SSD Clustering

Manny Ko

February 13, 2022

Abstract

1 Introduction

The premise of this investigation is to use a form of clustering to drive the training of deep learning models. With clusters, we hope to promote small intra-cluster variability and reduce the need for large capacity models by training a model for each cluster. With that a corresponding reducing in the amount of training data.

2 Clustering

There are two distinct set of approaches to doing clustering without training:

1. cluster the feature vectors of a pre-trained model
2. cluster the input images using:
   (a) "2nd-order" region statistics: e.g. region covariance 2.1 and Sigma-set 2.1.1
   (b) some form of non-Local Means (NLM)

There the algorithm to perform the clustering also deserve some investigation. K-means is certain a starting point but it is very sensitive to initialization. Next, the question of producing a quasi balanced cluster should also be considered.

If we are using a NLM approach, naive approach will make k passes for k clusters while testing for nearest center. At least we need a fast lookup structure so that the nearest center can be retrieved in $O(logn)$ or $O(1)$ time. Or use the very nice SigmaSet 2.1.1 or [KH10]

2.1 Region Covariance

Region covariance can be computed very efficiently using 'integral images/-sumarea table" [PT06, TPM06]. Let $F(x, y) = \phi(I, x, y)$ be the $W \times H \times d$ dimensional feature image extracted from I, where ϕ can be any mapping such as intensity, color, gradients, filter response etc.

The region covariance $\mathbf{C}_R$ is a $d \times d$ matrix of the feature points:

$$\mathbf{C}_R = \frac{1}{n-1} \sum_{k=1}^{n} (\mathbf{z}_k - \mu)(\mathbf{z}_k - \mu)^T \tag{1}$$

where μ is the mean. $\mathbf{C}_R$ is $d \times d$ instead of $n \times d$ if we are using the raw features. Also, RC does not have any information regarding the ordering and the number of points. This implies a certain scale and rotation invariance.

Covariance matrices do not lie on Euclidean space, a distance metric involving generalized eigenvalues follows from the Lie group structure of positive definite matrices:

$$\rho(\mathbf{C}_1, \mathbf{C}_2) = \sqrt{\sum_{k=1}^{n} ln^2 \lambda_i(\mathbf{C}_1, \mathbf{C}_2)} \tag{2}$$

where $\lambda_i(\mathbf{C}_1, \mathbf{C}_2)$ are the generalized eigenvalues of $\mathbf{C}_1$ and $\mathbf{C}_2$ computed from $\lambda_i \mathbf{C}_1 \mathbf{x}_i - \lambda_i \mathbf{C}_2 \mathbf{x}_i$ and $\mathbf{x}_i \neq 0$ are the generalized eigenvectors.

Using covariance improves computer detection and tracking of humans

2.1.1 Sigma set

[HCS+09] proposes a novel 2nd-order statistics based region descriptor, named "Sigma Set", in the form of a small set of vectors, which can be uniquely constructed through Cholesky decomposition on the covariance matrix. This is basically an optimized form of region covariance using sparsity.

for any matrix A that satisfies $C_R = AA^T$, the set of columns of A has the same 2nd order statistics as R. More specifically, we can construct the Sigma Set descriptor for region R through Cholesky decomposition

2.2 Kwatra2010

"Fast Covariance Computation and Dimensionality Reduction for Sub-Window Features in Images"

2.2.1 Faulkner2015

[FSS+15] "A Study of the Region Covariance Descriptor: Impact of Feature Selection and Image Transformations"

2.3 K-means

2.3.1 Balanced K-means

[MF14] in k-means assignment phase, the algorithm solves the assignment problem by Hungarian algorithm with time complexity $O(n^3)$.

2.3.2 Balanced K-means & min-cut

[CNMY14]

2.4 Non-Local Means

2.4.1 Qian2013

[QY13] nonlocal similarity and spectral-spatial structure of hyperspectral imagery into sparse representation. Non-locality means the self-similarity of image, by which a whole image can be partitioned into some groups containing similar patches. The similar patches in each group are sparsely represented with a shared subset of atoms in a dictionary making true signal and noise more easily separated.

2.4.2 Fu2017

[FLSS17]

3 SSD

4 Conclusion

5 Remarks

Acknowledgments. Finally, thank you to my family and friends for the support during this report.

References

[CNMY14] Xiaojun Chang, Feiping Nie, Zhigang Ma, and Yi Yang. Balanced k-Means and Min-Cut Clustering. In *ICCV '14*, 2014.

[FLSS17] Ying Fu, Antony Lam, Imari Sato, and Yoichi Sato. Adaptive Spatial-Spectral Dictionary Learning for Hyperspectral Image Restoration. *Int J of Computer Vision*, 122(2):228–245, 2017.

[FSS+15] Hayden Faulkner, Ergnoor Shehu, Zygmunt L. Szpak, Wojciech Chojnacki, Jules R. Tapamo, Anthony Dick, and Anton Van Den Hengel. A Study of the Region Covariance Descriptor: Impact of Feature Selection and Image Transformations. In *DICTA '15*, page 8, 2015.

[HCS+09] Xiaopeng Hong, Hong Chang, Shiguang Shan, Xilin Chen, and Wen Gao. Sigma set: A small second order statistical region descriptor. In *CVPR Workshops '09*, pages 1802–1809. Ieee, jun 2009.

[KH10] Vivek Kwatra and Mei Han. Fast covariance computation and dimensionality reduction for sub-window features in images. In *ECCV 2010*, volume 94043, 2010.

[MF14] Mikko I. Malinen and Pasi Fränti. Balanced k-means for clustering. In *S+SSPR '14*, volume 8621 LNCS, pages 32–41, 2014.

[PT06] Fatih Porikli and Oncel Tuzel. Fast construction of covariance matrices for arbitrary size image windows. *ICIP '06*, pages 1581–1584, 2006.

[QY13] Yuntao Qian and Minchao Ye. Hyperspectral imagery restoration using nonlocal spectral-spatial structured sparse representation with noise estimation. *IEEE J of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2):499–515, 2013.

[TPM06] Oncel Tuzel, Fatih Porikli, and Peter Meer. Region covariance: A fast descriptor for detection and classification. In *ECCV '06*, pages 589–600, 2006.