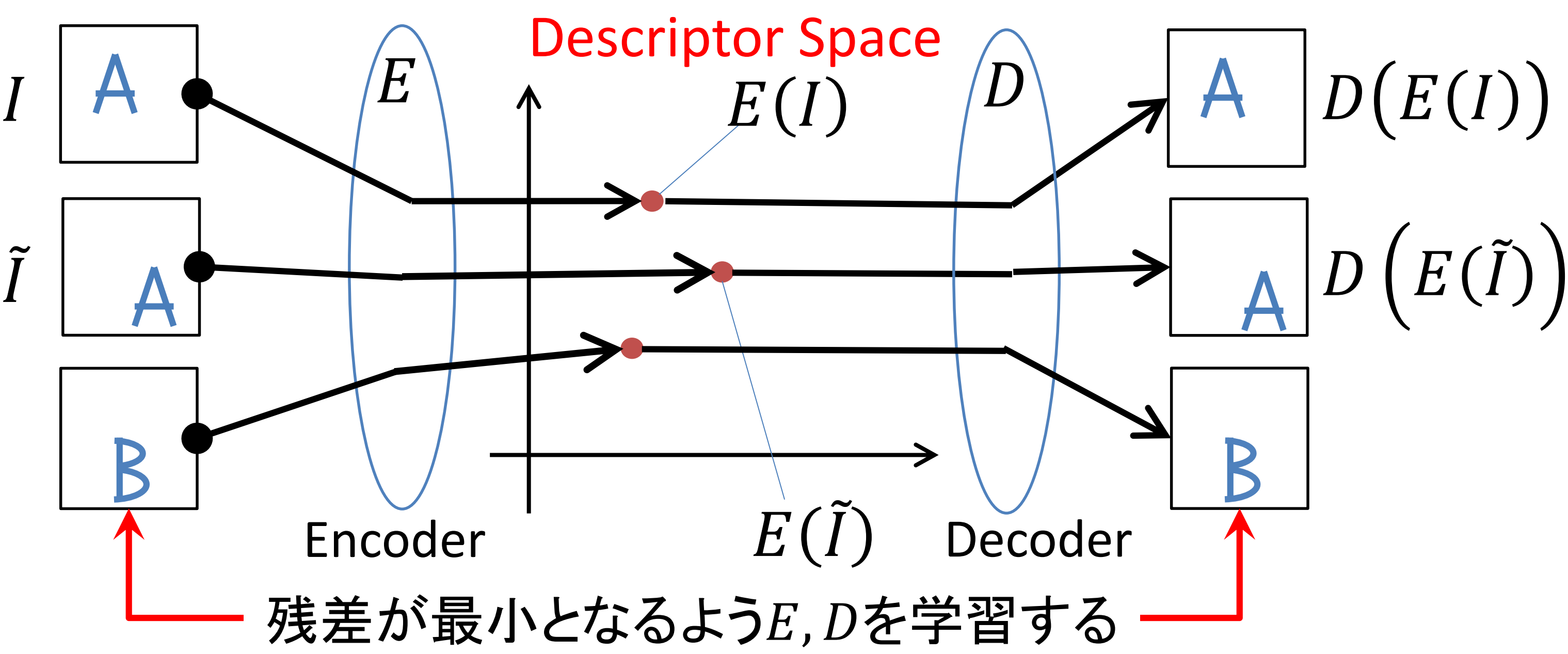


R Auto-encoder for generating a transform invariant descriptor and transform parameters

Tadashi Matsuo, Nobutaka Shimada (Ritsumeikan Univ.) E-mail: matsuo@i.ci.ritsumeai.ac.jp

1. Auto-encoder

多くの入力について再生性を担保しつつ次元削減する方法

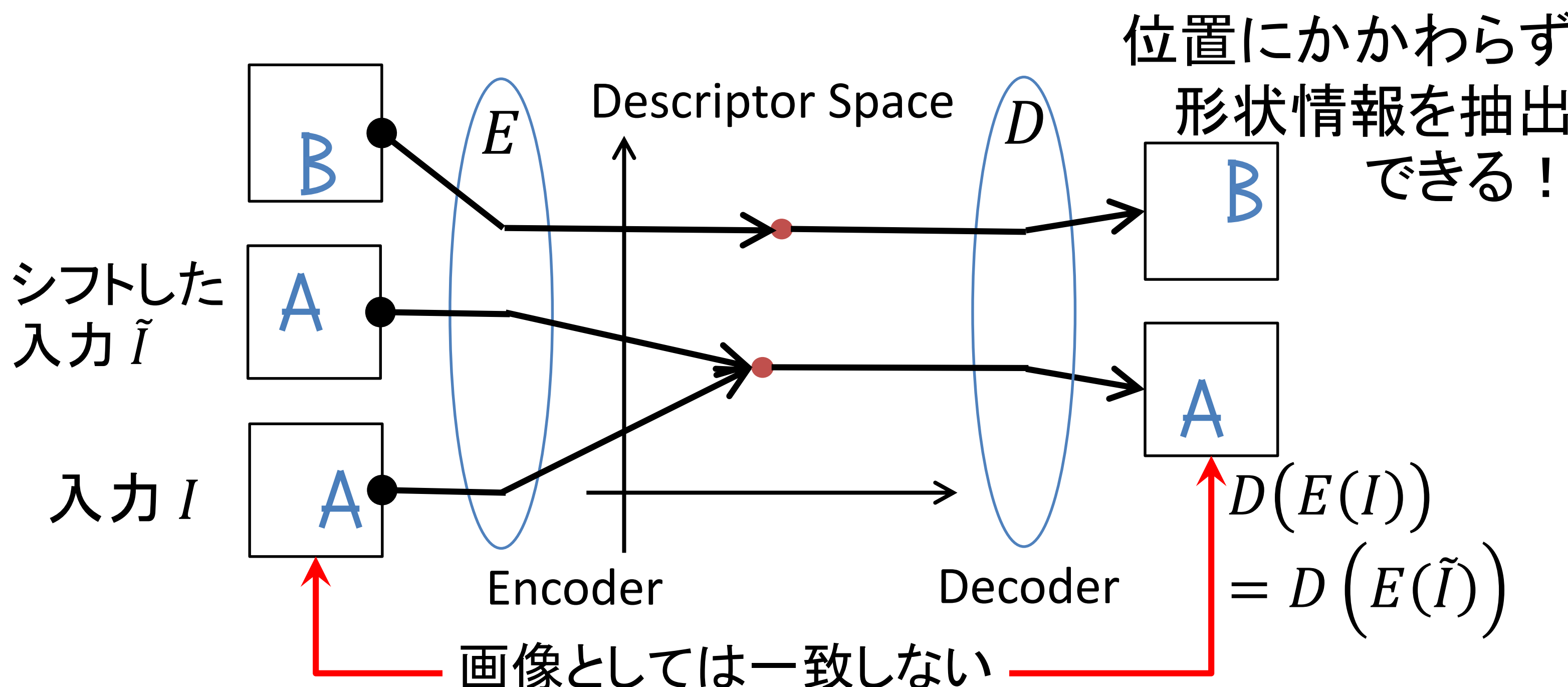


事前に教師ラベルを与えることなく、典型的なサンプルを精度良く表現するdescriptorが得られる

しかし同一形状でも位置が違くとdescriptorが異なる

2. シフト不変 auto-encoder

encoderがシフトに関して不変かつdecoderで形状が再現されるようにすれば、形状自体を表すdescriptorが得られる



通常目的関数 $\sum_I \|D(E(I)) - I\|_{L2}^2$ では学習できない

3. 目的関数

入力Iにどのシフト T_θ を適用しても復元結果は同じであるべきという制約

復元画像が入力Iを変換した $T_\theta(I)$ のいずれかと一致するべきという制約

$$\sum_I \lambda_{inv} \left\{ \sum_\theta \|D(E(I)) - D(E(T_\theta(I)))\|_{L2}^2 \right\} + \lambda_{res} \min_\theta \|D(E(I)) - T_\theta(I)\|_{L2}^2 + \lambda_{spa} \frac{\|E(I)\|_{L1}^2}{\|E(I)\|_{L2}^2}$$

考慮する変換パラメータ θ の個数が増えると総和計算、最小値探索の負荷が大きくなる
そこで、変換パラメータ θ の自由度増大による計算量の増加を抑える新しい目的関数を提案する

提案目的関数

$$\sum_I \lambda_{inv} \int_\theta \|D(E(I)) - D(E(T_\theta(I)))\|_{L2}^2 d\theta + \lambda_{res} \int_\theta \|D(E(I)) - T_\theta(I)\|_{L2}^2 W(I, \theta) d\theta + \lambda_{spa} \frac{\|E(I)\|_{L1}^2}{\|E(I)\|_{L2}^2}$$

各入力についての変換パラメータ空間上の重み関数 $W(I, \theta)$ を導入し目的関数を E, D, W について最小化する

- 連続パラメータを持つ変換にも適用できる。
- 入力を変換不変成分 $E(I)$ と変換パラメータの分布 $W(I, \theta)$ の対に分解できるようになる。
- 再現性項の変換パラメータに関する最小化に勾配法を活用できる。変換パラメータ θ についての微分計算は一般に難しいが重み関数のパラメータについての微分計算は可能。
- Monte Carlo法を用いることで計算量を削減できる。特に再現性項の計算は重みの大きい部分に着目することで少ない計算量で近似できる。

$$W(I, \theta) = \frac{1}{(2\pi)^{D/2} |\frac{1}{2}\Sigma_I|^{1/2}} \exp\left(-\frac{1}{2}(\theta - \mu_I)^T (\frac{1}{2}\Sigma_I)^{-1} (\theta - \mu_I)\right)$$

(Dは変換パラメータの次元)

変換パラメータ空間上のGauss関数を採用している
 $W(I, \theta)$ のパラメータ μ_I と Σ_I は I を入力とする neural network で生成する

Monte Carlo法による再現性項の近似

$$\int_\theta \|D(E(I)) - T_\theta(I)\|_{L2}^2 W(I, \theta) d\theta = \frac{D}{2^2} \int_\theta \|D(E(I)) - T_\theta(I)\|_{L2}^2 w(I, \theta) p_I(\theta) d\theta \approx \frac{D}{N} \sum_{\theta_n \sim N_D(\mu_I, \Sigma_I)} \|D(E(I)) - T_{\theta_n}(I)\|_{L2}^2 w(I, \theta_n)$$

4. 実験結果

MNISTの訓練画像をシフトして生成した画像群でシフト変換に関するauto-encoderを訓練

同一モデルを各方法で50000回更新した際に要した時間



従来法	提案法
16.9時間	5.9時間

従来法では変換パラメータの次元に関して指数的な計算量が必要となる。提案法はMonte Carlo法を用いることにより指数的な計算量を必要としない。

拡大縮小や回転などの変換の組み合わせへの適用