On averaging persistent Betti numbers

In the 90s our research group at the University of Bologna has studied the use of averages of persistent Betti numbers in degree 0 (i.e. size functions) for shape comparison. In plain words, skipping some technical details, the main idea was the following.

Let us consider a class C of objects sharing the same shape. Let us assume that n measurements $\varphi_r : S_r \to \mathbb{R}$ on n objects $\Sigma_r \in C$ are given. Each S_r is a topological space associated with the object Σ_r. As an example, S_r could be the surface of the solid object Σ_r.

Let $\beta_0^{\varphi_r}(u,v)$ be the persistent Betti number function in degree 0 of φ_r, i.e. $\beta_0^{\varphi_r}(u,v) := \dim i_*\left(H_0\left(\varphi_r^{-1}(-\infty,v]\right)\right)$ for $u<v$, where i_* is the homomorphism between homology groups that is induced by the inclusion $i:\varphi_r^{-1}(-\infty,u] \hookrightarrow \varphi_r^{-1}(-\infty,v]$. We can now consider the average persistent Betti number function

$$\beta_0^{av}(u,v) := \frac{1}{n}\sum_{r=1}^{n}\beta_0^{\varphi_r}(u,v)$$

as a shape fingerprint of the class C.

Every new object Σ associated with the filtering function $\varphi : S \to \mathbb{R}$ can be classified by comparing the persistent Betti number function $\beta_0^{\varphi}(u,v)$ of φ with the average functions $\beta_0^{j,av}$ associated with different classes C^j of objects.

A natural metric to make this comparison seems to be the L^1-metric. Unfortunately, a problem immediately arises, because we usually have

$$\int_{\mathbb{R}^2}\left|\beta_0^{\varphi}(u,v)-\beta_0^{j,av}(u,v)\right|\,du\,dv=\infty$$

for any j. Indeed, persistent Betti number functions in degree 0 usually differ from each other on a set of infinite measure. This is due to the fact that persistence diagrams in degree 0 always contain at least one point with infinite ordinate.

The following trick can solve this problem. For each point $(\bar{u},\bar{v})$ with $\bar{u}<\bar{v}$ (possibly $\bar{v}=\infty$), we fix an integrable function $f_{(\bar{u},\bar{v})} : \mathbb{R}^2 \to \mathbb{R}$. Now, we can consider the persistence diagram $Dgm_0(\varphi_r)$ and replace the function $\beta_0^{\varphi_r}$ with the function

$$\hat{\beta}_0^{\varphi_r} := \sum_{(\bar{u},\bar{v})\in Dgm_0(\varphi_r)} f_{(\bar{u},\bar{v})}.$$

After that, we consider the average function

$$\hat{\beta}_0^{av}(u,v) := \frac{1}{n}\sum_{r=1}^{n}\hat{\beta}_0^{\varphi_r}(u,v).$$

If we choose $f_{(\bar{u},\bar{v})}$ equal to the characteristic function of the triangle $\{(u,v)\in \mathbb{R}^2 : \bar{u}\le u<v<\bar{v}\}$, then $\hat{\beta}_0^{\varphi_r}\equiv\beta_0^{\varphi_r}$. However, in this case $f_{(\bar{u},\infty)}$ would not be integrable, and hence in practical applications the functions $f_{(\bar{u},\bar{v})}$ will be Gaussian-like functions.

We can classify the object Σ as an element of the shape class C^k if the L^1-distance

$$\int_{\mathbb{R}^2}\left|\hat{\beta}_0^{\varphi}(u,v)-\hat{\beta}_0^{j,av}(u,v)\right|\,du\,dv=\infty$$

takes its minimum for $j=k$.

This approach has been described in the paper

[DFL98] Pietro Donatini, Patrizio Frosini, Alberto Lovato, *Size functions for signature recognition*, Proceedings of SPIE, Vision Geometry VII, vol. 3454 (1998), 178–183.

Another application of this method has been described in the paper

[FFLZ98] Massimo Ferri, Patrizio Frosini, Alberto Lovato, Chiara Zambelli, *Point selection: A new comparison scheme for size functions (With an application to monogram recognition)*, Proceedings Third Asian Conference on Computer Vision, Lecture Notes in Computer Science 1351, vol. I, R. Chin, T. Pong (editors) Springer-Verlag, Berlin Heidelberg (1998), 329–337.

The previously described approach can be straightforwardly adapted to persistent Betti number functions in any degree.

For obvious copyright reasons not depending on us, the papers [DFL98], [FFLZ98] cannot be made freely available on the Web. However, if you are a colleague interested in them for research purposes, you can download copies for your personal and internal use by clicking on these links (username:*visitor* password:*1234*):

Size functions for signature recognition

Point selection: A new comparison scheme for size functions.