# Geometric Signatures for Fast Shape Processing

K. Clawson, S. Ibrahim, H. Van Dyke, T. Asaki, K. Vixie

*Washington State University* and *Center for Geometric Analysis and Data*

Washington State University
World Class. Face to Face.

## Concept: Finding shapes in streaming data

Geometric measure theory is just beginning to be exploited for its large potential to streaming data analysis. Recently, it has been used to develop a variety of promising tools for various sorts of data. Our interest in this problem spans theory and the application of this theory to analysis of streaming images. In this work, we demonstrate new multiscale signatures as well as fast surrogates for those signatures that permit processing of streaming data.

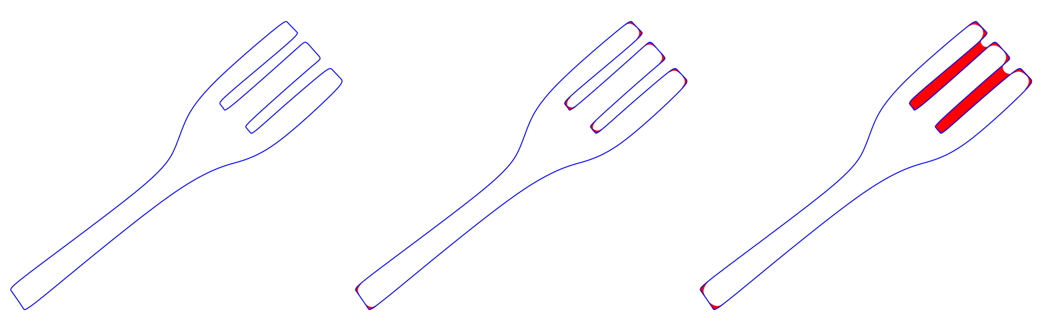
The multiscale flatnorm surrogates give us streaming speeds for shape recognition tasks: The development of scale based surrogates lead to new signatures for shapes and efficient methods for image streams. Moderate parallelization promises very high throughput speeds for image and shape analysis tasks that are focused on scale-based features.

## Flat Norm Scale Decomposition

The *flat norm with scale* $\lambda$, $\mathbb{F}_\lambda(T)$ of an oriented 1-dimensional set $T$ is given by

$$\mathbb{F}_\lambda(T) \equiv \min_S \{M(T - \partial S) + \lambda M(S)\} \qquad (1)$$

where $S$ varies over 2-dimensional regions and $M(U)$ indicates 1 or 2 dimensional volume of $U$.



These figures illustrate optimal flat norm decompositions. The boundary of the fork is the set $T$. The optimal decomposition is given by $S_\lambda$ and $T - \partial S_\lambda$. In these examples, $\lambda$ is decreasing left to right.

## $L^1$TV Computes the Flat Norm

The $L^1$TV functional

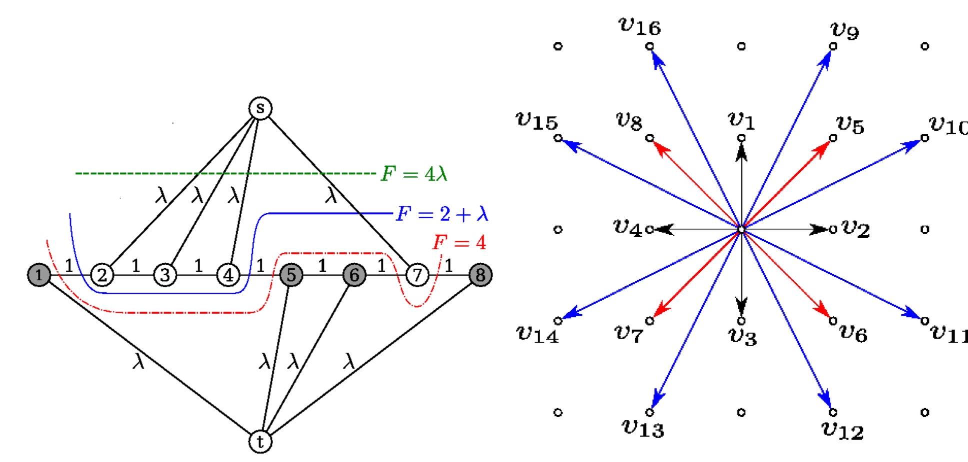
$$F(u; f, \lambda) \equiv \int_D |\nabla u| + \lambda \int_D |u - f|,$$

studied by Chan and Esedoglu[1] (and others), provides a means for computing the flat norm for characteristic functions. They showed that for sets $\Sigma$ and $\Omega$ supporting characteristic functions $u$ and $f$, respectively, the minimizer $u^*$ of $F(u; f, \lambda)$ has support $\Sigma^*$ which in turn minimizes

$$\mathcal{F}_\lambda(\Sigma; \Omega) = \text{Per}(\Sigma) + \lambda|\Sigma \Delta \Omega|$$

where $\Delta$ denotes the symmetric difference. The correspondence to the flat norm [2] is apparent with $\partial\Sigma = T - \partial S$ and $\Sigma\Delta\Omega = S$.

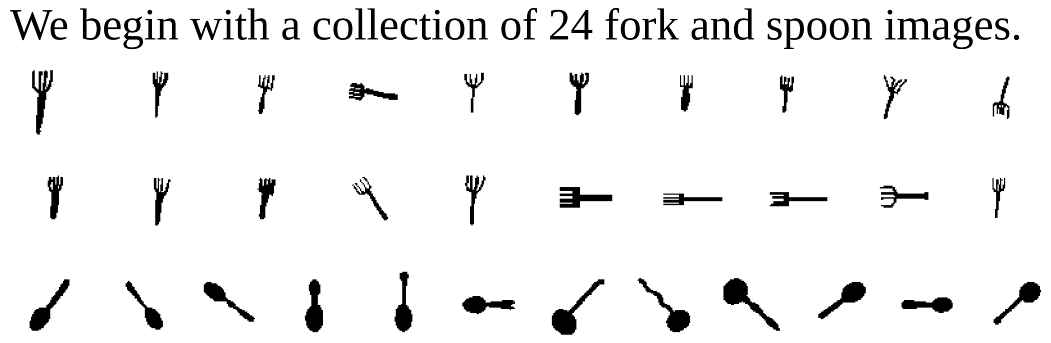
A relatively fast $L^1$TV computation on pixelized images is achieved using a graph cut formulation and an anisotropic gradient approximation.



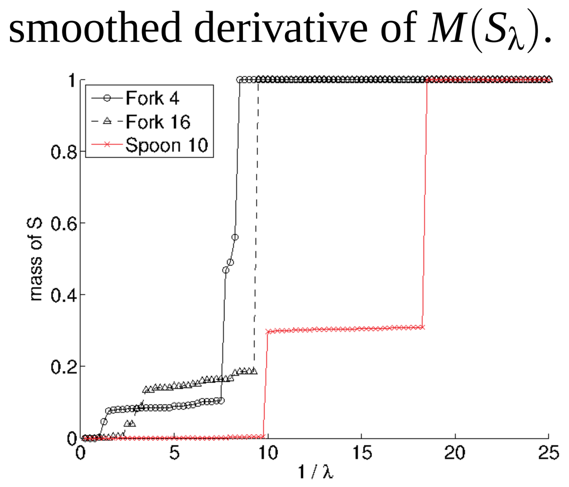

The capacity of any cut on this graph is equal to the value of $\mathcal{F}_\lambda$ for a particular region $\Sigma$. In the graph, $\Sigma\Delta\Omega$ corresponds to the set of all nodes $n$ for which either the edge $(n,s)$ or $(n,t)$ is in the cut. Any cut also incurs a penalty equal to the number of times the cutting line passes between image nodes. This penalty is equal to Per$(\Sigma)$. Thus, finding a cut with minimal capacity is equivalent to computing the flat norm.

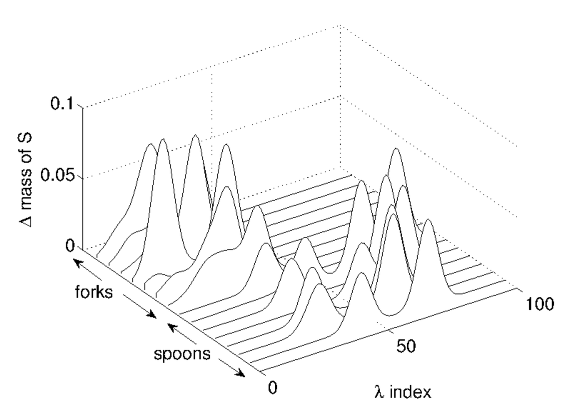
## Shape Classification Example [3]

We begin with a collection of 24 fork and spoon images.

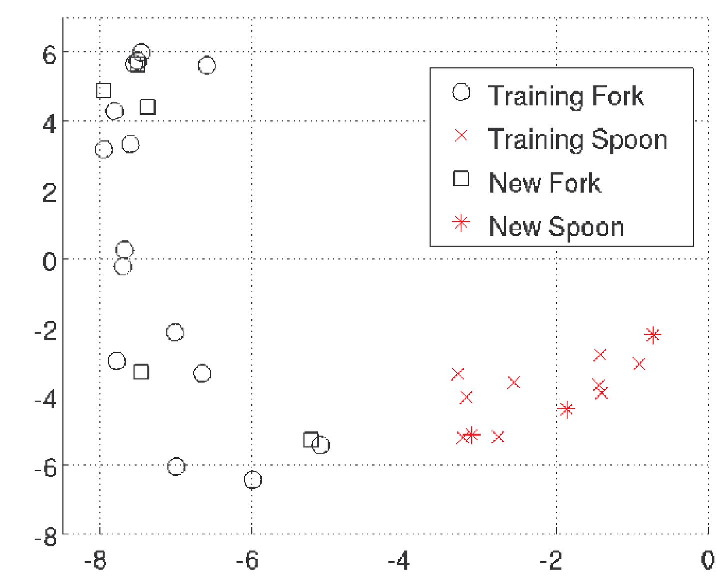


We then calculate the normalized multi-scale flat norm of each image for a sequence of $\lambda$ values. We consider the signatures $M(S_\lambda)$ which are always monotonic, and in particular, exhibit discontinuities corresponding to scales present in the object. The characteristic signature of any object is the smoothed derivative of $M(S_\lambda)$.




Classification is performed using clustering in $\mathbb{R}^m$. We solve the problem by using the singular value decomposition on the matrix of training data vectors and classify a data signature based on its projection onto the subspace spanned by the two most significant singular vectors. Ideally we would see a simple separation of the data into two groups. In the present example, sufficiently accurate spoon-fork classification is accomplished using projection onto only the first singular vector.




## Flat Norm Surrogates

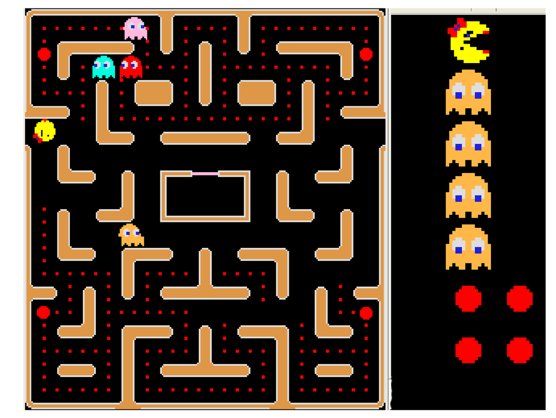
Algorithms for accurately and quickly computing the flat norm are not yet available for streaming video applications. As we pursue such algorithms, we are applying surrogate computations which capture some of the important scale-dependent properties of the flat norm but are relatively inexpensive to compute. Given binary image $B$ we generate a sequence of binary images $\{B_\lambda\}$ from which extract an area or perimeter derived signature $\{s_\lambda\}$.

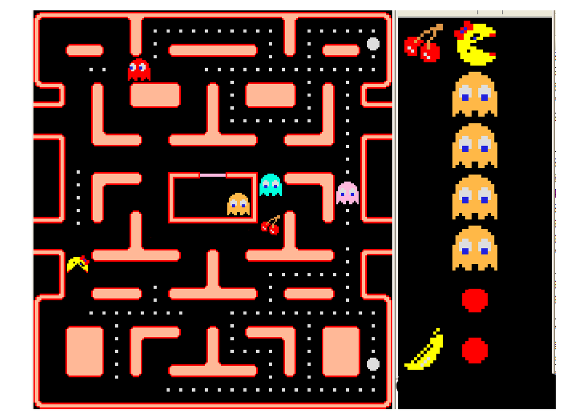
- **Convolution with Thresholding**. $B_\lambda = D(1/\lambda) * B$, where $D(r)$ is the normalized disk of radius $r$, and $s_\lambda =$ Area$(\{B_\lambda \geq 1/2\})$.
- **Heat Kernel Smoothing**. $B_{k+1} = K * B_k$, where $K$ is the heat kernel smoothing operator and $B_0 = B$. Then $s_k =$ Per$(\{B_k > t\})$, for some $t \in (0,1]$.
- **Pixel Erosion**. $B_{k+1} = B_k - \partial B_k$ with $B_0 = B$ and $s_k =$ Per$(B_k)$.

## Screen Capture Processing Example

As an illustration of a streaming video application we provide a live analysis demonstration of the classic video game Ms. Pac-Man. We consider the classification of dynamic objects in the scene given known perimeter signatures computed using the erosion surrogate.

The method performs relatively poorly when a given video frame contains objects for which characteristic signatures are not available such as happens when "collisions" occur among standard objects and when objects are partially visible at the edges of the screen.





## Streaming Video Implementation

There are several key features that allow us to analyze videos at streaming speeds:

- Live capture: we created a wrapper for FFmpeg/libav which allows video to be consumed from MATLAB
  - Flexible: Can switch between live (camera or screen) and prerecorded video by changing a single line of code.
- Surrogates are well-parallelizable:
  - Convolution thresholding
    - Running each scale on GPU (via CUDA) gives large speed-ups vs. equivalent CPU code
    - Each scale can be computed independently
  - Heat kernel/Pixel Erosion
    - Image can be divided into regions, processing has minimal communication overhead
  - General: Frames can be analyzed independently, but this introduces latencies. Not explored in current work.

## Key References

[1] T. F. Chan and S. Esedoḡlu, *Aspects of total variation regularized $L^1$ function approximation*, SIAM J. Appl. Math., 65 (2005), pp. 1817–1837.

[2] S. P. Morgan and K. R. Vixie, *$L^1$TV computes the flat norm for boundaries*, Abstract and Applied Analysis, 2007 (2007), pp. Article ID 45153, 14 pages. doi:10.1155/2007/45153.

[3] K. R. Vixie, K. Clawson, T. J. Asaki, G. Sandine, S. P. Morgan, and B. Price, *Multiscale flat norm signatures for shapes and images*, Applied Mathematical Sciences, 4 (2010), pp. 667–680.