten Farbe. (3) Je weniger der Anteil der hohen Teiltöne, angegeben durch die Spectral Skewness, desto größer der Blau- (r=0,49, p<0,001) und desto geringer der Rotanteil (r=-0,343, p=0,005). Somit lässt sich aus der Timbral Roughness, der Timbral Sharpness und der Spectral Skewness über die Korrelationsanalyse mit den Farbwerten ein RGB-Timbre Space erstellen (s. Abbildungen 2 und 3).

Auch anhand der Auswahlhäufigkeit von bestimmten Farbkategorien lassen sich bestimmte Schlüsse ziehen: (1) Je höher die Spectral Complexity, desto häufiger wurde Rot (r=0,424, p<0,001) und desto weniger wurde Blau (r=-0,416, p<0,001) gewählt. (2) Je höher die Spectral Skewness, desto häufiger wurde Blau (r=0,533, p<0,001) und desto seltener wurde Rot (r=-0,49, p<0,001) gewählt. (3) Je „rauer“ ein Klang, desto öfter wurde Rot (r=0,67, p<0,001) und seltener wurde Blau (r=-0,648, p<0,001) gewählt. (4) Je „schärfer“ sich der Klang anhört, desto mehr Rot (r=0,588, p<0,001) und desto weniger Blau (r=-0,617, p<0,001) wurde ausgewählt. (5) Je stärker der Klang „boomte“, desto mehr Farben aus der Blau-Kategorie (r=0,256, p=0,038) und desto weniger grüne Farben (r=-0,629, p<0,001) wurden gewählt.

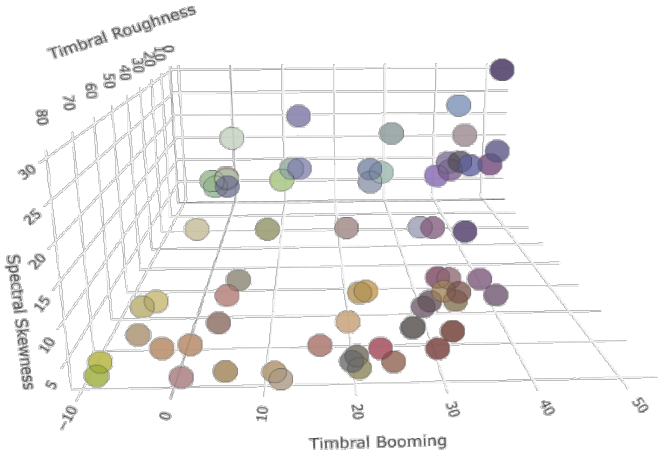

Abbildung 3: Darstellung eines Timbre Space mit den Achsen Timbral Roughness, Timbral Booming und Spectral Skewness korrelierend mit den RGB-Werten der assoziierten Farben: https://muwiserver.univie.ac.at/synthcolor/.

Abbildung 4: QR-Code zur Online-Version des RGB-Timbre Space und weiteren interaktiven Anwendungen.

Die Auswahlhäufigkeit von roten, blauen und grünen Farben und die RGB-Werte der gemittelten Farben ergeben sich in der gleichen Weise aus den fünf Audio Features.

Da mit steigendem Booming, also stärkerer Grundtönigkeit, des Klangs bevorzugt blaue Farben ausgewählt werden und mit steigender klanglicher Schärfe eher rote Farben gewählt werden, lässt sich eine zusätzliche Color Timbre Map zwischen diesen beiden Audio Features aufspannen. Dort

sieht man einen Verlauf der gemittelten Farben von Blau bei hohem Timbral Booming und niedriger Timbral Sharpness zu rot/gelben Farbtönen bei geringem Timbral Booming und hoher Timbral Sharpness (s. Abbildung 5).

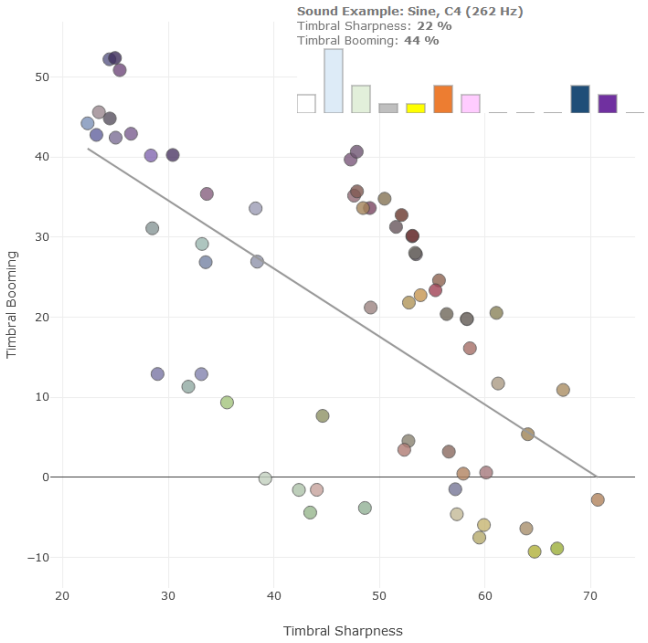

Abbildung 5: Color Timbre Map zwischen Timbral Booming und Timbral Sharpness, die Klänge sind als Punkte in der jeweils gemittelten Farbe dargestellt. Die Audio Features sowie die anteilig gewählten Farbkategorien werden rechts oben beispielhaft für einen Sinusklang auf C4 angezeigt. Interaktive Online-Version unter https://muwiserver.univie.ac.at/synthcolor/blue_red.htm

Zusammenfassung

Die hier vorgestellte Untersuchung bestätigte teilweise die Ergebnisse der 2023 veröffentlichten Pilotstudie [10] und brachte zusätzliche vorläufige neue Erkenntnisse; mit weiteren Versuchspersonen werden hier noch genauere Ergebnisse folgen:

Bekannte Corssmodal Correspondences wie die Verbindung zwischen Tonhöhe und assoziierter Farbhelligkeit konnten bestätigt werden, die in der Pilotstudie gefundene Verbindung zum Color Hue bei Sinus- und Dreieckswellen [10] jedoch nicht. Dennoch zeigte sich wieder, dass die teiltonarmen Sinus-, Dreiecks- und Flötenklänge eher mit blauen Farben und die teiltonreichen Rechteck- und Sägezahnklänge eher mit roten Farben assoziiert werden. Unter jeweils den teiltonarmen und teiltonreichen Wellenarten gibt es nur wenige Unterschiede in der Farbassoziation.

Eine Korrelationsanalyse mit zuvor ermittelten Audio Features wies fünf für nahezu alle Farbwerte besonders einflussreiche Klangmerkmale auf: Spectral Complexity, Spectral Skewness, Timbral Roughness, Timbral Sharpness und Timbral Booming. Diese lassen sich als Grundlage für einen RGB-Timbral Space verwenden, mit dem weitere Zusammenhänge zwischen Klangeigenschaften und Farbassoziationen dargestellt werden können. Bei vier der fünf Audio Features zeigte sich ein Rot-Blau-Gefälle in der Farbauswahlhäufigkeit, dass auch in der Darstellung der RGB-Werte stark auffällt.

Literatur

[1] Parise, C. V.; Spence, C.: Audiovisual crossmodal correspondences and sound symbolism. A study using implicit association test. In: *Experimental Brain Research* 220/3-4 (2012), 319-333.

[2] Ward, J., Huckstep; B., Tsakanikos, E.: Sound-Colour Synaesthesia. To What Extent Does It Use Cross-Modal Mechanisms Common to Us All? In: *Cortex* 42 (2006), 264-280.

[3] Walker, L.; Walker, P.; Francis, B.: A Common Scheme for Cross-Sensory Correspondences across Stimulus Domains. In: *Perception* 41 (2012), 1186-1192.

[4] Spence, C.: Simple and complex crossmodal correspondences involving audition. In: *Acoustical Science and Technology* 41/1 (2020), 6-12.

[5] Bernstein, I. H.; Eason, T. R.: Hue-Tone Sensory Interaction. A Negative Result. In: *Perceptual and Motor Skills* 33/3 (1971), 1327-1330.

[6] Anikin, A.; Johansson, N.: Implicit associations between individual properties of color and sound. In: *Attention, Perception, & Psychophysics* 81 (2019), 764-777.

[7] Qi, Y.; Huang, F.; Li, Z.; Wan, X.: Crossmodal Correspondences in the Sounds of Chinese Instruments. In: *Perception* 49/1 (2020), 81-97.

[8] Grill, T., Flexer, A., Cunningham, S.: Identification of perceptual qualities in textural sounds using the repertory grid method. In: *Proceedings of the 6th Audio Mostly Conference* (2011), 67-74.

[9] Palmer, S. E. et al.: Music-color associations are mediated by emotion. In: *PNAS* 110/22 (2013), 8836-8841.

[10] Feller, G.; Reuter, C. (2023). Klingt Sinus blau und Sägezahn rot? Eine Untersuchung zu Crossmodal Correspondences bei der Wahrnehmung von synthetischen Wellenformen. *Fortschritte der Akustik – DAGA 2023. 49. Deutsche Jahrestagung für Akustik* (S. 1163-1166), Hamburg.

[11] Bogdanov, D et al. (2013). Essentia: An Audio Analysis Library for Music Information Retrieval. *Proceedings of the 14th ISMIR Conference* (S. 493–498), Curitiba.

[12] Pearce, A., et al. (2017). Timbral attributes for sound effect library searching. *AES Proceedings*, Erlangen.