# Contribution Title*

Panagiotis Rigas[1][0000−1111−2222−3333], Charalambos Tzamos[2,3][1111−2222−3333−4444], George Ioannakis[2,3][1111−2222−3333−4444], and Ioannis Z. Emiris[3][2222−−3333−4444−5555]

[1] Princeton University, Princeton NJ 08544, USA
[2] Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany
`lncs@springer.com`
http://www.springer.com/gp/computer-science/lncs
[3] ABC Institute, Rupert-Karls-University Heidelberg, Heidelberg, Germany
`{abc,lncs}@uni-heidelberg.de`

**Abstract.** The abstract should briefly summarize the contents of the paper in 15–250 words.

**Keywords:** First keyword · Second keyword · Another keyword.

## 1 First Section

### 1.1 A Subsection Sample

Please note that the first paragraph of a section or subsection is not indented. The first paragraph that follows a table, figure, equation etc. does not need an indent, either.

Subsequent paragraphs, however, are indented.

**Sample Heading (Third Level)** Only two levels of headings should be numbered. Lower level headings remain unnumbered; they are formatted as run-in headings.

*Sample Heading (Fourth Level)* The contribution should contain no more than four levels of headings. Table 1 gives a summary of all heading levels.
Displayed equations are centered and set on a separate line.

$$x + y = z \tag{1}$$

Please try to avoid rasterized images for line-art diagrams and schemas. Whenever possible, use vector graphics instead (see Fig. 1).

**Theorem 1.** *This is a sample theorem. The run-in heading is set in bold, while the following text appears in italics. Definitions, lemmas, propositions, and corollaries are styled the same way.*

* Supported by organization x.

**Table 1.** Table captions should be placed above the tables.

| Heading level | Example | Font size and style |
|---|---|---|
| Title (centered) | **Lecture Notes** | 14 point, bold |
| 1st-level heading | **1 Introduction** | 12 point, bold |
| 2nd-level heading | **2.1 Printing Area** | 10 point, bold |
| 3rd-level heading | **Run-in Heading in Bold.** Text follows | 10 point, bold |
| 4th-level heading | *Lowest Level Heading.* Text follows | 10 point, italic |

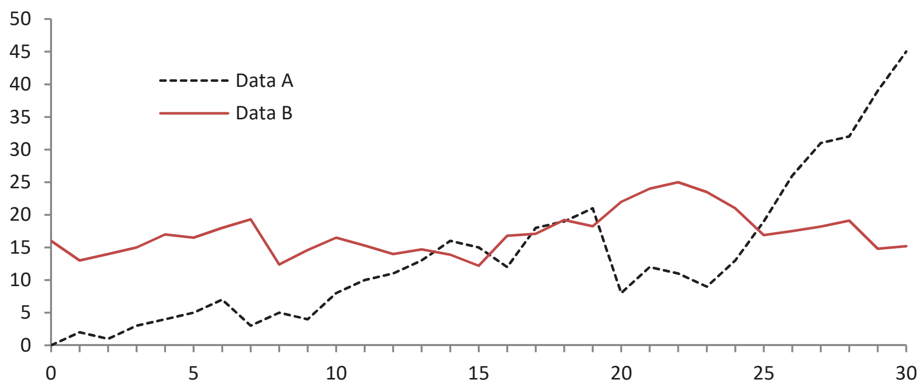


**Fig. 1.** A figure caption is always placed below the illustration. Please note that short captions are centered, while long ones are justified by the macro package automatically.

*Proof.* Proofs, examples, and remarks have the initial word in italics, while the following text appears in normal font.

For citations of references, we prefer the use of square brackets and consecutive numbers. Citations using labels or the author/year convention are also acceptable. The following bibliography provides a sample reference list with entries for journal articles [1], an LNCS chapter [2], a book [3], proceedings without editors [4], and a homepage [5]. Multiple citations are grouped [1–3], [1, 3–5].

## 2 Rotation clustering and compressing

In order to create spaces with objects that are relative Manhattan, we cluster the data by their relative rotation. Since we have one-dimensional clustering and random distributions, we choose to use kernel density estimation. The pdf can be derived as follows:

$$p(x) = \lim_{n\to\infty} p_n(x)$$

where $p_n(x)$ is equal to:

$$p_n(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{V_n}\phi\left(\frac{x-x_i}{h_n}\right) = \frac{1}{n}\sum_{i=1}^{n}\delta_n(x-x_i),$$

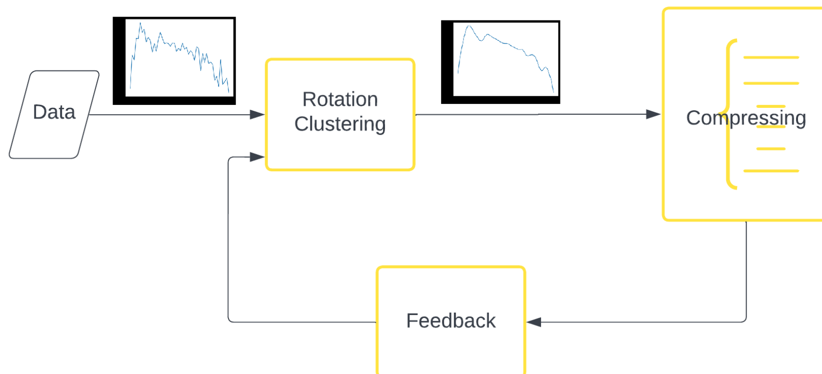


**Fig. 2.** Rotation clustering and compressing.

$$\delta_n(x) = \frac{1}{V_n}\phi\left(\frac{x}{h_n}\right)$$

For $\phi$ the following holds true:

$$\phi(x) \geq 0 \textbf{ and } \int_x \phi(x)\,dx = 1$$

Here, we choose $\phi$ to be equal to a gaussian distribution:

$$\phi\left(\frac{x - xi}{h_n}\right) = \frac{1}{h_n\sqrt{2\pi}}\epsilon^{-0.5\left(\frac{x-x_i}{h_n}\right)^2}$$

$h_n$ is equal to bandwidth and the variance of the gaussian distribution.

### 2.1 Bandwidth selection

The choice of bandwidth is crucial to the kernel density estimation. If the window is small or large, the curve produced may be undersmoothed, or oversmoothed respectively. Many methods where tested, such as rule of thumb, since our data are univariate and the density function is Gaussian, with bw being set equal to:

$$bw \approx 0.9\min(\hat{\sigma}, ICQ/1.34)n^{-1/5}$$

The Suleiman's rule of thumb, was discarded because the distribution is a sum of normal distributions. The second choice was to create a cost function and minimize it. We define The cost as:

$$J_m(x) = \sum_{cl=1}^{N_{cl}}\left(\sum_{i=1}^{N_g}\left(\frac{1}{n}\sum_{j=1}^{n_e\in i}\left(x_j^i\right)^m\right)^{\frac{1}{m}} \cdot w_{i,cl} - \left(\frac{1}{n}\sum_{j=1}^{n_e\in\max(w_{i,cl})}\left(x_j^i\right)^m\right)^{\frac{1}{m}} \cdot \max(w_{i,cl})\right)$$

The apply Newton's relaxed method of optimization.

$$bw_{k+1} = bw_k - \eta \frac{J^{(1)}(bw_k)}{J^{(2)}(bw_k)}$$

The derivatives are calculated with finite differences:

$$J^{(1)}(bw_k) \approx \frac{J(bw_k + s) - J(bw)}{bw_k} = \frac{1}{bw_k}\left(J(bw_k + s) - J(bw)\right)$$

$$J^{(2)}(bw_k) \approx \frac{J^{(1)}(bw_k + s) - J^{(1)}(bw)}{bw_k} =$$

$$\frac{1}{bw_k^2}\left(J(bw_k + 2s) - 2J(bw_k + s) + J(bw_k)\right)$$

The method that seems to produce the best results is done by remodeling the cost function as:

$$J_{new}(bw) = |J_{norm}(bw) - w_{Ncl} J_{Ncl,norm}(bw)|$$

Here the old cost is normalized and reduced by a weighted normalized cluster cost. The $w_{Ncl}$ is a hyper-parameter. This function produces a list of J values with indexes that point to a list of bw. We then can find the minimum cost corresponding bw for each cluster and choose the biggest bw, since it minimized the number of clusters. This has to be done for a field of bw values. One thing to consider is choosing the set of bw, which corresponds to a window in cost function. This is done by resizing the window to the area where error is not zero or infinity.

$$\max\left(\sum_{w_{cl}=w_0}^{w_n} \min J_n(bw, w_{cl})\right) = \max[bw_0, bw_1, \ldots, bw_N]$$

$$bw_N = bw_{opt}$$

## 2.2 Compression

The clusters produced by the process have a mean and a max value. The difference is related to the fact that each cluster is a weighted sum of the gaussians, each one contributing by a factor. We choose to compress each value of a cluster to each mean and not max, since the mean is represents the data more accurately than max, which is the mean of the gaussian with the most contribution.

$$C_i(j) = \delta(x - x_{m_i}) R_{m_i} \forall i, j$$

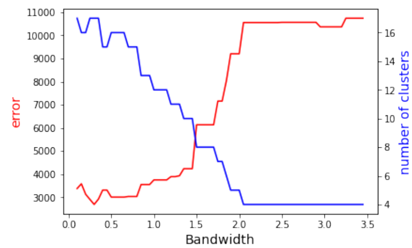


**Fig. 3.** Cost function and cost function for number of clusters per bw.

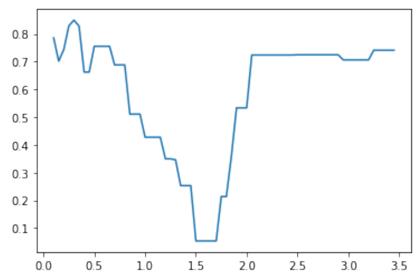


**Fig. 4.** New cost function for $w_{Ncl} = 1.5$.

## References

1. Author, F.: Article title. Journal **2**(5), 99–110 (2016)
2. Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016). https://doi.org/10.10007/1234567890
3. Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999)
4. Author, A.-B.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010)
5. LNCS Homepage, http://www.springer.com/lncs. Last accessed 4 Oct 2017