A person’s height is influenced by many factors, such as her parents’ heights and how well she was nourished as a child. The features of an organism that can be seen are called phenotypes. Phenotypes that are influenced by multiple genes or the environment; that have many different features; and that vary smoothly over many different values; are called complex phenotypes. These features are called dimensions. Many agronomically valuable phenotypes are complex, such as how much grain can be produced by a field of corn. To understand these phenotypes better so that we could improve yield, we have to be able to recognize when two phenotypes differ and by how much they differ. One way to recognize such differences is by a computational technique called clustering. Clustering groups similar things together by some criterion. For example, different varieties of corn have different yields, and the way their yields change for different amounts of fertilizer also changes. Say two varieties of corn maximize their yields at different fertilizer amounts. Once we can compare different yields and amounts of fertilizer for different corn varieties, we can recognize these similarities.

Phenotypes described by only a few dimensions can be impossible to cluster reliably if the numbers for the different dimensions are not comparable to each other. This thesis demonstrates a novel way to make the dimensions comparable to each other and applies the method to 90 different varieties of corn grown under nine different combinations of water and fertilizer.