Unsupervised Learning and Similarity Comparison of Water Mass Characteristics with Gaussian Mixture Model for Visualizing Ocean Data

Authors : Jian-Heng Wu, Bor-Shen Lin

Abstract : The temperature-salinity relationship is one of the most important characteristics used for identifying water masses in marine research. Temperature-salinity characteristics, however, may change dynamically with respect to the geographic location and is quite sensitive to the depth at the same location. When depth is taken into consideration, however, it is not easy to compare the characteristics of different water masses efficiently for a wide range of areas of the ocean. In this paper, the Gaussian mixture model was proposed to analyze the temperature-salinity-depth characteristics of water masses, based on which comparison between water masses may be conducted. Gaussian mixture model could model the distribution of a random vector and is formulated as the weighting sum for a set of multivariate normal distributions. The temperature-salinity-depth data for different locations are first used to train a set of Gaussian mixture models individually. The distance between two Gaussian mixture models can then be defined as the weighting sum of pairwise Bhattacharyya distances among the Gaussian distributions. Consequently, the distance between two water masses may be measured fast, which allows the automatic and efficient comparison of the water masses for a wide range area. The proposed approach not only can approximate the distribution of temperature, salinity, and depth directly without the prior knowledge for assuming the regression family, but may restrict the complexity by controlling the number of mixtures when the amounts of samples are unevenly distributed. In addition, it is critical for knowledge discovery in marine research to represent, manage and share the temperature-salinity-depth characteristics flexibly and responsively. The proposed approach has been applied to a real-time visualization system of ocean data, which may facilitate the comparison of water masses by aggregating the data without degrading the discriminating capabilities. This system provides an interface for querying geographic locations with similar temperature-salinity-depth characteristics interactively and for tracking specific patterns of water masses, such as the Kuroshio near Taiwan or those in the South China Sea.

Keywords : water mass, Gaussian mixture model, data visualization, system framework

Conference Title : ICMDMIS 2021 : International Conference on Marine Data Management and Information Systems

Conference Location : Dubai, United Arab Emirates

Conference Dates : June 29-30, 2021