AAAMS : Anisotropic Agglomerative Adaptive Mean-Shift

Rahul Sawhney[1]
rahul.sawhney@gatech.edu
Henrik I. Christensen[1]
hic@gatech.edu
Gary R. Bradski[2]
gbradski@magicleap.com

[1] Institute of Robotics & Intelligent Machines, Georgia Tech
[2] Magic Leap Inc.

Mean Shift today, is widely used for mode detection and clustering. The technique though, is challenged in practice due to assumptions of isotropicity and homoscedasticity. Isotropic/scalar bandwidths tend to smooth anisotropic patterns and affect partition boundaries, while homoscedastic / global bandwidths are inappropriate when clusters (or modes) at different scales need to be identified.

We present an adaptive Mean Shift methodology that allows for anisotropic clustering, through unsupervised local bandwidth selection. The bandwidth matrices evolve naturally, adapting locally through agglomeration, and in turn guiding further agglomeration. The online methodology is practical for low-dimensional feature spaces, preserving better detail and clustering salience. Additionally, conventional Mean Shift either critically depends on a per instance choice of bandwidth, or relies on offline methods which are inflexible and/or again data instance specific. The presented approach, due to its adaptive design, also alleviates this issue - with a default form performing generally well. The methodology though, allows for effective tuning of results.

In the proposed approach, clusters arise on the fly, as a consequence of agglomeration of extant clusters. Local bandwidths which evolve anisotropically every iteration, are associated with each cluster; by design, all members of a cluster converge to the same local mode. By evolving as function of a cluster's aggregated trajectory points, these bandwidths are able to adapt to the underlying mode structure (shape, scale, orientation) - and in turn, guide future cluster trajectory and agglomeration. This results in robust mode detection and with increased partition saliency (*Figs. 1, 2(a)*). The supplementary presents a convergence proof when anisotropic bandwidths vary between Mean shift iterations, as is the case here.

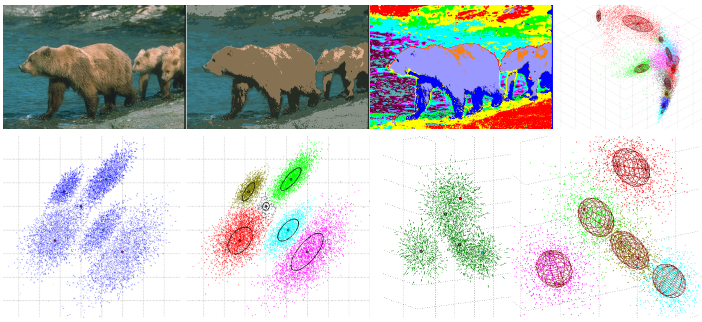


Figure 1: Single domain clustering examples over color data (top row with segment and label images are shown in middle, 11 clusters), and simulated gaussian mixtures (second row) in 2D & 3D respectively. $1-sigma$ final trajectory set bandwidths have been overlaid at converged mode positions. Post processing was disabled.

The approach involves running Mean Shift fixed point iterations at cluster levels, over a single data point per cluster. Starting out as trivial clusters (solitary data points), the clusters agglomerate between iterations. By algorithm design, clusters are merged only when they are tending towards the same mode. All member points of a cluster, u, which will eventually converge to a common local mode, share a common bandwidth, Σ_u - referred to as the local bandwidth. This bandwidth evolves every iteration, adapting to the structure of the local mode and to an extent, its basin.

The standard MS fixed point iteration, is reformulated through local bandwidth based decomposition, as a fixed point update over clusters :

$$u^{\tau+1} = f(u^{\tau}),\ where\ u^{\tau=0} \equiv x_{u,u} \quad (1a)$$

$$f(u^{\tau}) = \Big(\sum_{\forall g \in G} \frac{1}{c_g}\Sigma_g^{-1} \sum_{\forall i | x_{i,g} \in Ne_x(u^{\tau})} K'(\|u^{\tau} - x_i\|_{\Sigma_g})\Big)^{-1} \dots$$
$$\times \Big(\sum_{\forall g \in G} \frac{1}{c_g}\Sigma_g^{-1} \sum_{\forall i | x_{i,g} \in Ne_x(u^{\tau})} K'(\|u^{\tau} - x_i\|_{\Sigma_g}) x_i\Big) \quad (1b)$$

For ascertaining cluster merges, the data points in the vicinity of a cluster u's trajectory, u^{τ}, are considered. If a data point, y, in vicinity of u^{τ}, is ascertained to be heading to the same mode as u^{τ}, then by transitivity - all the members of its parent cluster, $\Pi(y)$, are heading to that mode too - the clusters u and $\Pi(y)$, can then be merged. The cluster which is higher up the mode (higher density) assimilates the other cluster into itself, thus accelerating convergence. This also helps in avoiding spurious merges.

The bandwidth, Σ_u, of a cluster, u, is updated every iteration utlizing T_u - the set of trajectory points arising from its constituent members.

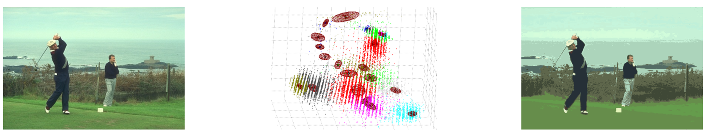
$$\Sigma_u = \frac{\sum_{\forall v \in T_u} \rho(v) v v^{\tau}}{\sum_{\forall v \in T_u} \rho(v)} - \eta_u \eta_u^T + \xi I,\ where\ \eta_u = \frac{\sum_{\forall v \in T_u} \rho(v) v}{\sum_{\forall v \in T_u} \rho(v)} \quad (2)$$

$\rho(v)$ is the data density in the immediate vicinity of a point $v \in T_u$. η_u and Σ_u are then the expectation, and variance of the localized distribution. *Eq.* 2 results in conservative but more localized and robust bandwidth estimates – more immune to long tails.

So starting with an initial base scalar, σ_{base}, the bandwidth matrices evolve by themselves. The nice part is that just a low base value suffices for reasonably dense data, with the bandwidths scaling data driven thereon and adapting to the local structure's scale, shape and orientation. σ_{base} thus becomes indicative of the minimum desired detail in the data space. This is opposed to traditional Mean Shift - where the bandwidth scalar is indicative of the scale at which data space has to be partitioned.

As *Figs. 1, 2(a)* indicate, reasonable local bandwidths arise, robustly identifying modes and salient clusters, by adapting according to local structure.

(*a*) Clustering (23 clusters) over color data (left) by the proposed approach. Segment image is shown on right.



(*b*) Comparitive results with standard MS (left) and variable-bandwidth isotropic MS, (VarMS, right), at similar clustering levels, 25 & 27 respectively, are shown. MS with correctly chosen bandwidth detected more coherent modes than VarMS, but looses partition saliency (bushes, water, sky in background). VarMS better adapts to scales but oversegments at places, and smooths over others (face). Both smoothed over details, failed to detect some modes at lower scales (trouser edges, maroon on shirt & shoes).

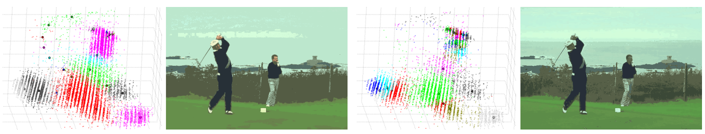


Figure 2: Exemplar illustrative result of our approach, AAAMS (a), is shown along with conventional MS results (b), at comparable clustering levels. As is indicated by the plots and segment images, AAAMS effectively adapts to local scale and preserves anisotropic details. This results in more salient yet parsimonius partitions.

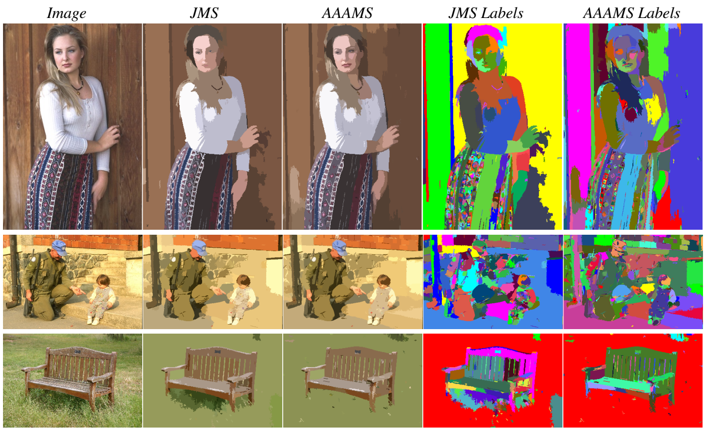


Figure 3: Example AAAMS results are shown, along with comparisons with standard joint domain Mean Shift (JMS). A single parameter set was used for AAAMS to show its adaptivity on varied images. At similar clustering levels, AAAMS preserved more details and affected more salient segmentations.

Promising qualitative and quantitative results were attained over image and point datasets - indicating the efficacy of the presented approach.

Future work would focus on experimenting with different merging schemes, and on more varied data spaces.