

# q-SNE: Visualizing Data using q-Gaussian Distributed Stochastic Neighbor Embedding

Motoshi Abe<sup>1</sup>, Junichi Miyao<sup>1</sup>, Takio Kurita<sup>1</sup>

<sup>1</sup>Hiroshima University in Japan



### Abstract

- We present a novel technique called a q-Gaussian distributed stochastic neighbor embedding (q-SNE) to improve t-distributed stochastic neighbor embedding (t-SNE). The q-SNE uses q-Gaussian distribution instead of t-distribution of t-SNE.
- Since the q-Gaussian distribution includes Gaussian distribution, t-distribution, and others, so the q-SNE leads to more powerful and flexible visualization on 2 or 3-dimensional mapping than the t-SNE.

## Results of q-SNE

We show the embedding of MNIST and COIL-20 dataset in Fig.3 and in Fig4 respectively. If **q** becomes larger, the ebedding becomes more clustered.





• Code on Python is here (https://github.com/i13abe/q\_SNE).

#### q-SNE

We propose an extension of the t-SNE by using q-Gaussian distribution instead of t-distribution of the t-SNE. We call the proposed method q-Gaussian distributed stochastic neighbor embedding (q-SNE). The q-Gaussian distribution is derived by the maximization of the Tsallis entropy under appropriate constraints and is a generalization of the Gaussian distribution. Let s be a 1-dimensional observation. The q-Gaussian distribution for the observation s is defined as

$$P_{q}(s;\mu,\sigma^{2}) = \frac{1}{Z_{q}} \left( 1 + \frac{q-1(s-\mu)^{2}}{3-q} \right)^{-\frac{1}{q-1}}$$
(1)

where  $\mu$  and  $\sigma$  are the mean and the variance, respectively. The normalization factor  $Z_q$  is given by

$$Z_{q} = \begin{cases} \sqrt{\frac{3-q}{q-1}} \text{Beta}\left(\frac{3-q}{2(q-1)}, \frac{1}{2}\right)\sigma, & 1 \le q < 3\\ \sqrt{\frac{3-q}{1-q}} \text{Beta}\left(\frac{2-q}{1-q}, \frac{1}{2}\right)\sigma, & q < 1 \end{cases}$$
(2)

where Beta() is the beta function. We show the graph of Gaussian distribution, t-distribution, and q-Gaussian distribution with different phyperparameter q in Fig.1. We show the architecture of q-SNE in Fig.2.



Figure 1: The graph of Gaussian distribution, t-distribution, and q-Gaussian distribution.

#### Figure 3: The embedding of t-SNE and q-SNE on MNIST.





Figure 4: The embedding of t-SNE and q-SNE on COIL-20.

presentation of ICPR2020