

性差・年齢差を越えた 音声のモデル化とその応用

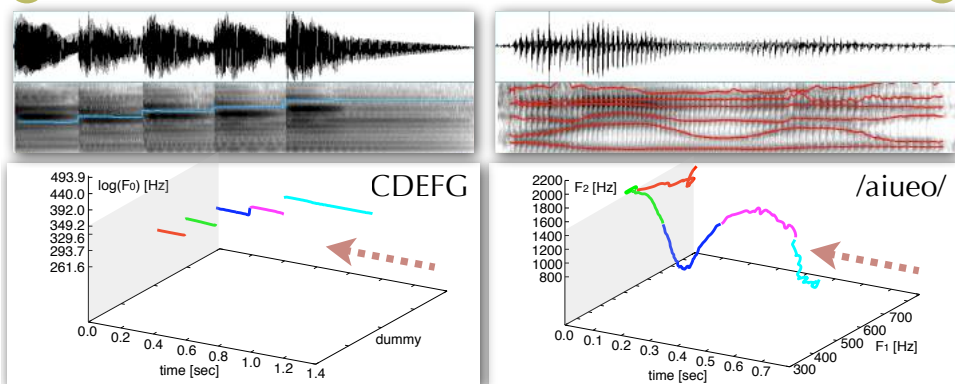
峯松 信明

大学院工学系研究科

工学部電子情報工学科

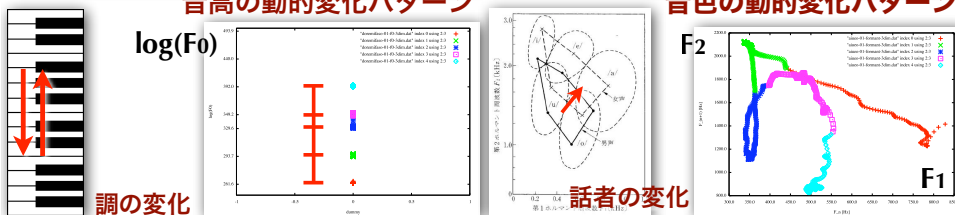


音高の相対音感／音色の相対音感



音高の動的変化パターン

音色の動的変化パターン



調の変化

話者の変化

本講義の流れ

抽象化・モデル化と音声コミュニケーション技術

音声の物理学・基礎編

● 音声の生成からその物理的（音響的）側面まで

● 人間が受け取るメディア情報の多様性と不変性

● 視・聴・触・味・嗅、多様に変形しても同じだと分かる。

● 多様性の「そぎ落とし」と内在する不変性

● その数学的な回答とそれに基づく音声のモデル化

● 多様性に打ち勝つ音声技術の構築と応用

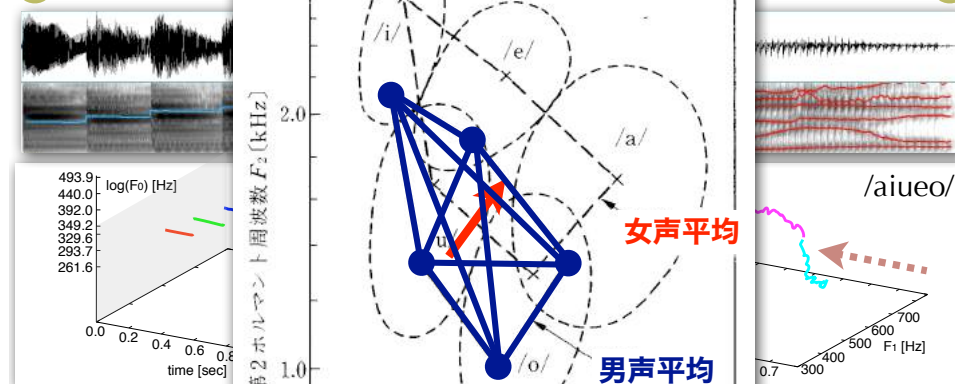
● 音声認識，外国語発音分析を例にとって

● 「そぎ落とし」が気付かせてくれる異分野との接点

● あれと，あれは，同じ原理なのか？

● 惑星とリンゴは同じである。音声とXXは同じである。

音高の相対音感／音色の相対音感



女声平均

男声平均

音高

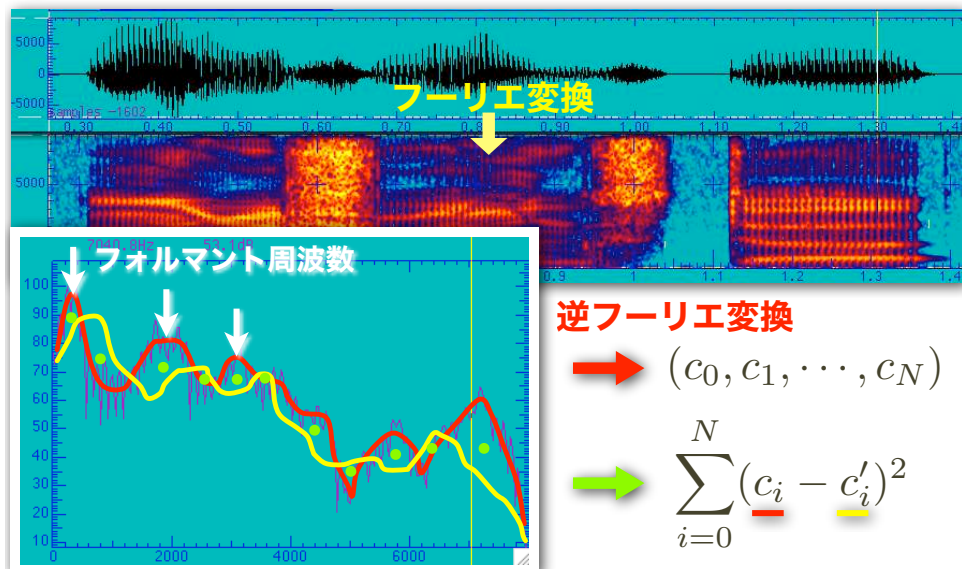
調の変化

音色の動的変化パターン

話者の変化

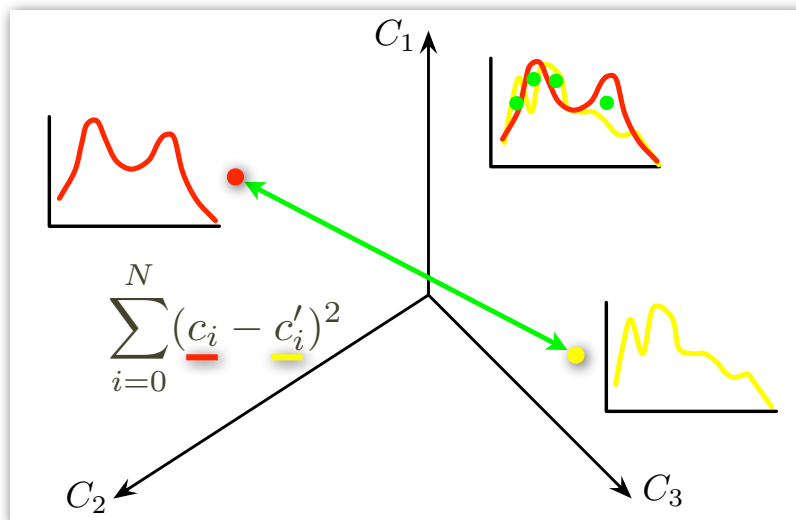
スペクトル包絡の効率的なベクトル表現

音声波形→スペクトラム→ケプストラム



スペクトル包絡の効率的なベクトル表現

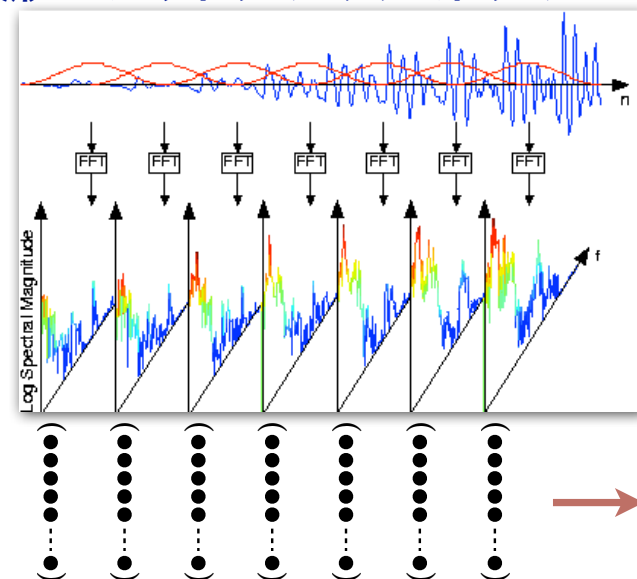
ケプストラム空間における「点」と「点間距離」



点=スペクトル包絡, 点間距離=スペクトル間差異

スペクトル包絡の効率的なベクトル表現

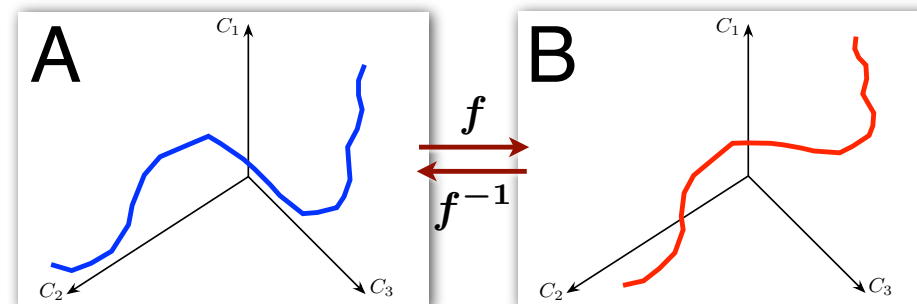
音声波形→スペクトラム→ケプストラム



変換不変な音響量の数学的探求

話者の違い=空間写像 (話者・声質変換)

話者Aの音響空間 ↔ 話者Bの音響空間



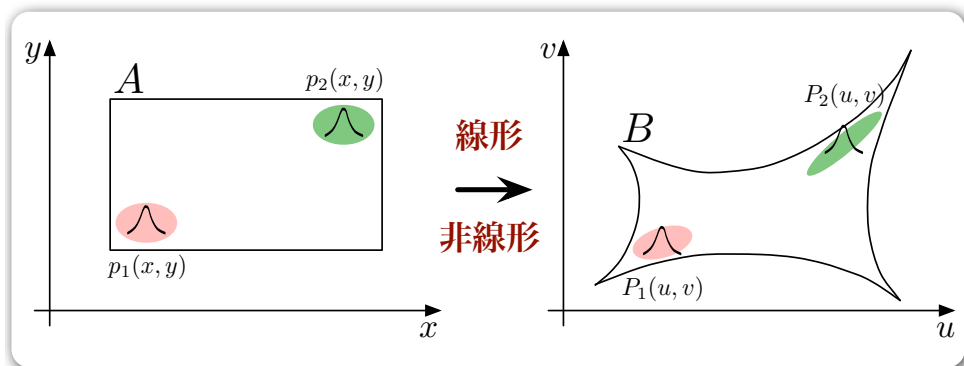
話者Aの声を65億全ての話者の声へ変形する

- 65億 x 65億 の写像関数が定義可能
- 話者不変のコントラスト=写像不変のコントラスト
- 任意の写像に対して不変なるコントラスト量は存在するのか?

変換不変な音響量の数学的探求

二人の話者空間（一対一対応）における不変音響量

- 音響事象を点ではなく、分布として表現する。
- 空間Aの分布 p は空間Bの分布 P へと写像される。
- p と P は異なる物理特性を持つ（[あ] と [あ]）
- 両空間において不変な物理量はどこに？不変コントラストは存在する？



変換不変な音響量の数学的探求

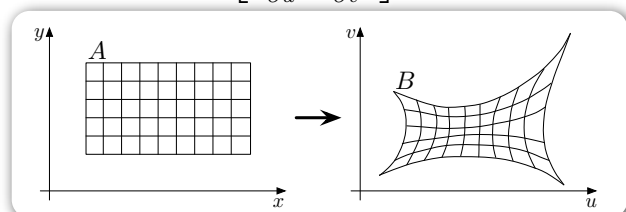
変数変換と確率密度分布関数

- 一変数： $x = x(t)$ ($x_1 = x(t_1), x_2 = x(t_2)$)

$$1.0 = \int_{x_1}^{x_2} p(x) dx = \int_{t_1}^{t_2} p(x(t)) \frac{dx(t)}{dt} dt = \int_{t_1}^{t_2} q(t) x'(t) dt$$
- 二変数： $x = x(u, v), y = y(u, v)$

$$1.0 = \iint_A p(x, y) dx dy = \iint_B q(u, v) |J(u, v)| du dv$$

$$J(u, v) \equiv \frac{\partial(x, y)}{\partial(u, v)} \equiv \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$$



変換不変な音響量の数学的探求

変数変換と積分

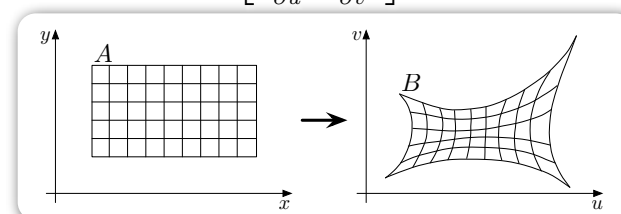
- 一変数： $x = x(t)$ ($x_1 = x(t_1), x_2 = x(t_2)$)

$$\int_{x_1}^{x_2} f(x) dx = \int_{t_1}^{t_2} f(x(t)) \frac{dx(t)}{dt} dt = \int_{t_1}^{t_2} g(t) x'(t) dt$$
- 二変数： $x = x(u, v), y = y(u, v)$

$$\iint_A f(x, y) dx dy = \iint_B g(u, v) |J(u, v)| du dv$$

$$J(u, v) \equiv \frac{\partial(x, y)}{\partial(u, v)} \equiv \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$$

$$\begin{aligned} x &= 3u + 2v - 5 \\ y &= 4u + 5v + 3 \\ x &= r \cos \theta \\ y &= r \sin \theta \end{aligned}$$



変換不変な音響量の数学的探求

バタチャリヤ距離（=分布間距離尺度の一つ）

- 座標変換による式の変形 $x = x(u, v), y = y(u, v)$
- $BD(p_1(x, y), p_2(x, y))$

$$= -\log \iint \sqrt{p_1(x, y) p_2(x, y)} dx dy$$

$$= -\log \iint \sqrt{q_1(u, v) q_2(u, v)} |J(u, v)| du dv$$

$$= -\log \iint \sqrt{q_1(u, v) |J(u, v)| \cdot q_2(u, v) |J(u, v)|} du dv$$

$$= -\log \iint \sqrt{P_1(u, v) P_2(u, v)} du dv$$

$$= BD(P_1(u, v), P_2(u, v))$$

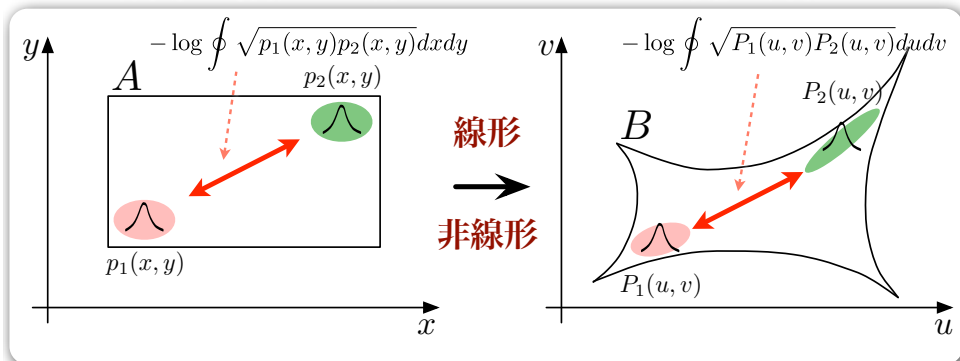
$$q_1(u, v) = p_1(x(u, v), y(u, v)), \quad J = \text{Jacobian}$$



変換不変な音響量の数学的探求

二人の話者空間（一対一対応）における不変音響量

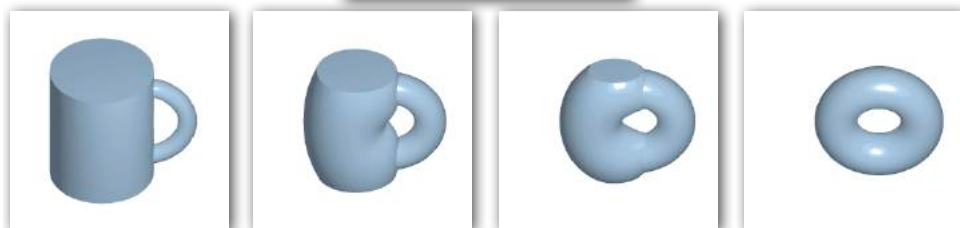
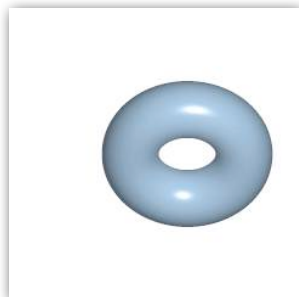
- 音響事象を点ではなく、分布として表現する。
- 空間Aの分布 p は空間Bの分布 P へと写像される。
- p と P は異なる物理特性を持つ（[あ] と [あ]）
- 両空間において不変な物理量はどこに？不変コントラストは存在する？
- 各事象は可変、しかし、少なくともバタチャリヤ距離は不変。



変換不変な音響量の数学的探求

位相幾何学（トポロジー）における不変量

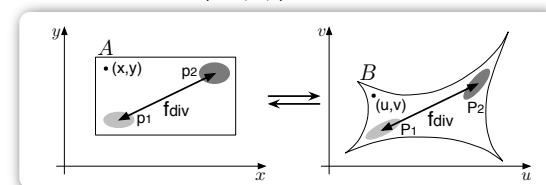
- 連続かつ可逆な任意の変形を施しても不変なる幾何学的性質



変換不変な音響量の数学的探求

変換不変量の一般式はあるのか？

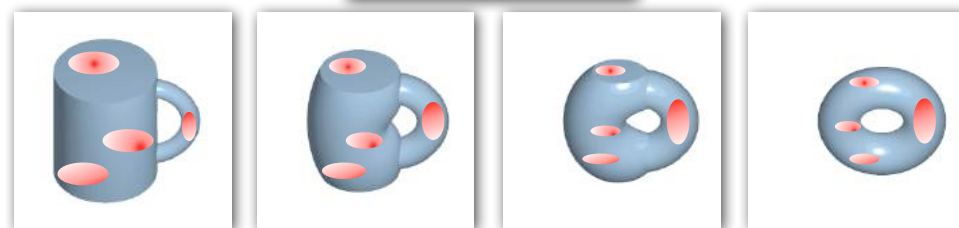
- f-divergence 不変性の十分性
 - $f_{div}(p_1, p_2) = \int p_2(x) g\left(\frac{p_1(x)}{p_2(x)}\right) dx$
 - $g(t) = t \log(t) \rightarrow f_{div} = \text{KL} - \text{div.}$ $g(t) = \sqrt{t} \rightarrow -\log(f_{div}) = \text{BD}$
 - $f_{div}(p_1, p_2) = f_{div}(P_1, P_2)$
- f-divergence 不変性の必要性
 - $\int M(p_1(x), p_2(x)) dx$ が任意の可逆&連続変換に対して不変となる条件
 - この場合、 $M = p_2(x) g\left(\frac{p_1(x)}{p_2(x)}\right)$ であることが必要。



変換不変な音響量の数学的探求

位相幾何学（トポロジー）における不変量

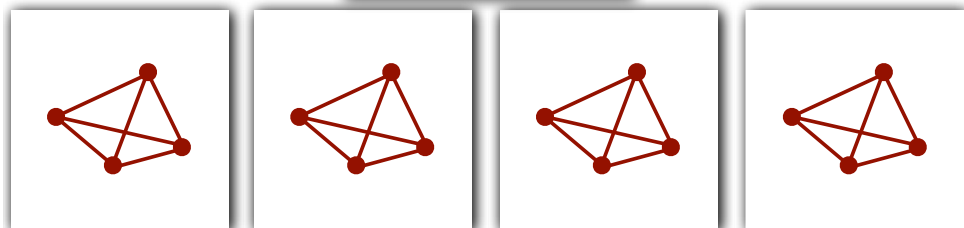
- 連続かつ可逆な任意の変形を施しても不変なる幾何学的性質



変換不変な音響量の数学的探求

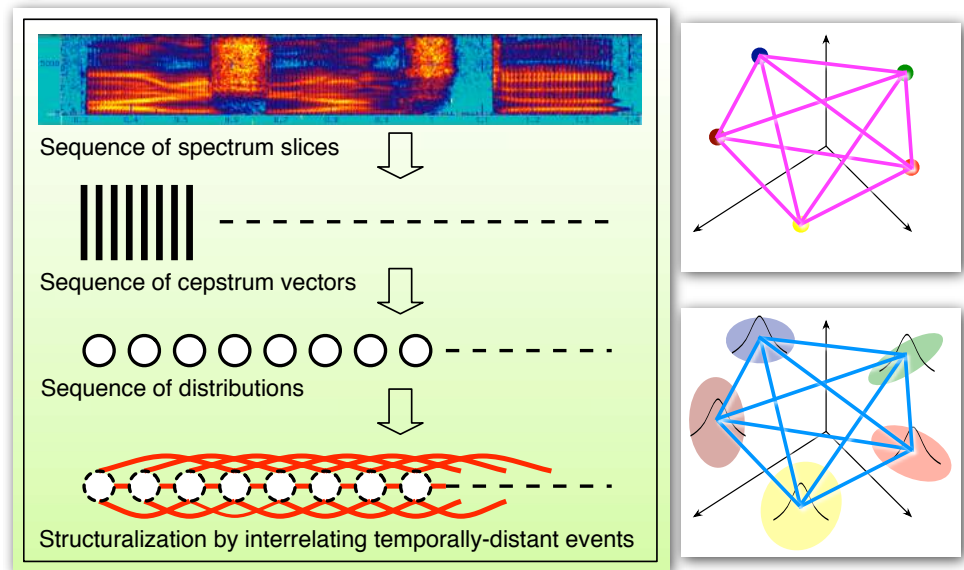
位相幾何学（トポロジー）における不変量

連続かつ可逆な任意の変形を施しても不変なる幾何学的性質



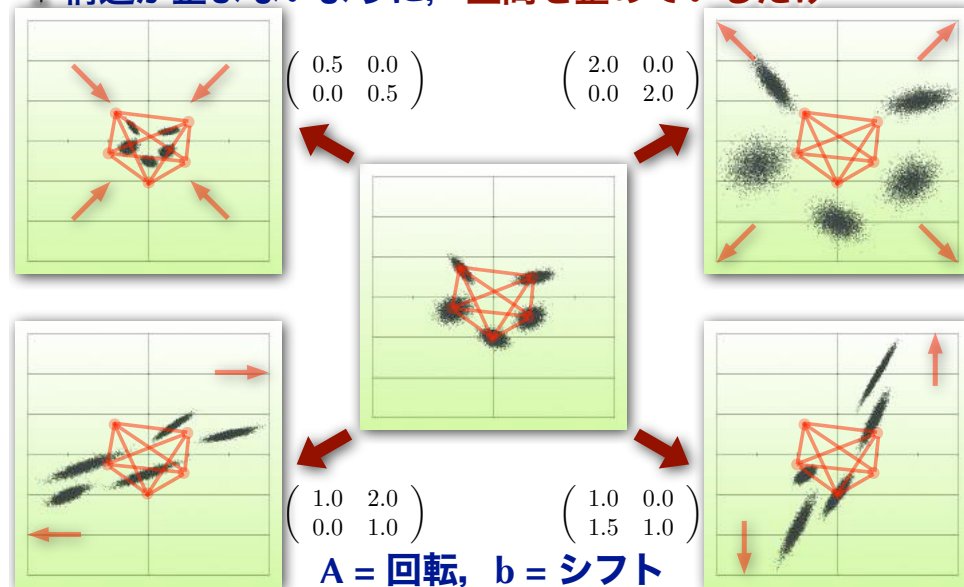
分布間距離群としての音声表象

頑健に話者不変な音声表象＝構造的・全体的表象



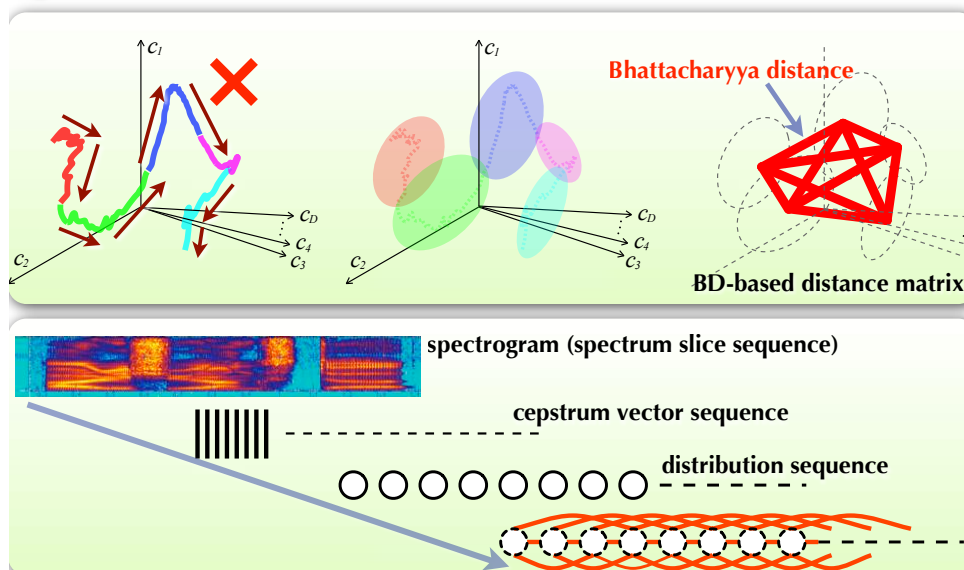
分布間距離群としての音声表象

構造が歪まないように、空間を歪めているだけ・・・



分布間距離群としての音声表象

ケプトラム系列 → 分布系列 → 距離行列



音色の偏差とその認知的不変性

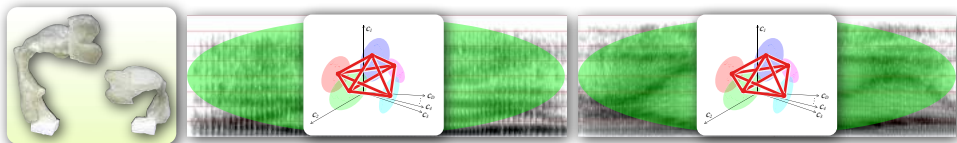
色み・音高の恒常・不変的認知

- 🎧 コントラスト情報に基づく処理が重要
- 🎧 コントラスト群から成る全体的パターン処理が要素同定を可能



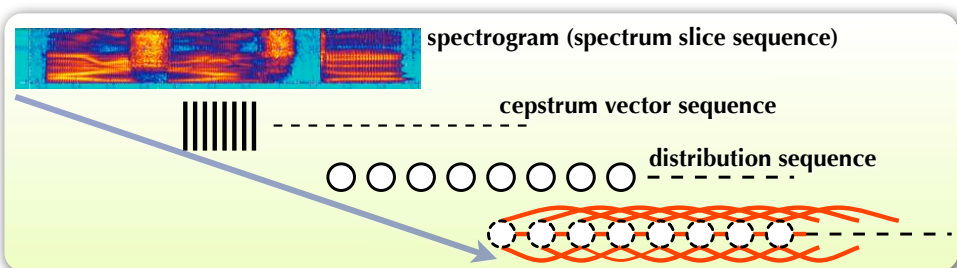
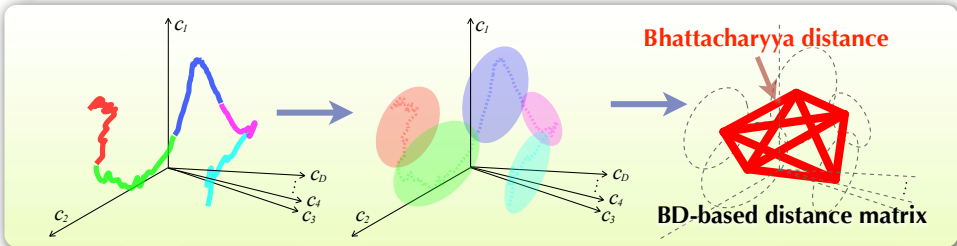
音色の恒常・不変的認知

- 🎧 コントラスト情報に基づく処理が重要
- 🎧 コントラスト群から成る全体的パターン処理が要素同定を可能



音声の構造的表象の工学的・実験的検証

f-div. (BD)に基づく一発声の構造化



本講義の流れ

抽象化・モデル化と音声コミュニケーション技術

音声の物理学・基礎編

- 🎧 音声の生成からその物理的（音響的）側面まで

人間が受け取るメディア情報の多様性と不変性

- 🎧 視・聴・触・味・嗅、多様に変形しても同じだと分かる。

多様性の「そぎ落とし」と内在する不変性

- 🎧 その数学的な回答とそれに基づく音声のモデル化

多様性に打ち勝つ音声技術の構築と応用

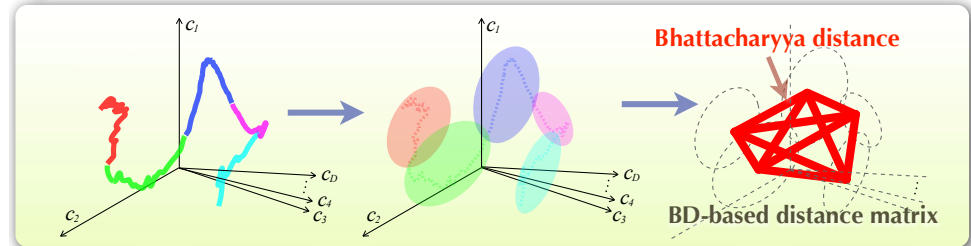
- 🎧 音声認識，外国語発音分析を例にとって

「そぎ落とし」が気付かせてくれる異分野との接点

- 🎧 あれと，あれは，同じ原理なのか？
- 🎧 惑星とリンゴは同じである。音声とXXは同じである。

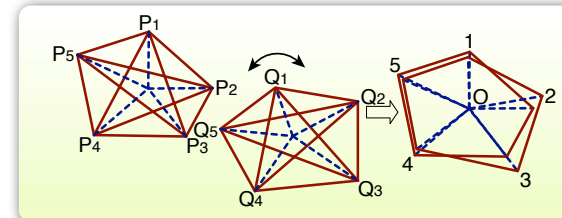
音声の構造的表象の工学的・実験的検証

f-div. (BD)に基づく一発声の構造化



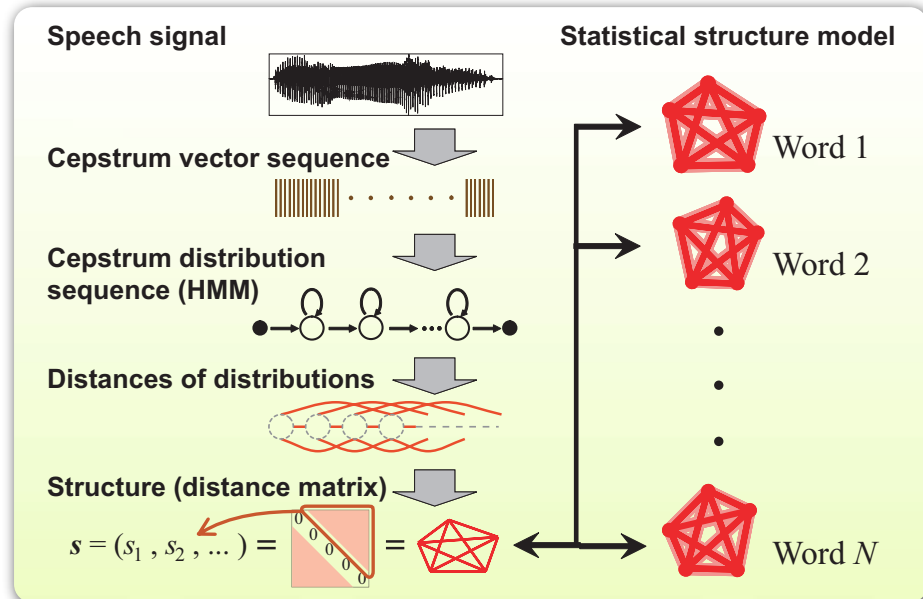
2発声（＝2距離行列）間の音響照合

- 🎧 2距離行列間のユークリッド距離



- 🎧 回転：声道長差異
- 🎧 シフト：マイク差異
- 🎧 話者適応・環境適応後のスコアが適応処理無しで算出
- 🎧 話者性を削除した音声表象

音声の構造的表象の工学的・実験的検証



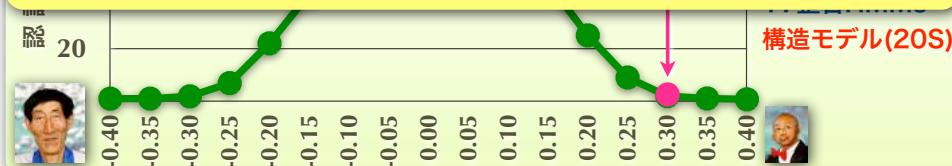
音声の構造的表象の工学的・実験的検証

孤立単語音声の認識実験

- 二つの問題とその解決
- 強すぎる不変性→マルチストリーム構造化による都合のよい不変性へ

但し、孤立音の同定は不可能
そこにはコントラストが無いから

孤立提示された音を音韻同定する能力は
音声言語運用には不要なのかもしれない



音声の構造的表象の工学的・実験的検証

孤立単語音声の認識実験

- 二つの問題とその解決
- 強すぎる不変性→マルチストリーム構造化による都合のよい不変性へ
- 高すぎる次元数→線形判別分析 (LDA) による次元数削減
- 孤立単語認識実験による提案手法の評価
- 日本語五母音を並び替えて作成される120単語の孤立単語認識



音色の偏差とその認知的不変性

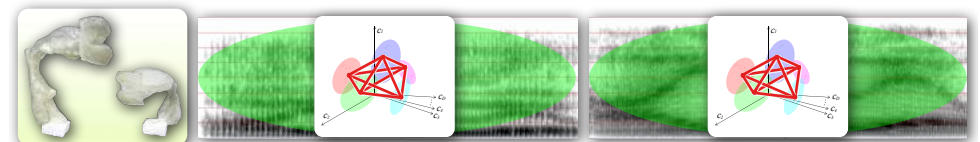
色み・音高の恒常・不変的認知

- コントラスト情報に基づく処理が重要
- コントラスト群から成る全体的パターン処理が要素同定を可能



音色の恒常・不変的認知

- コントラスト情報に基づく処理が重要
- コントラスト群から成る全体的パターン処理が要素同定を可能



本講義の流れ

抽象化・モデル化と音声コミュニケーション技術

音声の物理学・基礎編

- 音声の生成からその物理的（音響的）側面まで

人間が受け取るメディア情報の多様性と不変性

- 視・聴・触・味・嗅、多様に変形しても同じだと分かる。

多様性の「そぎ落とし」と内在する不変性

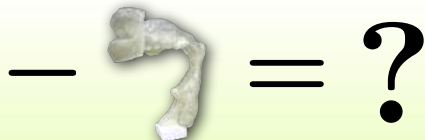
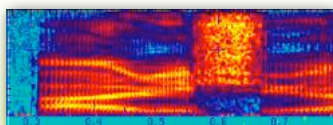
- その数学的な回答とそれに基づく音声のモデル化

多様性に打ち勝つ音声技術の構築と応用

- 音声認識, 外国語発音分析を例にとって

「そぎ落とし」が気付かせてくれる異分野との接点

- あれと、あれは、同じ原理なのか？
- 惑星とリンゴは同じである。音声とXXは同じである。



- 個々の発声における、不適切な発音部位を検出する。

発音評価技術構築における根本的な問題

- 教師音声データベースから個々の音素の音響モデルを構築する。
- 学習者は小学生かもしれない。
- 音韻の違い＝話者の違い＝管形状の違い＝音色の違い

学習者にあった音響モデルの構築

- 話者適応技術
 - 大人用モデル／子供用モデル
- 発音評定？声帯模写評定？

峯松の声から峯松を消したい

- そんなことができるのか？



そもそもの出発点はここ

学習者の英語発音を評価する技術の構築

- 発音の習熟度を定量的にスコア化する。
- 個々の発声における、不適切な発音部位を検出する。

発音評価技術構築における根本的な問題

- 教師音声データベースから個々の音素の音響モデルを構築する。
- 学習者は小学生かもしれない。
- 音韻の違い＝話者の違い＝管形状の違い＝音色の違い

学習者にあった音響モデルの構築

- 話者適応技術
 - 大人用モデル／子供用モデル
- 発音評定？声帯模写評定？

峯松の声から峯松を消したい

- そんなことができるのか？



で、こういうのができました。

峯松の普通の英語発音はどちらに近いのか？

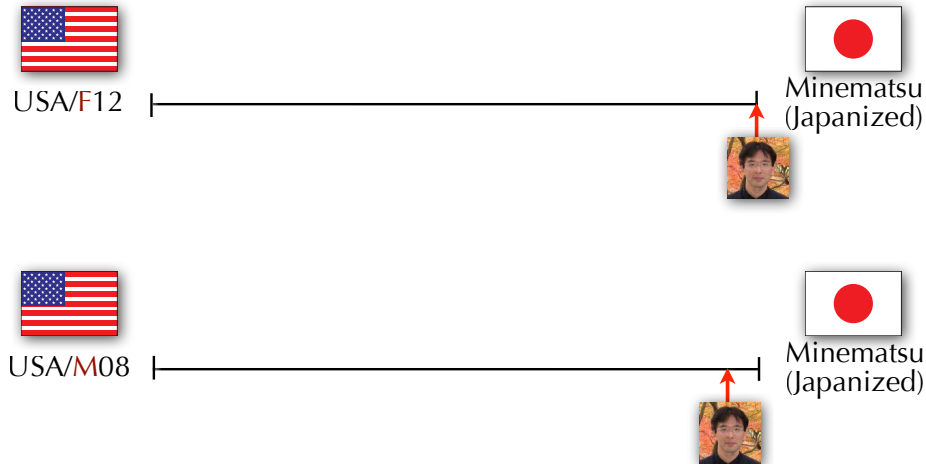
speaker	USA/F12	✗ Minematsu	○ Minematsu
gender	female	✗ male	○ male
age	??	✗ 36	○ 36
mic	Sennheiser	✗ cheap mic	○ cheap mic
room	recording room	✗ living room	○ living room
AD	SONY DAT	✗ PowerBook	○ PowerBook
proficiency	perfect	△ good	✗ Japanized

Those answers will be straight forward if you think them through carefully first.

で、こういうのができました。

母語話者発音と日本語的峯松発音の音響モデルを利用

各モデルから峯松英語発声が観測される確率 $P(o|M)$



で、こういうのができました。

母語話者発音と日本語的峯松発音の音響モデルを利用

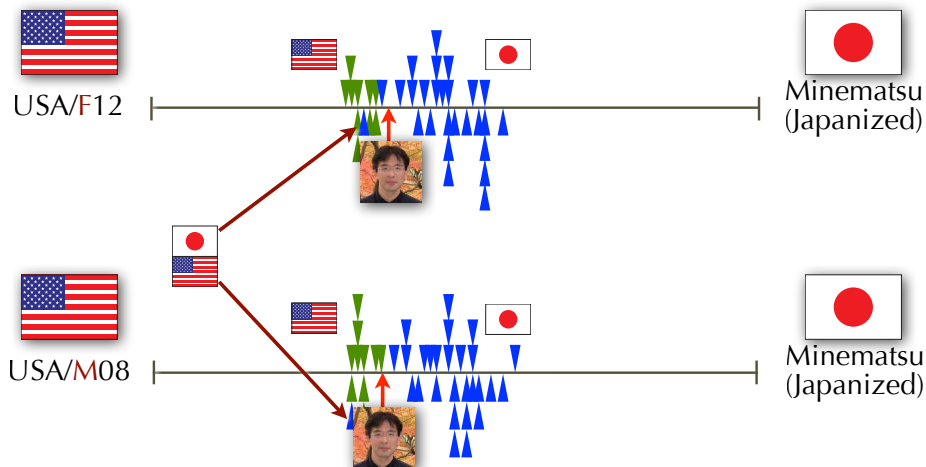
各モデルから峯松英語発声が観測される確率 $P(M|o)$

$$\begin{aligned}
 P(M|o) &= P(p_1, \dots, p_N | o) \\
 &= \frac{P(o | p_1, \dots, p_N) P(p_1, \dots, p_N)}{\sum_{p_i} P(o | p_1, \dots, p_N) P(p_1, \dots, p_N)} \\
 &\approx \frac{P(o | p_1, \dots, p_N)}{\sum_{p_i} P(o | p_1, \dots, p_N)} \\
 &\approx \frac{P(o | p_1, \dots, p_N)}{\max_{p_i} P(o | p_1, \dots, p_N)} \\
 &= \frac{P(o|M)}{\max_M P(o|M)} \\
 &= \text{GOP (Goodness Of Pronunciation)}
 \end{aligned}$$

で、こういうのができ

母語話者発音と日本語的峯松発音

母語発音構造, 日本語的峯松発音構造と, 峯松発音構造を比較



proficiency

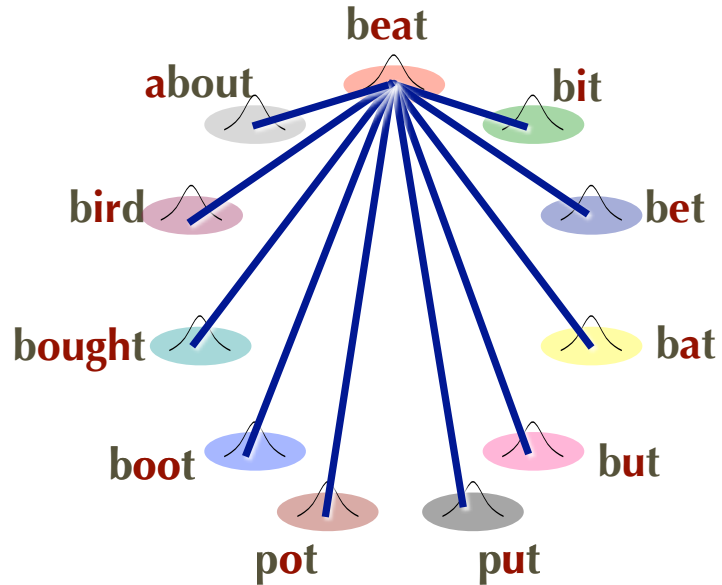
perfect

good

japanized

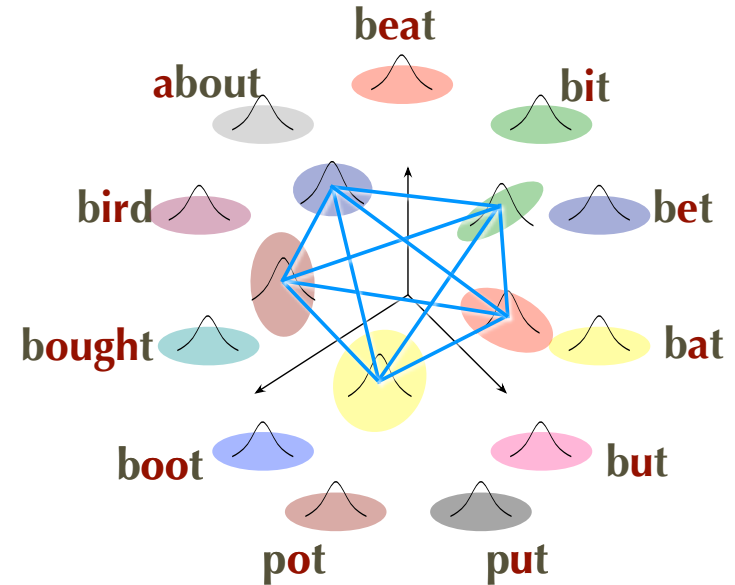
英語母音の発音トレーニング

個々の母音ではなく、母音群の体系を評価対象



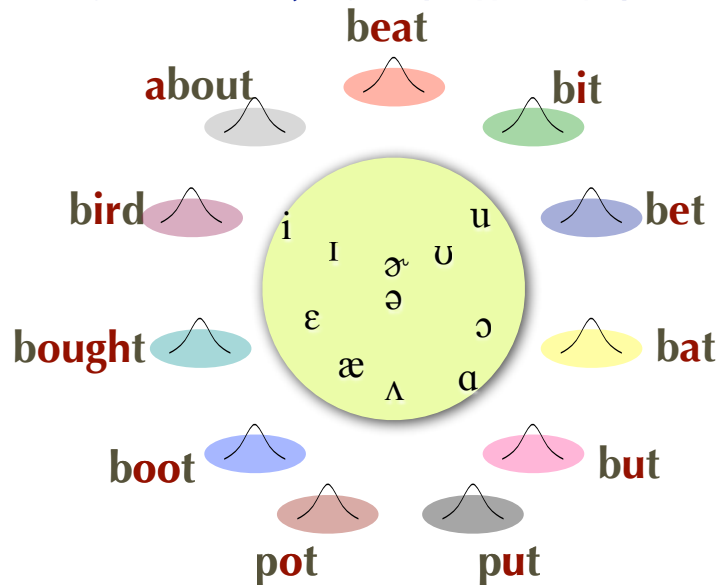
英語母音の発音トレーニング

個々の母音ではなく、母音群の体系を評価対象



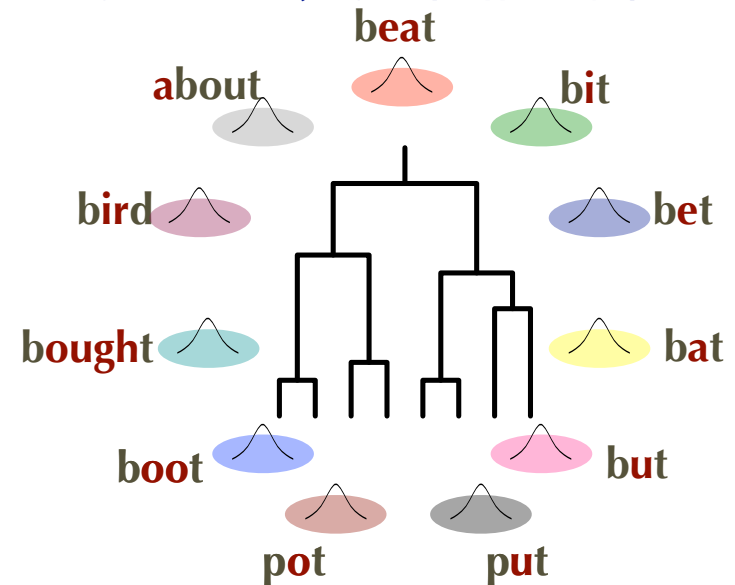
英語母音の発音トレーニング

個々の母音ではなく、母音群の体系を評価対象



英語母音の発音トレーニング

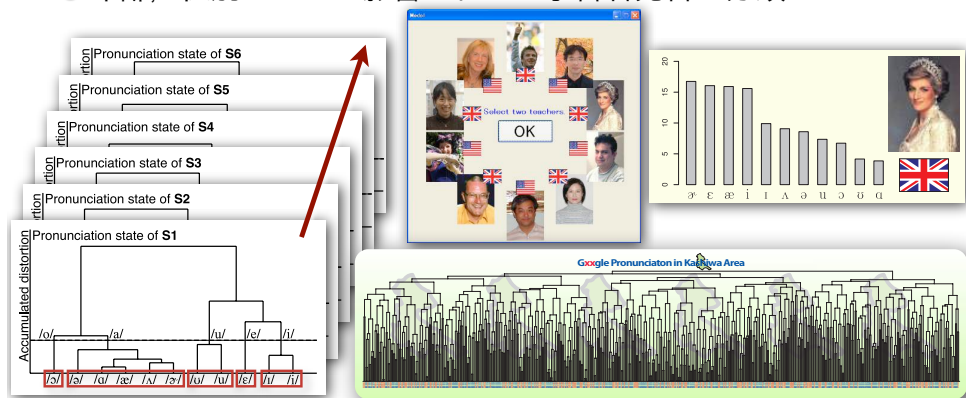
個々の母音ではなく、母音群の体系を評価対象



母音体系に見る発音習熟度

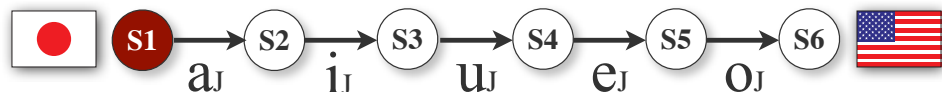
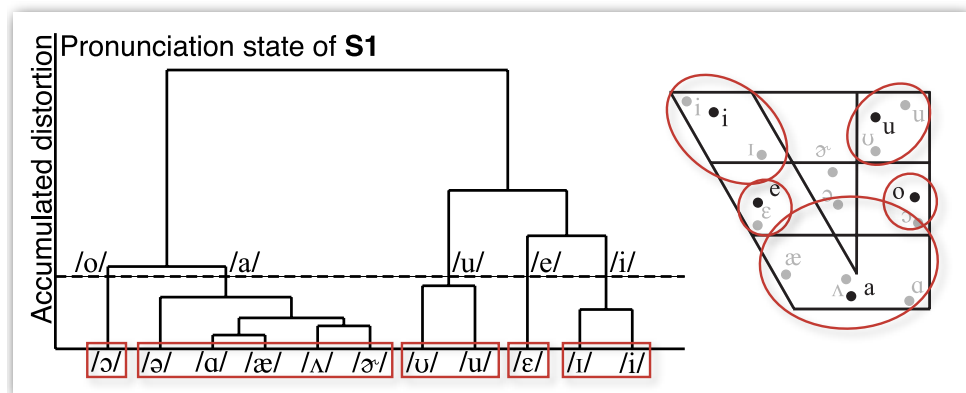
発音の構造表象に基づいて構築した技術・機能

- 学習者の母音体系がどのような状態にあるのかを記録する
- お手本の母音構造と比較して、矯正対象の母音を教示する
- 複数の教師からお手本としたい教師（相手）を選ぶ
- 年齢、性別などには影響されない学習者発音の分類



学習者の発音状態の記録とその遷移

S1からS6への遷移の様子の記事



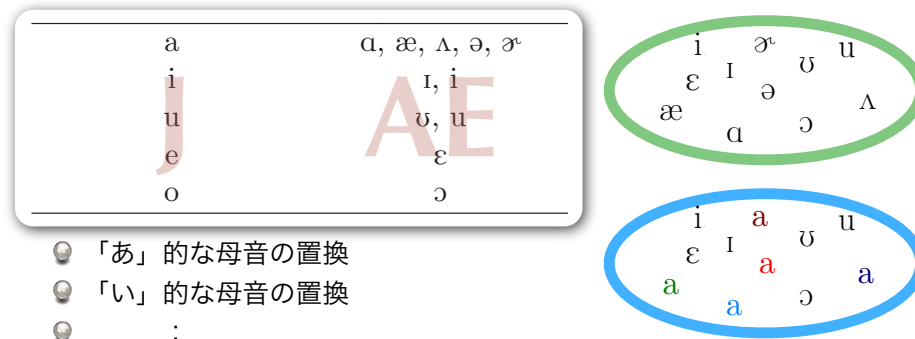
学習者の発音状態の記録とその遷移

分析対象の音声資料

- 英語劇経験者による英語&日本語単語音声
 - beat, bit, bet, bat, but, pot, bought, boot, put, bird, about x 1
 - ばと, びと, ぶと, べと, ぼと x 5

日本人学習者の母音発音の模擬的生成

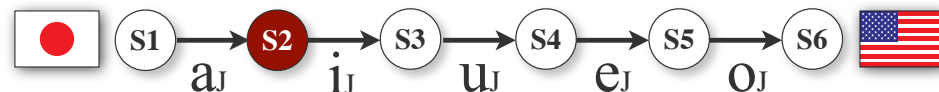
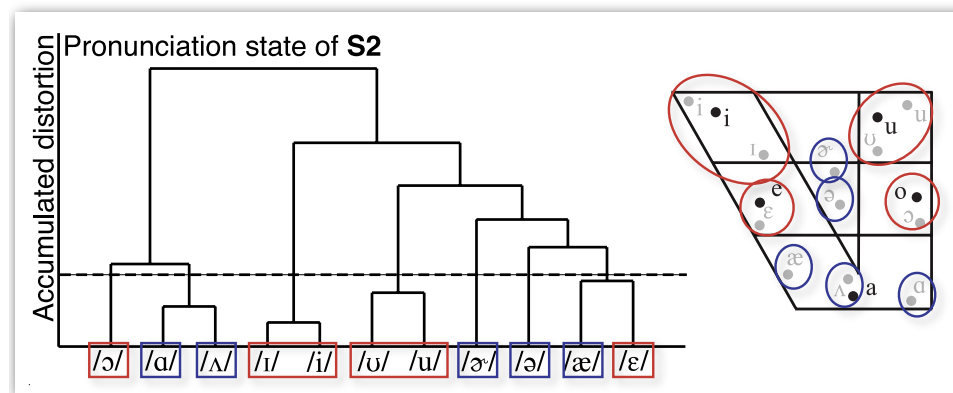
- 幾つかの米語母音と日本語母音の置き換え



- 「あ」的な母音の置換
- 「い」的な母音の置換
- :

学習者の発音状態の記録とその遷移

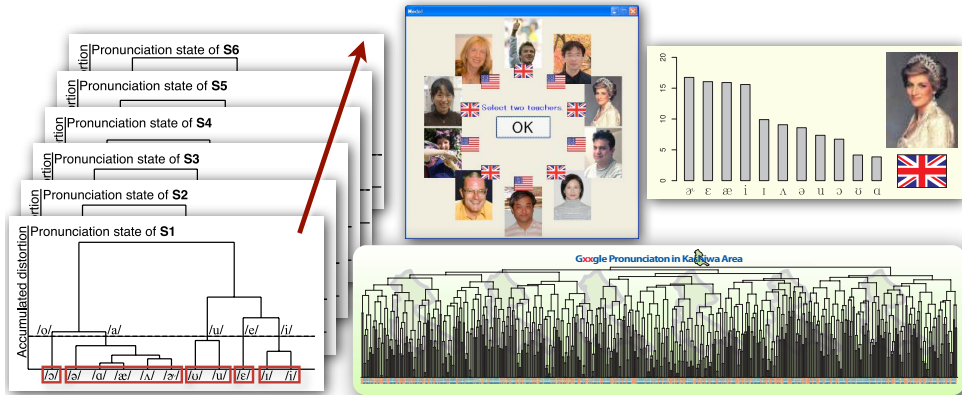
S1からS6への遷移の様子の記事



母音体系に見る発音習熟度

発音の構造表象に基づいて構築した技術・機能

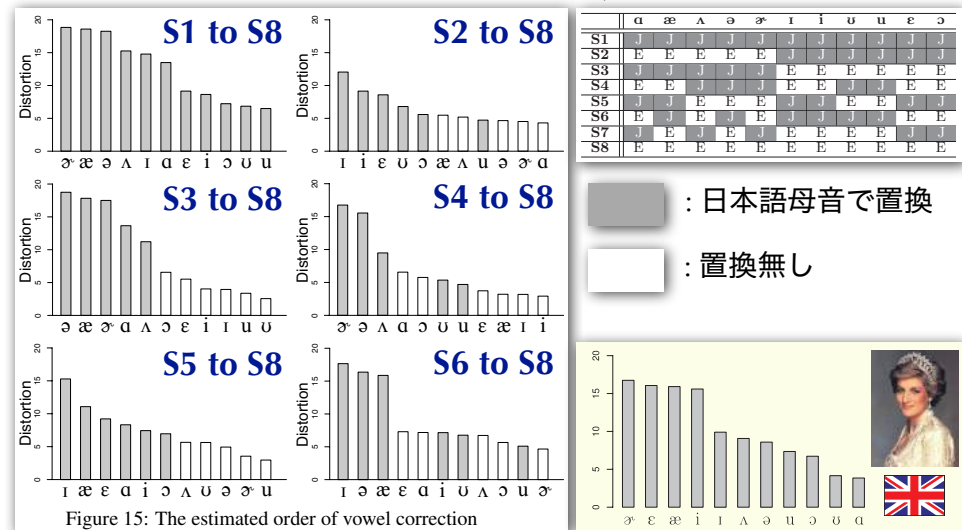
- 学習者の母音体系がどのような状態にあるのかを記録する
- お手本の母音構造と比較して、矯正対象の母音を教示する
- 複数の教師からお手本としたい教師（相手）を選ぶ
- 年齢、性別などには影響されない学習者発音の分類



優先的に矯正すべき母音の教示

S8 : 米語, S1-S7 : 日本語混入米語

- 教師と学習者の母音距離行列のみから、優先的矯正対象を推定



優先的に矯正すべき母音の教示

全体的構造歪みと局所的構造歪み

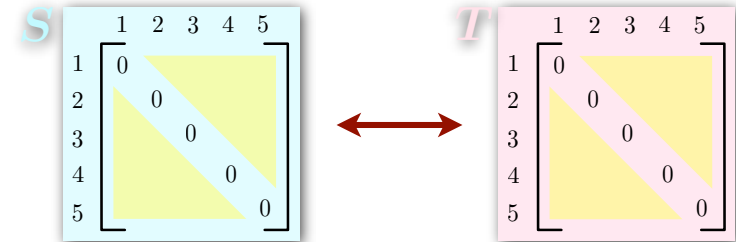
- 全体的構造歪み→発音習熟度の推定

$$TD(S, T) = \sqrt{\frac{1}{M} \sum_{i < j} (S_{ij} - T_{ij})^2}$$

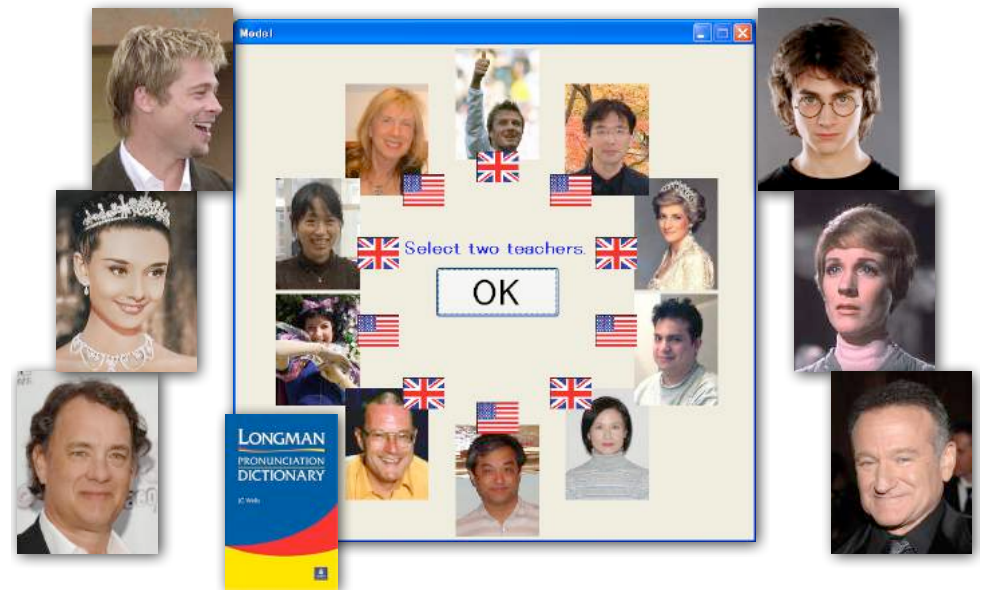
- 局所的構造歪み→矯正すべき発音部位の推定

$$LD(S, T, i) = \sum_{j=1}^M |S_{ij} - T_{ij}|$$

- LD = 大 ↔ 当該母音の矯正度 = 大



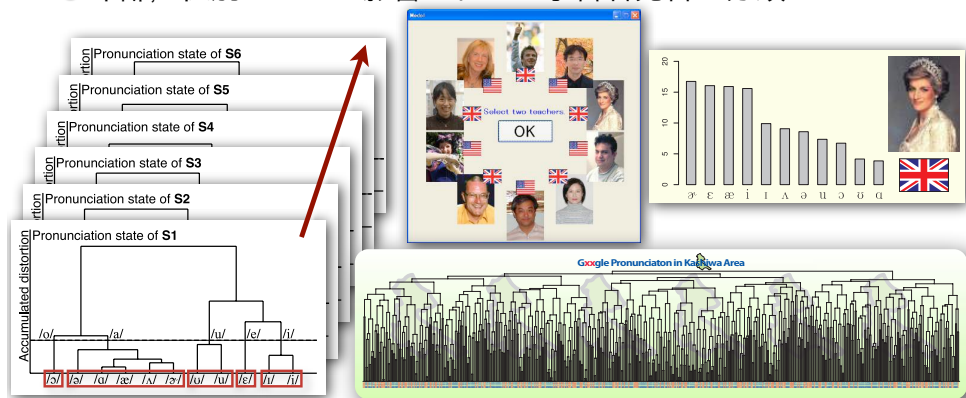
お手本となる相手を選ぶインタフェース



母音体系に見る発音習熟度

発音の構造表象に基づいて構築した技術・機能

- 学習者の母音体系がどのような状態にあるのかを記録する
- お手本の母音構造と比較して、矯正対象の母音を教示する
- 複数の教師からお手本としたい教師（相手）を選ぶ
- 年齢、性別などには影響されない学習者発音の分類



2通りの学習者分類

12人の帰国子女による96人の学習者シミュレーション

- 米語11母音+日本語5母音を含む単語音声収録
 - beat, bet, bat, but, × 1回
 - ばと, びと, ぶと, × 5回
- 米語母音の一部を日本語母音で置換
- どの米語母音を置換するのか? 8通りの置換パターン
 - 12人の帰国子女 × 8通りの置換パターン = 96人の学習者
 - [A~L] × [1~8] = 96個の11×11 距離行列

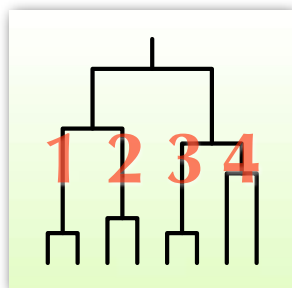
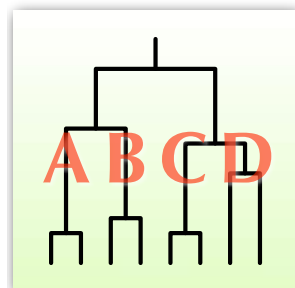
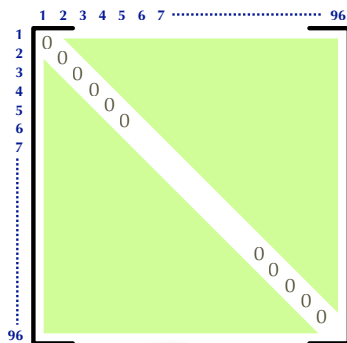
あ	a, æ, ʌ, ə, ɔ
い	i, i
う	u, u
え	ε
お	o

	a	æ	ʌ	ə	ɔ	i	ɪ	u	u	ε	o
S1	J	J	J	J	J	J	J	J	J	J	J
S2	E	E	E	E	E	J	J	J	J	J	J
S3	J	J	J	J	J	E	E	E	E	E	E
S4	E	E	J	J	J	E	E	J	J	E	E
S5	J	J	E	E	E	J	J	E	E	J	J
S6	E	J	E	J	E	J	J	J	J	E	E
S7	J	E	J	E	J	E	E	E	E	J	J
S8	E	E	E	E	E	E	E	E	E	E	E

2通りの学習者分類

96 x 96 の距離行列に基づく学習者分類

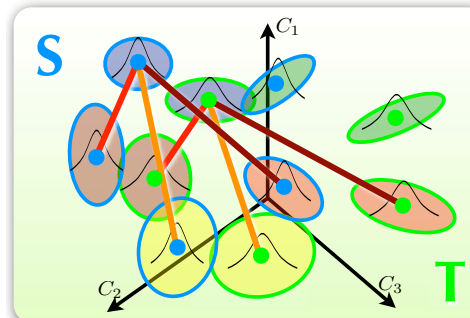
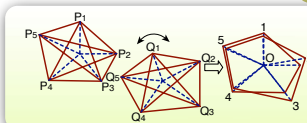
- 話者: A~L
- 発音: 1~8



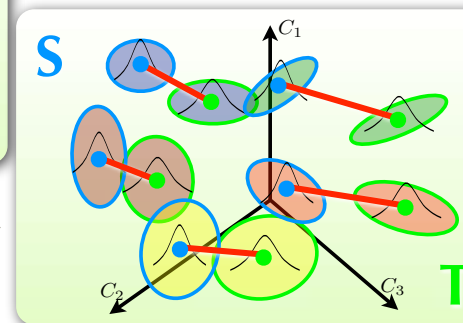
2通りの学習者分類

2通りの構造間距離

- コントラストの比較に基づく構造間距離
- 音の実体の比較に基づく構造間距離



$$\sqrt{\frac{1}{M} \sum_{i < j} (S_{ij} - T_{ij})^2}$$



$$\sqrt{\frac{1}{M} \sum_i BD(v_i^S, v_i^T)}$$

2通りの学習者分類

[illegible]

様々な英語の発音

1. British - Southern English - East London - Cockney

9. British - Scottish (unsure of specific type)

3. British - Southern English - Formal RP (Received Pronunciation)

音声模倣の二面性 ～音真似と? 真似～

- 👉 これを通して幼児は言語を獲得する。
- 👉 動物学的には非常に稀な行為。霊長類では人間だけ。
- 👉 他の動物では小鳥、クジラ、イルカくらいか。

- 九官鳥の音声模倣
 - 💡 車, ドア, 椅子, 犬, 猫, 音を真似る。人の声も音でしかない。
 - 💡 良い九官鳥を聞くと, 飼い主が分かる。
- 幼児の音声模倣
 - 💡 動物学的には奇妙な模倣行為
 - 💡 いくら良い子でも, 声から父親を割り出せずにお巡りさんは困る。



世界中の英語話者の分類に向けて

- 👤 英語は国際語
- 👤 地方訛り，母語訛り，様々な変容を受けるのは不可避免的
- 👤 訛りの最小単位は，国？地方？都市？町？ → 個人!!!
- 👤 英語利用者数＝15億人 → 15億通りの英語発音がある!!!

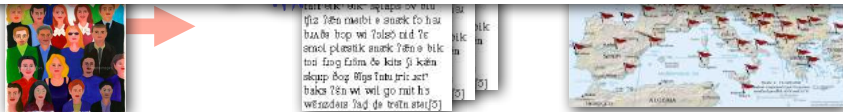
Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

世界中の英語話者の分類に向けて

世界諸英語 (World Englishes)

👤 英語は国際語

[pə'lis kal stələ azy xei tu brɪŋk dɪs fɪŋk wɪθ heɪ
frəm ðə stɔɪə sɪks spuːns ɒf fɪɛʃ snəʊ pɪs faɪy θɪk
sleɪps ʌv bluː tʃiːs ɛnt meɪbi ʌ snækˈ fɔɪ xei bɪlɔː bəp
wi ɔlzo nɪdˈ ə smɒl pləstɪk snēk æntˈ ə bɪk tɔɪ frɒg
fɔɪ ðə kɪts ʃɪ kæ skuːp dɪs θɪŋk ɪntu ˌtriː ʔret beks ɛnt
vi wɪl goʊ mɪt heɪ vɛnzdeɪ æt ðə tren steɪʃən]



もう少し具体的に・・・

TED speakers を学習者を中心に俯瞰するブラウザ

- 👤 TED speakers が SAA パラグラフを読み上げてくれれば
- 👤 その書き起こしを提供してくれれば

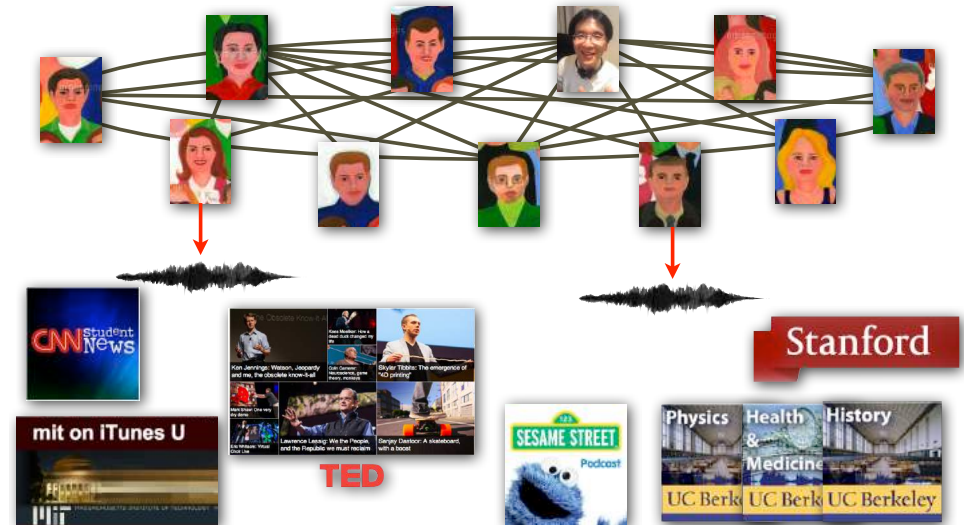
👤 任意の学習者一人を中心に据えて、彼らを俯瞰する



ネット上に存在する様々な訛った英語

世界諸英語ブラウザの構築

👤 特定パラグラフ読み上げを大勢の英語話者に依頼して・・・

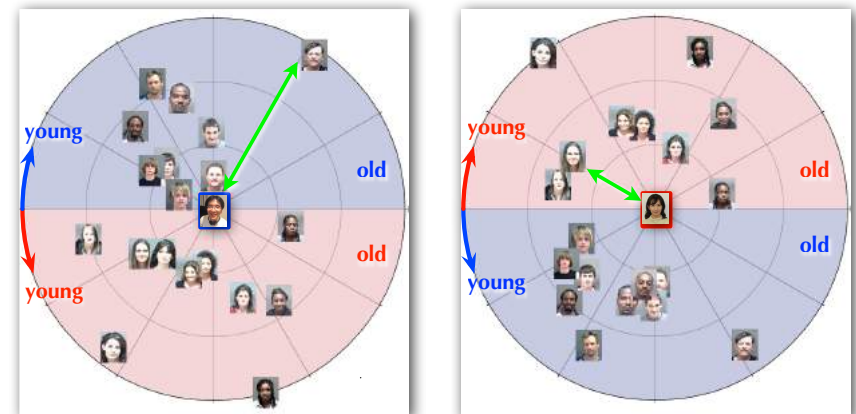


もう少し具体的に・・・

TED speakers を学習者を中心に俯瞰するブラウザ

- 👤 TED speakers が SAA パラグラフを読み上げてくれれば
- 👤 その IPA 書き起こしを提供してくれれば

👤 任意の学習者一人を中心に据えて、彼らを俯瞰する



こういう世界が作れたらいいな・・・

各自が自分の訛りをサーバーに登録する時代・・・

- CEFR的な言語パスポートの登録&管理



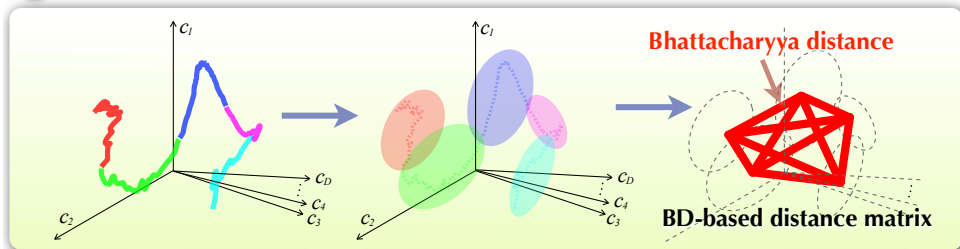
例えば、ホテルのフロントで・・・

例えば、レストランの注文時に・・・



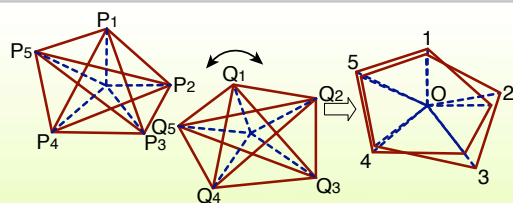
音声の構造的表象の工学的・実験的検証

f-div. に基づく一発声の構造化



2発声 (= 2距離行列) 間の音響照合

- 2距離行列間のユークリッド距離



- 回転：声道長差異
- シフト：マイク差異
- 話者適応・環境適応後のスコアが適応処理無しで算出
- 話者性を削除した音声表象

本講義の流れ

抽象化・モデル化と音声コミュニケーション技術

音声の物理学・基礎編

- 音声の生成からその物理的（音響的）側面まで

人間が受け取るメディア情報の多様性と不変性

- 視・聴・触・味・嗅、多様に変形しても同じだと分かる。

多様性の「そぎ落とし」と内在する不変性

- その数学的な回答とそれに基づく音声のモデル化

多様性に打ち勝つ音声技術の構築と応用

- 音声認識，外国語発音分析を例にとって

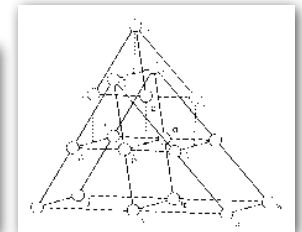
「そぎ落とし」が気付かせてくれる異分野との接点

- あれと，あれは，同じ原理なのか？
- 惑星とリンゴは同じである。音声とXXは同じである。

100年ほど前の出来事です

彼らは数学的に導いたのか？感覚しただけなのか？

- Ferdinand de Saussure (1857-1913)
 - Language is a system of differences.
 - The important thing in the word is not the sound alone but the phonic differences that make it possible to distinguish the word from the others.
- Roman Jakobson (1896-1982)
 - We have to put aside the accidental properties of individual sounds and substitute a general expression that is the common denominator of these variables.



形と形の比較 ～分子構造の比較～

DNA・RNA・タンパク質と音声の構造解析

立体構造を比較する(4)

— 距離に基づくRMSD (変換を求める必要がない)

$$RMSDd(A,B) = \sqrt{\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^m (R^{a_{ij}} - R^{b_{ij}})^2}$$



- m点の相互距離のそれぞれ(例えばi-j間)について、AとBでどのくらい一致しているかを比較指標とする。
- 利点: 変換を求める必要がない。理解しやすい。
- 欠点: 鏡像の関係にある構造を区別できない。

2.2 予測結果の評価基準

タンパク質の立体構造予測を行う際に、評価する基準は2つある。1つはエネルギー値、もう1つは構造の形である。構造が既知のタンパク質ならば、シミュレーション結果がどの程度その構造に似ているかが重要となる。2つの構造の差異を定量化するために用いられる量がRMSD(Root Mean Square Deviation)である。RMSDは2つの分子構造を重ね合わせて、対応する各原子のずれの二乗を平均したものの平方根で定義される。式(1)にRMSDの求め方を示す。

$$RMSD(A,B) = \sqrt{\frac{1}{N} \sum_{i=1}^N (a_i - b_i)^2} \quad (1)$$

RMSDの単位はÅで、値が小さいほど2つの構造がよく似ていることになる。

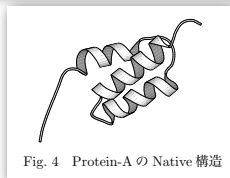
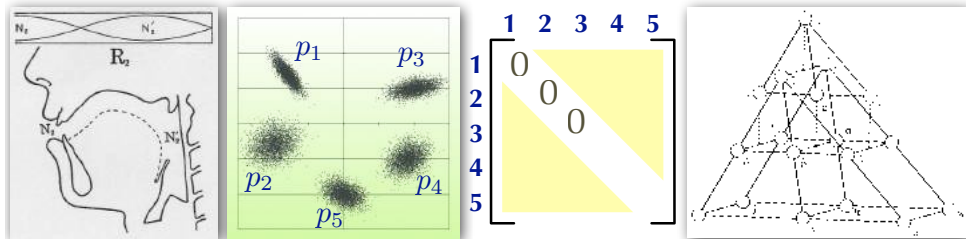


Fig. 4 Protein-A の Native 構造

再度、音声の構造的表象

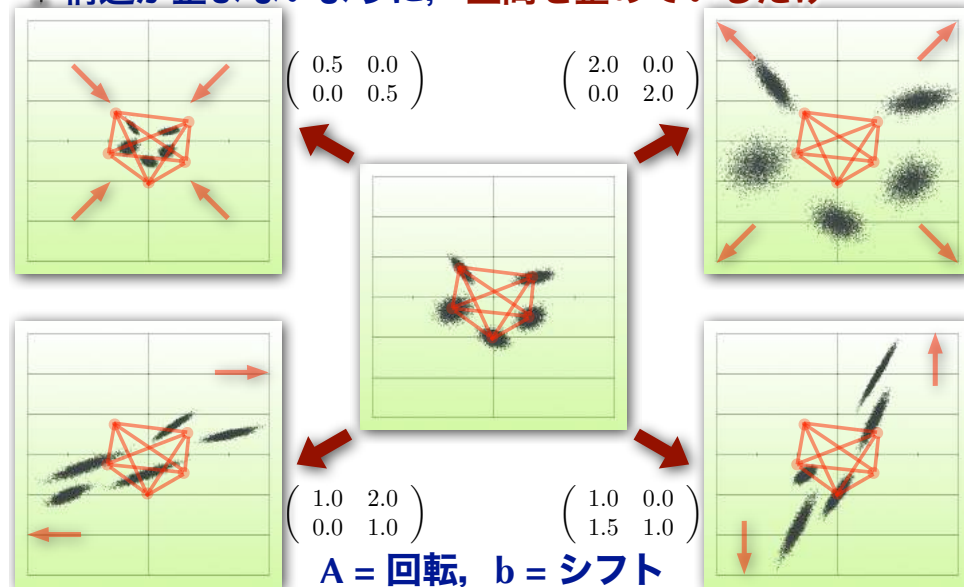
音声の構造的表象についておさらい

- 定常波として存在する音声を対象とする。
- 音声スペクトルは常に揺れている。これを分布で表現する。
- N個の音声事象はN個の分布となる (分布群)。
- 2事象間の関係を $\oint \sqrt{p_1(x) \cdot p_2(x)} dx = \cos \theta$ で定量化
- 2事象間の距離を $-\log(\cos \theta)$ で定量化
- 全ての2事象距離を計算し、NxNの距離行列を算出
- 変換・写像不変の情報の構造的表象



分布間距離群としての音声表象

構造が歪まないように、空間を歪めているだけ・・・



A = 回転, b = シフト

う～ん、似すぎている・・・?

価電子群の配置 = 分子構造の計算

$$|\phi_1(x)|^2 = p_1(x)$$

- 電子は原子核の周りに定常波として存在する。
- その定常波の2乗はその電子の確率密度関数 (分布) である。
- N個の価電子はN個の分布となる (分布群)。
- 2事象間の重なり具合を $\oint \phi_1(x) \phi_2(x) dx$ で定量化 (重なり積分)
- 全ての2事象間を計算し、行列化 (重なり行列)
- 重なり積分の関数として距離を算出。NxNの距離行列
- 変換・写像不変の情報の構造的表象?

