

Distance: A New Class of Methods

Abstract

Distance is that which separates two objects. In pattern recognition distance is that which separates elements, sample points, or even patterns. This notion of distance is inseparable from related topics which include the models which establish the parametric space of measuring, methods by which distance measures are applied, and the context in which the data is interpreted.

Classical statistical methods in pattern recognition rely heavily on the distribution of samples within a sample space. Often these are distributed in a normal or Gaussian distribution. Statistics ... not surprisingly ... and probabilistic aspects are indicated when handling such data. Classical syntactic methods in pattern recognition depend on the identification of features within a field and correlation of these features to match the structure dictated by a grammar. Family history data fits only marginally into either of these categories. There are probabilistic and statistical aspects to the existing data. There are certain statistical distributions and a certain structure to the collection, one of births, deaths, marriages, children, families, stationarity (stationary-ness), and mobility.

Typically the success criteria associated with distance in these methods are aggressive and rigorous: least squares, maximum likelihood, minimum error. Here weak methods are addressed. These weak methods generally arise from short thoughts or ideas generally applicable to human behavior. Careful consideration may allow such innovation to be injected into distance measure, or possibly into some other part of the model, methods, context decimation.

Continuing analysis suggests an alternate taxonomy for these methods. The *corridor methods* in pattern recognition are presented. The foundational principles for the corridor methods been known since the rise of civilization, but have only been formalized in the mathematics of art rather than in the sciences of pattern recognition. When one walks cross country, the elements of the landscape become visible in the distance, loom closer, are passed on the right or the left, recede into the distance, and disappear. The whole of this experience presents a metaphor for object processing. Chief among the advantages of these methods is the intrinsic handling of data which is identifiable but missing. These methods also draw heavily on established or new procedures in image processing.

Family historical data ... and all historical data for that matter ... is fragmentary. However that the lives and generations are very connected and continuous is clear. Considering these two elements in juxtaposition, the corridor methods address missing or more correctly occulted data in these fragmentary fields where the more classical solutions may falter.

Current research is addressing the nature of the conscious experience and the principles of intelligence extended to the most fundamental levels. Opportunities are being identified for hypothesis formation, establishing contexts for fragmentary information, and the measurement of distance.