Look at the Driver, Look at the Road: No Distraction! No Accident!

Cheng Guan

July 8, 2018

1. Archetypes and Archetype Hull

According to this paper, I know that archetype was originally called psychology [2] and was used to describe a very typical and generally understood paradigm of a set of objects. Prototype analysis aims to find a compact "pure type", *i.e.*, the so-called archetypes, Thus, the typical pattern of this set of objects is covered in the archetypes, while the other objects are just a combination of simulations or archetypes.

1.1. Concept of Archetypes

The concept of archetypes actually is not far away from us, which exists in a variety of subjects including literature, philosophy, and psychology marketing [3] and statistics[1]. The idea of archetypes is recently introduced into pattern recognition and informatics [6, 4]. Understanding the concept of archetypes in face recognition is intuitive: people can easily remember someone with a very distinctive facial appearance; some people are thought of being very similar to a few distinctive faces. Such phenomena serve as the evidence of identifying archetypes and correlating unknown faces to known archetypes in a recognition process.

1.2. Archetype Seeking

The time efficiency of finding archetypes is an important concern. It is known that for an input data set χ of n points with d dimensions, the time complexity for computing a convex hull enclosing χ is as high as [5, 6]. For a set of tens of thousands of images with high-dimensional descriptors, exactly solving this convex hull problem quickly becomes computationally intractable. To this end, a few algorithms have been designed to achieve approximate solutions, of which the simplex volume maximization algorithm [6] can provide a good approximate solution in a linear time complexity $O\left(n\right)$.

With a mild assumption that all simplex vertices are equidistant, the Cayley-Menger determinant gives a simplified volume formula as Eq.1

$$\begin{aligned} V_{ol}(S)^2 =& \frac{a^{2(m-1)}}{2^{m-1}((m-1)!)^2}\left(\frac{2}{a^4}\sum_{j=1}^{m-1}\sum_{j'=j+1}^{m-1} d_{j,m}^2 d_{j',m}^2\right. \\ & \left. + \frac{2}{a^2}\sum_{j=1}^{m-1} d_{j,m}^2 - \frac{m-2}{a^4}\sum_{j=1}^{m-1} d_{j,m}^4 - (m-2)\right) \end{aligned} \tag{1}$$

2. Archetype Hull Ranking

In this section, the authors address the problem that they raised in Section 3: how to correlate an input sample with archetypes. Note that this sample may be an unseen sample outside the set of available training samples, and that these archetypes can be stored in memory for real-time computations.

Since the archetype hull encloses most training samples in χ , we can project any input sample $x \in \mathbb{R}^d$ onto S, thereby obtaining a new (lossy) representation of x in terms of the stored archetypes in υ. Formally, they reconstruct the input x using a convex combination of the archetypes in υ as Eq.2

$$\begin{aligned} &\min_{z(x)\in\mathbb{R}^m} \left\|x - Uz\left(x\right)\right\| \\ &z\left(x\right) \geq 0, \quad \mathbf{1}^\top \mathbf{z}\left(\mathbf{x}\right) = \mathbf{1} \end{aligned} \tag{2}$$

References

[1] M. J. Eugster and F. Leisch. Weighted and robust archetypal analysis. *Computational Statistics & Data Analysis*, 2011. 1

[2] C. G. Jung. *The archetypes and the collective unconscious*. Routledge, 2014. 1

[3] S. Li, P. Wang, J. Louviere, and R. Carson. Archetypal analysis: A new way to segment markets based on extreme individuals. In *ANZMAC*, 2003. 1

[4] C. Seiler and K. Wohlrabe. Archetypal scientists. *Journal of Informetrics*, 2013. 1

[5] C. Thurau. Nearest archetype hull methods for large-scale data classification. In *CVPR*, 2010. 1

[6] C. Thurau, K. Kersting, and C. Bauckhage. Yes we can: Simplex volume maximization for descriptive web-scale matrix factorization. In *CIKM*, 2010. 1