



東京大学 工学部 計数工学科 応用音響学

## D3 - ベクトル量子化

嵯峨山 茂樹 <[sagayama@hil.t.u-tokyo.ac.jp](mailto:sagayama@hil.t.u-tokyo.ac.jp)> 他

東京大学 工学部 計数工学科

資料所在 [http://hil.t.u-tokyo.ac.jp/~sagayama/applied\\_acoustics/](http://hil.t.u-tokyo.ac.jp/~sagayama/applied_acoustics/)

---

謝辞： システム情報第一研究室勉強会資料を部分利用 (亀岡担当分)

- クラスタリング
- $k$ -means アルゴリズム
- ベクトル量子化
- LBG アルゴリズム
- セグメンタル  $k$ -means アルゴリズム



# ベクトル量子化とは？ (量子化の種類)

- **スカラー量子化 (Scalar Quantization)**  
スカラー量の各サンプルを量子化する。(普通の量子化)
- **ベクトル量子化 (Vector Quantization)**  
ベクトル量の各データを量子化する。
  - 観測パターンを有限個の代表的なパターンの1つ(クラスの中心点)に置き換える。
  - 任意の観測パターンは、有限個のシンボルに変換される。
  - 1次元信号を短時間区間ごとに分析し、スペクトル等の特徴パラメータベクトルに変換して量子化する、などに用いられる。
- **マトリクス量子化 (Matrix Quantization)**  
2次元信号の小2次元平面のサンプル集合を一括して量子化する、など。



# ベクトル量子化の概念 (模式図)

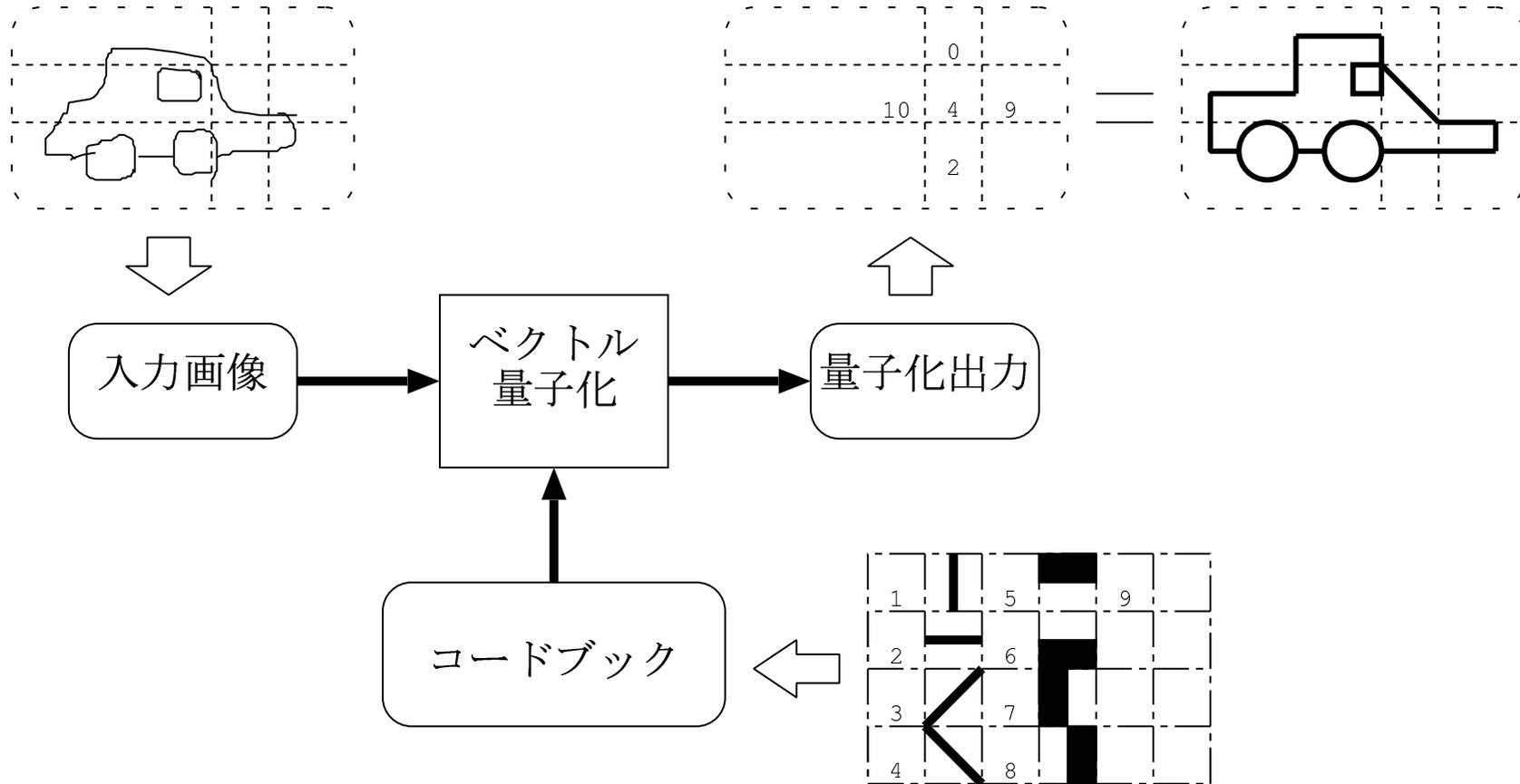


図1. ベクトル量子化器 概念図



# ベクトル量子化器 (Vector Quantizer)

## ■ ベクトル量子化符号帳 (codebook)

- ベクトル量子化符号 (code) — クラスタ番号
- ベクトル量子化符号語 (codeword) ベクトル — セントロイド

符号	符号語ベクトル
1	$m_1$
2	$m_2$
3	$m_3$
⋮	⋮
$n$	$m_n$

## ■ ベクトル量子化符号帳の学習 (training)

- 大量の学習データをクラスタリングしてセントロイドを得る。
- 学習規範 (criterion): 全ての学習データに関する量子化誤差の総和が最小



# ベクトル量子化

- 入力ベクトル( $x$ )を代表ベクトル( $y$ )で近似。

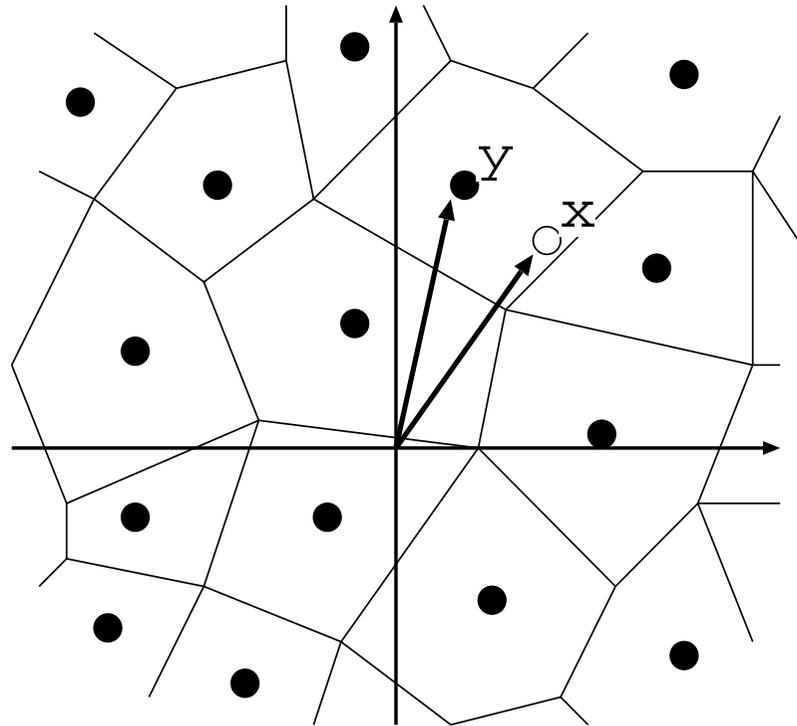


図2. ベクトル量子化



# ベクトル量子化の意義・用途

- 特徴ベクトルの時系列で表現される音声の符号化や認識などに広く利用される。
- 音声認識、話者認識に用いられる。
- 画像圧縮、画像符号化（jpeg等）にも用いられる。
- 構造的パターン認識法の入力として使用できる。
- 音声波形やパラメータの組には連続性や相関があり、それらをまとめて符号化することによって、独立に符号化するときよりも情報量を削減できる。
- 波形の時系列や音声の特徴パラメータに適用され、ベクトルの低ビット表現化が可能であることから、主にデータ量の削減化法として用いられる。
- スカラ量子化や直交変換符号化に比べて無駄が少ない。



# ベクトル量子化の概念図

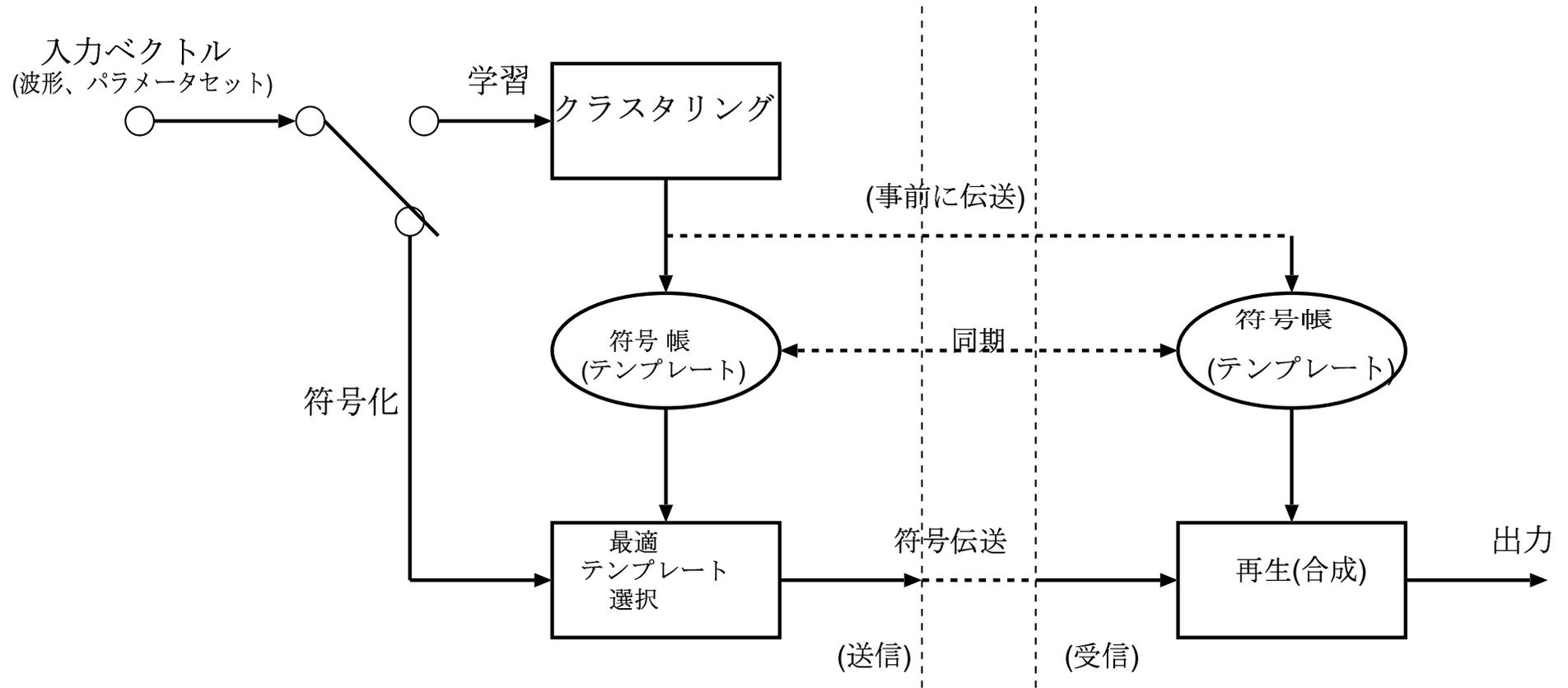


図3. ベクトル量子化の原理



# ベクトル量子化の原理

---

## ■ 学習 :

一定時間の波形あるいはパラメータの組 (ベクトル) をあらかじめ多数用意し、クラスタリングを適用し、代表パターン (セントロイド、テンプレート) の集合を生成。それぞれの代表パターンに符号を与え、符号と代表パターンとの対応を示す表 (符号帳) を蓄える。

## ■ 符号化 :

符号化すべき波形あるいはパラメータの組を、符号帳の各代表パターンと比較して、最も類似度の高いものを選び、その符号を与える。

## ■ 復号化 :

符号化に用いた同じ符号帳を参照して、伝送された符号に対応する代表パターンを再生する。



# コードブック作成のアルゴリズム

- 代表点はベクトルの統計的分布を反映するようにして選ばなければならない。
- 各サンプルを歪みが最小になるように代表的なベクトルに置き換える。
  - 代表点をコードワード (コードベクトル)
  - 代表点の集合をコードブック (符号帳)  
 $\{y_0, y_1, \dots, y_{N-1}\}$  (サイズ  $N$  のとき)
- コードブック作成の為に任意の  $k$  個のベクトル (訓練サンプル)  $\{x_1, x_2, \dots, x_k\}$  のセントロイド (中心) ベクトル  $x$  を次式で定義する<sup>1</sup>。

$$x = C(x_1, x_2, \dots, x_k) = \operatorname{argmin}_x \frac{1}{n} \sum_{i=0}^{n-1} d(x, x_i) \quad (1)$$

<sup>1</sup>  $d(x, x_i)$  は、 $x$  と  $x_i$  とのベクトル間距離を表す。  $\operatorname{argmin}_x f(x)$  は、 $f(x)$  を最小にする  $x$  を意味する。



# LBG アルゴリズムとは?

---

- Y. Linde, A. Buzo and R. M. Gray が提案
- 最もよく用いられるベクトル量子化器の設計アルゴリズム
- 繰り返し2分割と、 $k$ -means アルゴリズムの組合せ



# LBG アルゴリズム

- コードブック作成のため, ベクトルセントロイドと分割部分集合を繰り返し計算により更新していく手法

- 手法の概要

1. コードブックサイズを1とする

2. コードブックのベクトルを近接する二つのベクトルに分ける

- (a) コードブックの初期集合の設定

- (b) サンプルベクトルを部分集合に分割

- (c) 各部分集合内のベクトルセントロイドを求める

- (d) 平均ひずみが閾値より小さくなるまで (b) と (c) を繰り返す

3. コードブックサイズが所望の個数になれば終了.

それ以外なら 2. に戻る

- 手法の特徴

- $k$ -means アルゴリズムと類似

- 中間分割数でのセントロイドの値が得られる (分割数は2のべき乗)



# コードブック作成アルゴリズム

## 1. 初期化

訓練サンプル集合  $\{\mathbf{x}_j : j = 0, 1, \dots, n - 1\}$

コードブックサイズ  $N$

コードブック初期集合  $A_0^{(N)} = \{\mathbf{y}_0^{(0)}, \mathbf{y}_1^{(0)}, \dots, \mathbf{y}_{N-1}^{(0)}\}$ ,  $m = 0$

## 2. コードブックを固定して分割を最適化。

$\{\mathbf{x}_j\}$  をコードブック  $A_m^{(N)}$  によって、 $N$  個の部分集合  $\{S_i : i = 0, 1, \dots, N - 1\}$  に分割する。全ての  $t$  について  $d(\mathbf{x}_j, \mathbf{y}_i^{(m)}) < d(\mathbf{x}_j, \mathbf{y}_t^{(m)})$  なら、 $\mathbf{x}_j \in S_i$

## 3. 入力ベクトルと代表ベクトルの歪み（ユークリッド距離の2乗）が十分小さくなれば、停止し、コードブック $A_m^{(N)}$ を出力。それ以外は、4.へ。

## 4. 分割を固定してコードブックを最適化（重心に最適化）

コードブック  $A_{m+1}^{(N)} = \{\mathbf{y}_0^{(m+1)}, \mathbf{y}_1^{(m+1)}, \dots, \mathbf{y}_{N-1}^{(m+1)}\}$  を求め、 $\mathbf{y}_i^{(m+1)} = C(\{S_i\})$ ,  $m := m + 1$  (式(1)参照) として、2.に戻る。



# LBG アルゴリズムの特徴

---

- 符号化レートと量子化に伴う平均ひずみとのトレードオフ。
- ベクトル量子化時の量子化誤差（歪み）はテンプレートの個数だけでなく、選び方にも依存。
- 母集団の分布の偏りを利用して、ベクトル量子化時をサンプル全体として最小化。
- 本質的にクラスタリング手法と変わらない。  
 $k$ -means アルゴリズムは、クラスタ基準がパターン分類能力最大であり、LBG アルゴリズムは、クラスタ基準が歪み最小。



# セグメンタル $k$ -means アルゴリズム

## 概要

連続時系列信号（音声、字を書く時の筆圧等）を、クラスタリングする技法。使用例としては、HMMの各状態の初期割り当てを決定するときを使用すると考えられる。

## アルゴリズム

初期に適当な分割を行ってから以下の行程を繰り返す。

1. 各クラスにおける代表値を、そのクラス内の各点からの距離の総和が小さいものにする。
2. 各点と代表値を比較して、距離の総和が小さくなるようにクラスの大きさを設定する。但し、クラスの大きさを変化させる時に、必ず連続的になるようにする。

代表値、クラスの大きさともが変動しなくなった時点で終了する。