ハーモニックエントロピーの基本

協和の強さを数値化をめざして

- 背景
- 前置き シャノンエントロピー
- 概要
- 定義
- height(高さ/協和距離)
- 確率分散
- 領域積分確率
- 単純重み付き確率
- 実例
- ハーモニックレニイエントロピー

背景

ハーモニックエントロピーて何だ？

和音がどれだけハモるのかを
取り敢えず数値化したかった

基本は２和音（ダイアド）について聴覚的協和感を評価するものと見れる

複雑な和音→情報量が多そう→エントロピー高
単純な和音→情報量が少なそう→エントロピー低

パウル・エリッチが考案、ウィリアム・セタレスらにより研究された

人間の聴覚システムの働き

- 和音の構成音が融合される感覚
- ミッシングファンダメンタルの出現
- うなりの少なさ
- 周期性バズという音の荒々しく聞こえる効果の出現

本来の目的：

音響心理学での仮想的な基音に関する事象の測定のため

言葉の綾・混同やめてくれ

協和は音楽的協和と混同されることが多く、
音響心理の文献で「感覚的協和」と称され
何を指しているかもはや曖昧に・・・

聴覚的協和：西洋音楽では音律論で取り沙汰された
例　バーバーショップ音楽・弦楽器のアンサンブル
　　ミーントーン・ウェルテンペラメント・12EDO

聴覚的協和を最大化する試みと見てとれる

シャノンの
情報エントロピー

情報量

確率が少ないほど情報量が多い

発生確率　1　→情報量　0

発生確率　1/2　→1ビットで表せるから情報量　1

発生確率　1/4　→2ビットで表せるから情報量　2

発生確率　1/n　→$\log_2 n$　ビットで表せる

$$\boxed{-\log_2(P(x))}$$

シャノンエントロピー

情報量の期待値＝(確率×情報量)の総和

情報量の不確定性とか乱雑さの値
取りうるパターン X とその中での確率 P があったとき
情報量の期待値は・・・

$$\boxed{H(X) = -\sum_{x \in X} P(x) \log_2 P(x)}$$

概要

近似音程の確率

音程が何かしらの有理音程(純正音程)に近似して聞かれる
→協和している

何かしらの有理音程から遠くどっちつかずの場合
→不協和である

何らかの音程がいずれかの有理音程として聞かれる確率を
定義して、シャノン情報量の期待値を評価する

情報量による評価

協和的

シャノン情報量が少ない

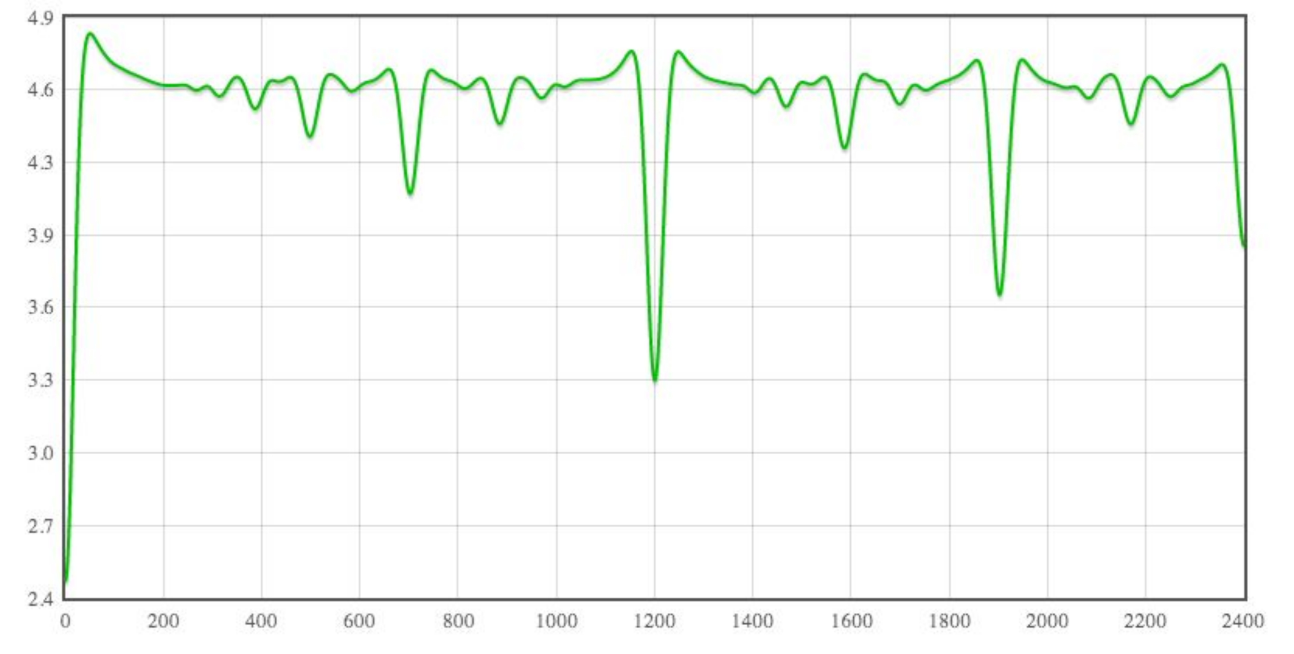
→無数にある有理音程から、特定の有理音程であると判断できると解釈

不協和的

シャノン情報量が多い

→簡単な有理音程からは遠く、どの有理音程として聞かれる確率も低い

グラフで理解しよう


4.9
4.6
4.3
3.9
3.6
3.3
3.0
2.7
2.4
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400

定義

確率の導入

ある音程 c が有理音程 j として聴取される確率

$$P(J = j|C = c)$$

例

400￠の音程が純正音程5/4として聴取される確率

$$P(J = \frac{5}{4}|C = 400\mathrm{cent})$$

確率の導入

いちいち確率を丁寧に書いてるのは大変なので
慣習的にはこう書く→

$$P(j|c)$$

とりあえず、この確率 P の性質について

- j が簡単な比率であるものほど大きい
  →単純な周波数比に近似して聞かれる可能性は高い
- c が j に近いほど大きい
  →精度が高ければ近似して聞かれる可能性は高い

ハーモニック エントロピー

c セントの音程があったとき、そのシャノン情報量を算出

ある音程 c が近似して聞かれる有理音程のパターン J と
その中での確率 P があったとき
情報量の期待値は・・・

$$H(J|c) = -\sum_{j \in J} P(j|c) \log P(j|c)$$

height（高さ/調和距離）

無限にある有理音程

すべての有理音程の集合 J の中にある一つの音程 j を選んで期待値を出し、全部足したかったのに・・・

無限にある純正音程の期待値を片っ端から足し算するのは実際にはムリ！！！

$$H(J|c) = -\sum_{\boxed{j \in J}} P(j|c) \log P(j|c)$$

純正音程の上限

プログラム上でグラフを書くにも・・・

どうにかして有限個の純正音程にしなくてはならない
使う音程、どうやって制限する？

→複雑すぎる音は諦めよう。

上限を決める為の二つの「高さ」/調和距離関数

有理音程 n/d に対して

Tenney Height

分母と分子を掛け算してルート

$$\sqrt{nd}$$

慣例的には N<10000

Weil Height

分母と分子の数そのものを制限
（要するにN次ファレイ数列でやる）

$$\max(n, d)$$

慣例的には N<100

二つの高さ関数の実例

Tenney Height

例

3/2の場合

$$\sqrt{2 \times 3} \approx 2.45$$

Weil Height

例

3/2の場合

2と3だったら

3のほうがでかいから3

確率分散

音程認知がずれる確率とは

ダイアドがどれ程ずれて認知されうるかを一旦定義する

→一般的方法としてガウス分布を使えばいい

慣習的には標準偏差σ =17セント

ラプラス分布とかを使うこともあるが、
とりあえず一般化正規分布を使っておけば大丈夫

まだここではある音程 c が有理音程 j として聴取される確率ではない

$$\boxed{S(x,c) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(x-c)^2}{2\sigma^2})}$$

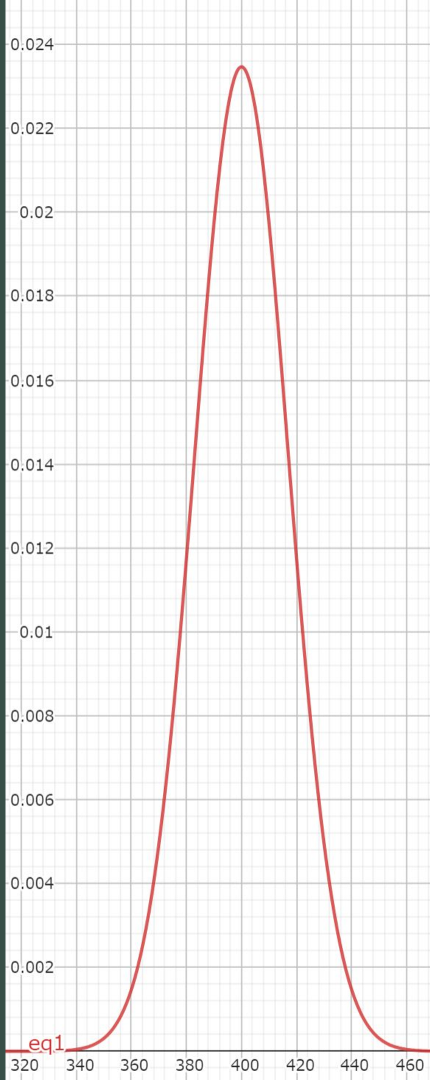
音程認知がずれる確率とは

$$S(x,c) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(x-c)^2}{2\sigma^2})$$

c が他の有理音程 x セントとして聞かれる確率を表す

現状、変数 x は c が何¢として誤認されるかを意味する

400セントが他の音程として聞かれる確率のグラフ(ガウス分布)→
5/4として聞かれる確率は0.017程度となっている

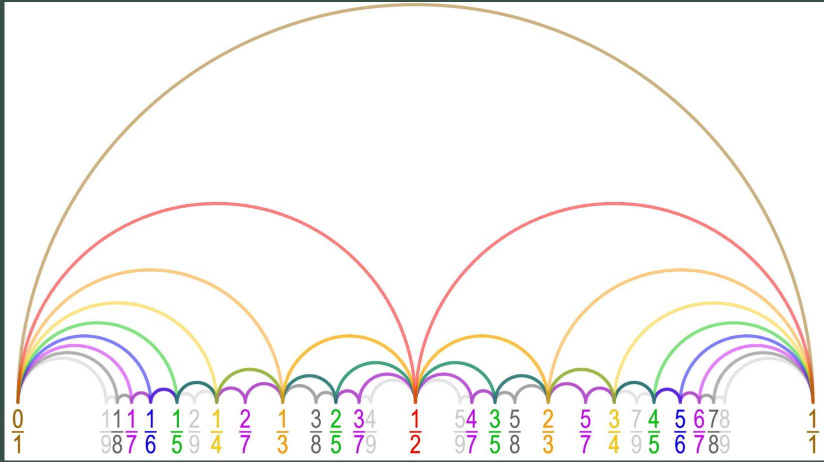


領域積分確率

例の確率の算出法（ウィリアム・セタレスの文献）

- j が簡単な比率であるものほど大きい
  ファレイ数列を見ると、簡単な有理数は
  隣の値との距離が遠い

- c が j に近いほど大きい
  ガウス分布で c が x に近いほど大きくなる



つまり

① j を x セントに変換し
②分布を有理音程ごとの領域として中央値で区分し
③区分ごとガウス分布の面積を調べる

これこそがまさしく $P(j|c)$ では？

何がしたいか

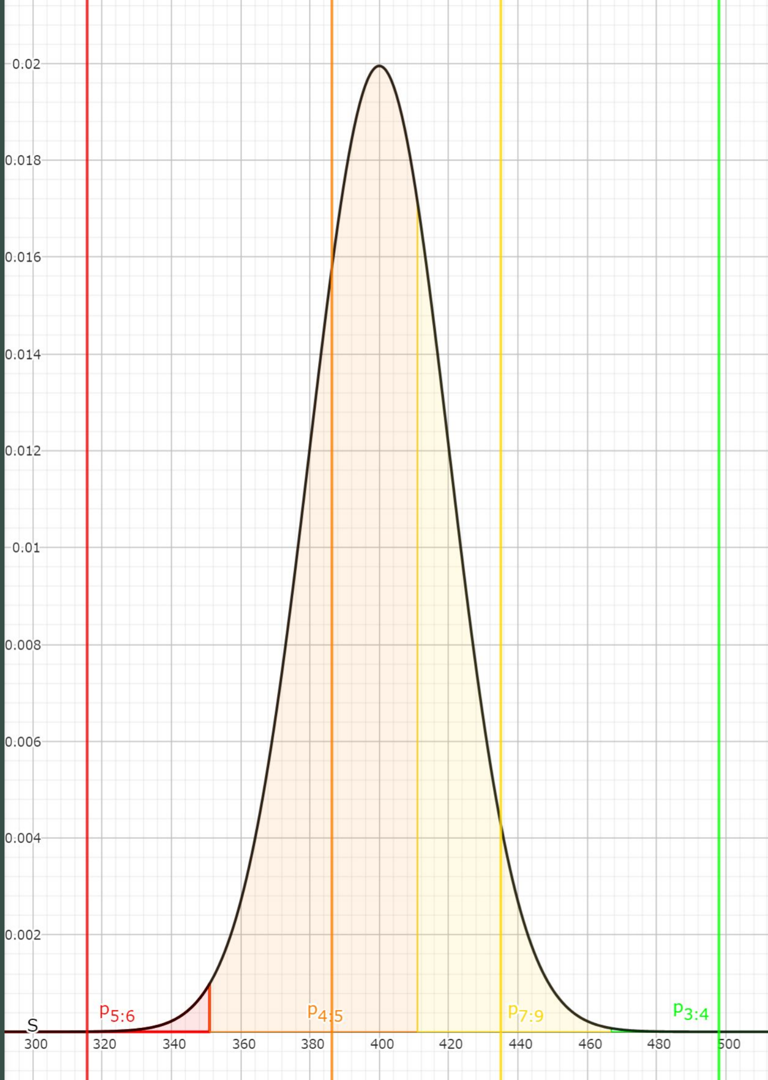
例

400セントがほかの音程として認識される
確率を、各有理音程ごとの領域に区切ってみた

この領域は簡単な有理音程ほどでかい

区分した部分を積分して、出てきた面積こそが
ある音程 c が
有理音程 j として聴取される確率

N=9のとき
5/4 →70%
9/7 →29%
6/5 →0.71%
4/3 →0.04%



領域積分確率

一連の流れを数式にすると

$$Cent(x) = 1200 \log_2 x$$

$$j_u = \mathrm{med}(Cent(j_n), Cent(j_{n+1}))$$

$$j_l = \mathrm{med}(Cent(j_{n-1}), Cent(j_n))$$

$$P(j|c) = \int_{j_l}^{j_u} S(x, c) dx$$

選ばれた純正音程をセントに変換する関数があって、
区分する領域の中央値セントを求めて、その範囲を積分する

現段階のまとめ(ウィリアム・セタレスの方法)

$$Cent(x) = 1200 \log_2 x$$

$$j_u = \mathrm{med}(Cent(j_n), Cent(j_{n+1}))$$

$$j_l = \mathrm{med}(Cent(j_{n-1}), Cent(j_n))$$

$$P(j|c) = \int_{j_l}^{j_u} S(x,c)dx$$

$$S(x,c) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-c)^2}{2\sigma^2})$$

$$H(J|c) = -\sum_{j \in J} P(j|c) \log P(j|c)$$

単純重み付き確率

領域積分確率

片っ端から積分させるのは大変！

そこで、パウル・エリッチによる経験的な近似法を使用

Tenney Heightが N 未満で、有理音程 n/d が持つ領域は
N が十分小さい場合に

$$\frac{1}{\sqrt{nd}}$$

に凡そ比例するっぽい

※任意の大きさのNにおいて これが成立するかは未だに不明だが、
恐らくハーモニックエントロピーを推定するには十分

領域積分確率

Weil Heightの場合でも似たようなことが一応起こる

Weil Heightが N 未満で、有理音程 n/d が持つ領域は
N が十分小さい場合に

$$\frac{1}{\max(n, d)}$$

になんとなく比例するっぽい

※やっぱりこれらは経験則なので、だれか証明してほしい

領域積分確率

いずれにせよ

Heightが N 未満で、有理音程 j が持つ領域幅は
N が十分小さい場合に、j の高さを $||j||$ とすると

$$\frac{1}{||j||}$$

に凡そ比例するっぽい

未正規化確率

一旦はこの確率について$Q(j|c)$と置く

$$\boxed{Q(j|c) = \frac{S(Cent(j), c)}{||j||}}$$

この確率は合計1にならず正規化されていないため、修正する

$$\boxed{P(j|c) = \frac{Q(j|c)}{\sum_{j \in J} Q(j|c)}}$$

→確率の代用完了！

現段階のまとめ(パウル・エリッチの代替法)

$$Cent(x) = 1200 \log_2 x$$

$$S(x, c) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-c)^2}{2\sigma^2})$$

$$Q(j|c) = \frac{S(Cent(j), c)}{||j||}$$

$$P(j|c) = \frac{Q(j|c)}{\sum_{j \in J} Q(j|c)}$$

$$H(J|c) = -\sum_{j \in J} P(j|c) \log P(j|c)$$

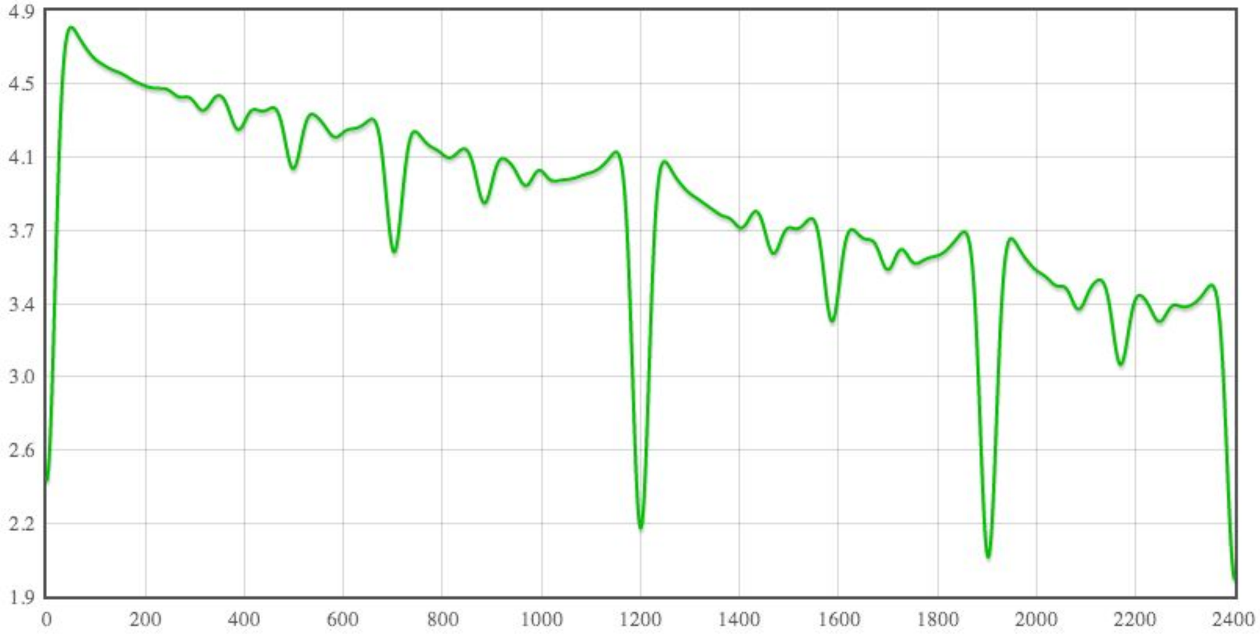
实例

Tenney Height単純重み付きHE

$$\sigma = 17, N = 10000$$

Weil Height単純重み付きHE

$\sigma = 17, N = 100$



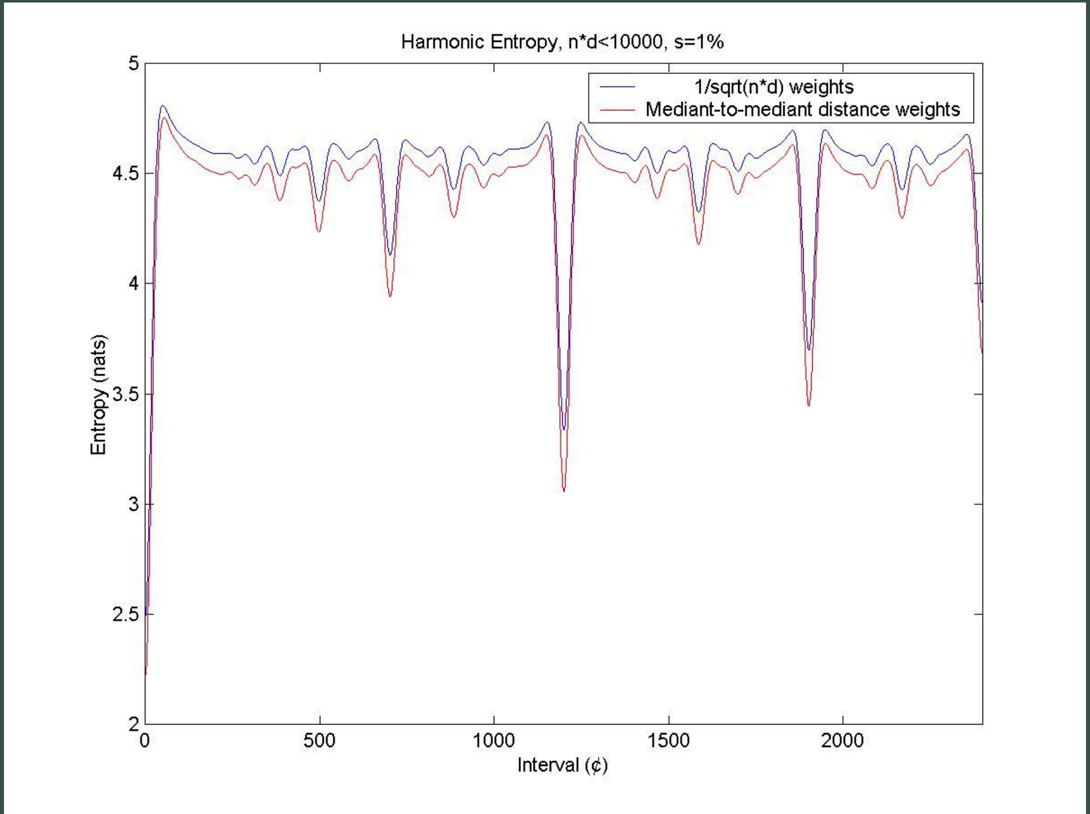
広い音程ほど重みが大きくなる弊害あり

パウル・エリッチによる比較

中央値領域積分重み
(mediant-to-mediant)

平方根重み
(1/sqrt(n*d))

極大と極小の位置はほぼ一致

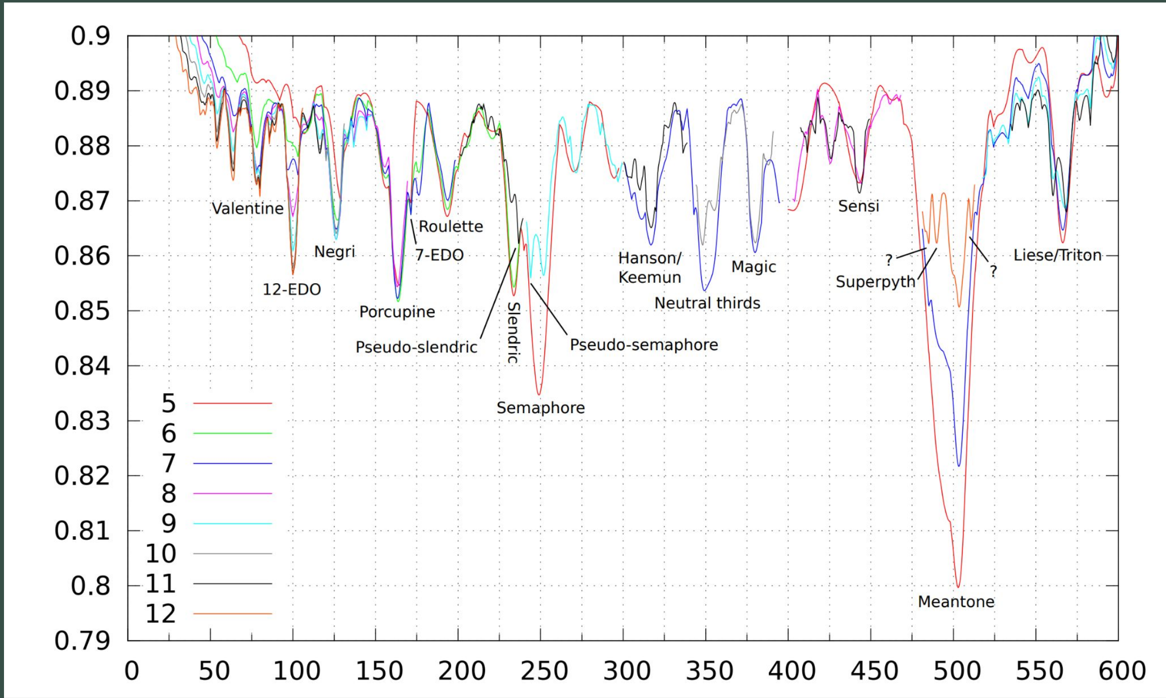


各ジェネレータごとの平均HE

一般的な標準偏差による確率での平均ハーモニックエントロピー
様々なMOSスケールについて比較可能

5音音階、7音音階ともに
明らかにMeantoneが強い
(697¢付近が最小に→31EDO近辺)

完全5度を多く持ち、
ほかの簡単な音程を
次々ととっている為

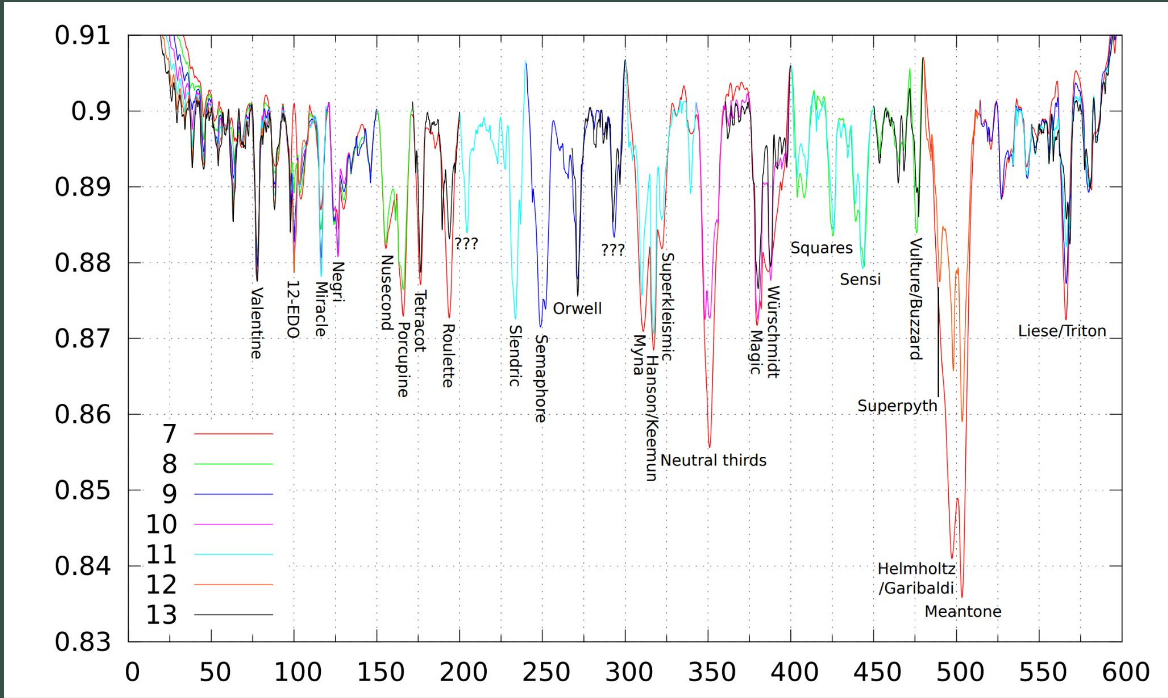


各ジェネレータごとの平均HE

より狭い標準偏差による確率での平均ハーモニックエントロピー
より高精度に谷が出現

またしても
明らかにMeantoneが強い
(697￠付近が最小に→31EDO近辺)

音数が増えると、
ほかの音階の選択肢も
見えてくる



ハーモニック
レニイエントロピー

確率を知っていると確率が変わる(ベイズ確率)

Aの箱→白い碁石×？、黒い碁石×？
Bの箱→白い碁石×？、黒い碁石×？

白い碁石が出たとき、Aの箱から取り出した確率

→箱をランダムに選ぶので 0.5

Aの箱→白い碁石×3、黒い碁石×1
Bの箱→白い碁石×2、黒い碁石×2

白い碁石が出たとき、Aの箱から取り出した確率

$$P(A|W)\frac{0.5\times0.75}{0.5\times0.75+0.5\times0.5}=0.6$$

→0.6

暗号学での用途

セキュリティの分野で重要なのは
暗号解読に使われる鍵が予測不能であること

コンピュータで使用される疑似乱数は予測できることがあるため、
敵に事前に疑似乱数の知識がある場合、
敵に解読されてしまう確率が高まる

事前知識を加味したエントロピーの概念が必要

シャノンエントロピーでは力及ばず

HEでの用途

暗号学の考えを音楽的な場面として読み替えると
聞こえたダイアドが、脳(敵)によって解読されようとしている

jとして認知される音程がcである確率を知っている
「優秀な聴覚」を想定した場合
cが聞こえたとき、jとして識別される可能性が高まる

このような場合は積極的に確率の高い有理音程を選ぼうとするので

シャノンエントロピーでは力及ばず

具体例

3/2として認知される音程が700￠である確率→かなり高いと思う
700￠が聞こえたとき、3/2として推定する
可能性がさらに高まる

恐らくこの時、認知的には
標準偏差σが大きくともほとんどの人が
16/11(648￠)や
20/13(746￠)はあまり選択しないはずである

情報量は確率をもとに修正されなければならない

ハーモニック レニイ エントロピー

Harmonic Rényi Entropy

マイク・バッタグリア氏による拡張概念
どれだけ優秀な聴覚システムかを示すパラメータ
「α」を追加

$$H_\alpha(J|c) = \frac{1}{1-\alpha}\log\sum_{j\in J} P(j|c)^\alpha$$

シャノンエントロピーを含めた一般化と見てとれる

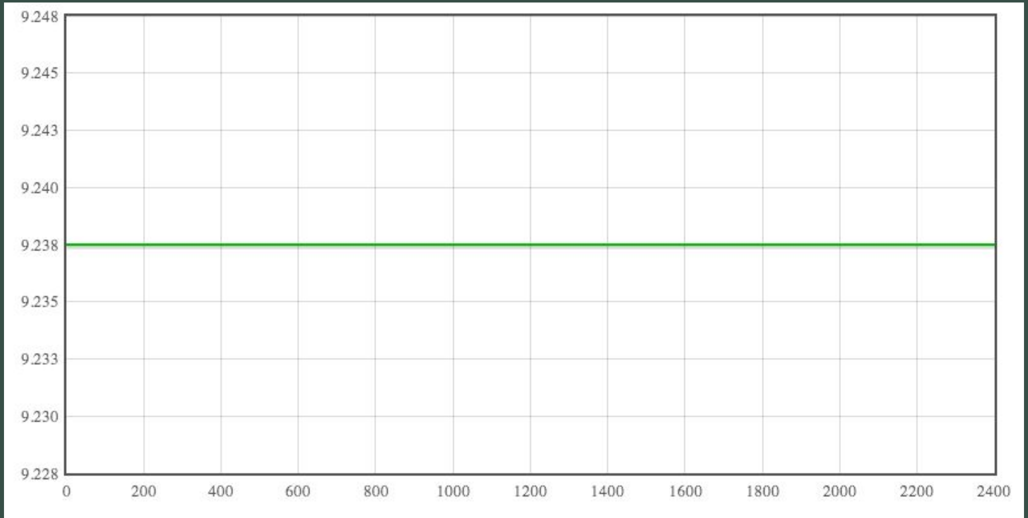
α=0 ハーモニック ハートレイ エントロピー

Harmonic Hartley Entropy

聞こえた音からランダムな有理音程しか選択できない
最もマヌケな聴覚

$$\boxed{H_0(J|c) = \log|J|}$$

一様に値が大きい→



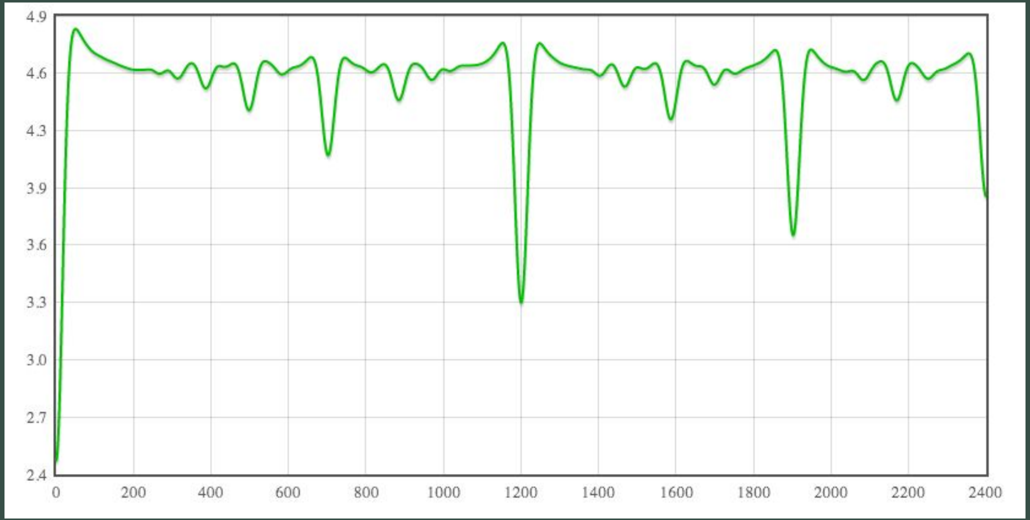
$\alpha \to 1$ ハーモニック シャノン エントロピー

Harmonic Shannon Entropy

積極的に音程を選択しようとしない場合のエントロピー

$$\boxed{H_1(J|c) = -\sum_{j \in J} P(j|c) \log P(j|c)}$$

何となく値が高い→

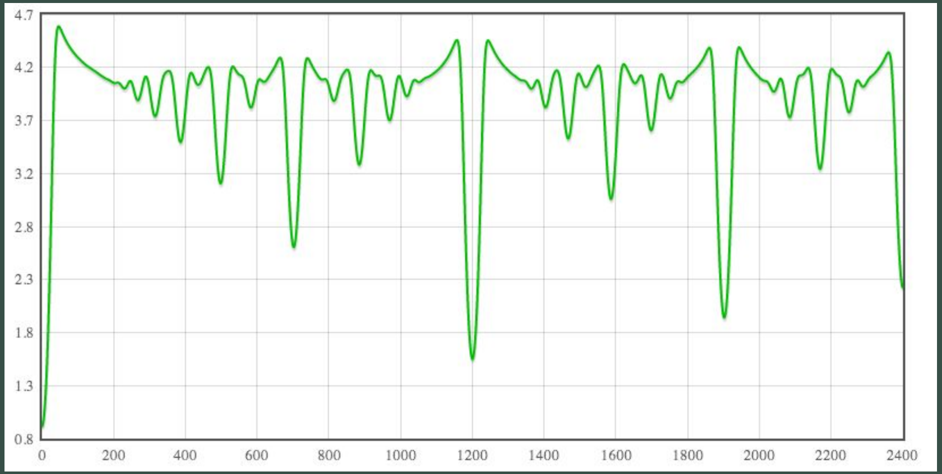


α=2 ハーモニック 衝突 エントロピー

Harmonic Collision Entropy
脳が候補を二回選んで、結局それが一致する確率の情報量の期待値

$$\boxed{H_2(J|c) = -\log \sum_{j \in J} P(j|c)^2 = -\log P(J_1 = J_2|c)}$$

より谷の部分が明瞭になる→



$\alpha \to \infty$ ハーモニック 最小エントロピー

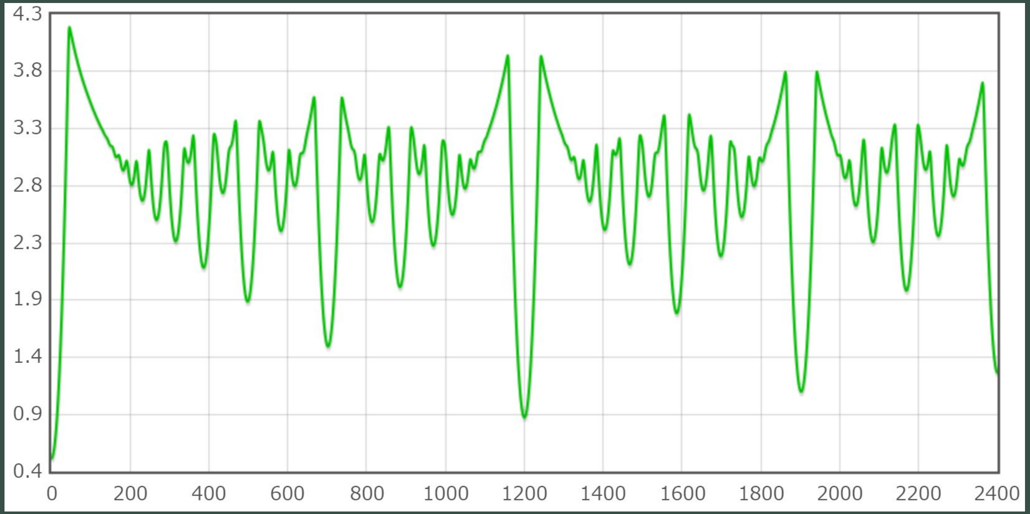
Harmonic Min-Entropy

優秀な聴覚を相手にしたエントロピー

$$\boxed{H_\infty(J|c) = -\log \max_{j \in J} P(j|c)}$$

深い谷ができる→

アプリで計算できないのでα=10でのグラフ



マイク・バッタグリア氏のウェブアプリ

https://www.mikebattagliamusic.com/HE-JS/HE.html

ハーモニックレニイエントロピーが
計算できる便利なウェブアプリ

記号の置き換えに注意

標準偏差 σ→s
レニイエントロピーのパラメータ α→a

ありがとうございました。

画像引用一覧

https://en.xen.wiki

https://www.mikebattagliamusic.com/HE-JS/HE.html