Adjustment Learning and Relevant Component Analysis

Noam Shental[1], Tomer Hertz[1], Daphna Weinshall[1], and Misha Pavel[2]

[1] School of Computer Science and Engineering and the Center for Neural Computation, The Hebrew University of Jerusalem, Jerusalem, Israel 91904
[2] Department of Electrical and Computer Engineering, Oregon Health and Science University, Beaverton OR 97006
{fenoam,tomboy,daphna}@cs.huji.ac.il, pavel@ece.ogi.edu

Abstract. We propose a new learning approach for image retrieval, which we call *adjustment learning*, and demonstrate its use for face recognition and color matching. Our approach is motivated by a frequently encountered problem, namely, that variability in the original data representation which is not relevant to the task may interfere with retrieval and make it very difficult. Our key observation is that in real applications of image retrieval, data sometimes comes in small chunks - small subsets of images that come from the same (but unknown) class. This is the case, for example, when a query is presented via a short video clip. We call these groups *chunklets*, and we call the paradigm which uses chunklets for unsupervised learning *adjustment learning*. Within this paradigm we propose a linear scheme, which we call Relevant Component Analysis; this scheme uses the information in such chunklets to reduce irrelevant variability in the data while amplifying relevant variability. We provide results using our method on two problems: face recognition (using a database publicly available on the web), and visual surveillance (using our own data). In the latter application chunklets are obtained automatically from the data without the need of supervision.

1 Introduction

We focus in this paper on a fundamental problem in data organization and data retrieval, which is often ignored or swept under the carpet of "feature selection". More specifically, our domain is image retrieval and image organization based on between-image distance. It is assumed that the specifics of the image representation and between-image distance were pre-determined, typically based on some engineering design that took into account a very large scope of tasks and data. Within this domain, data organization essentially requires graph-based clustering. Retrieval is based on the distance from a query image to stored images, and is essentially nearest neighbor classification.

We identify the following problem: Assuming that the data is represented compactly, a typical distance (whether Euclidean in feature space or other) is affected by all the variability that is maintained in the data representation indiscriminantly. However, for a particular query some of this variability is irrelevant.

A. Heyden et al. (Eds.): ECCV 2002, LNCS 2353, pp. 776–790, 2002.

For example, if we submit as a query the image of Albert Einstein sticking his tongue out, the relevant features in the image may be the expression (if we want to retrieve other silly faces), or the relevant features may be the facial features (if we want to retrieve other pictures of Albert Einstein).

Irrelevant variability is also a problem in data organization, e.g., in clustering (which has been used in vision for image retrieval [7]). Typically we collect data and rely on the pre-determined representation and distance function to cluster the data. However, depending on the context of the data collection, some variability is always irrelevant. Unaccounted for, this will damage the results, since clustering only gives as good results as the underlying representation and distance function.

More precisely, we first define data variability to be *correlated with a specific body of data and a specific task*, if the removal of this variability from the data deteriorates (on average) the results of clustering or retrieval. We then define *irrelevant variability* as data variability which satisfied the following conditions:

- it is normally maintained in the data (i.e., it is used in the representation and/or distance function explicitly or implicitly);
- it is **not** *correlated with the specific task* according to the definition above and the task at hand.

Intuitively, in a task where irrelevant variability is large, performance using the general and indiscriminating pre-determined distance measure can be poor. For example, when a certain irrelevant environmental feature dominates all other features because of its large variability, nearest neighbor classification using the indiscriminating distance function could perform poorly. This is the problem addressed in the present paper.

We propose an unsupervised adjustment scheme, which will modify the distance function so as to enhance its discriminative capabilities on the given test data. For this purpose we identify and use relatively small sets of data points, in which the class label is constant, but unknown. We call these small sets of data points *Chunklets.*

We justify our use of *chunklets* for data clustering and retrieval by the following observations:

Data organization: we observe that very often data naturally arrives in chunks in which the message is held constant. For example, in speaker identification, short utterances of speech are likely to come from a single speaker. In surveillance, short video clips obtained by tracking a moving target are likely to come from a single target. In both examples, the magnitude of the variance due to the irrelevant variability is comparable if not larger than the relevant variability. In both examples, we can automatically detect chunks of data where the message is constant while the context changes.

Data retrieval: in image retrieval based on nearest-neighbor classification chunklets may also come by naturally. For example, often a query includes a number of images, especially when the user is allowed to use some of the retrieved images in the first stage to refine the answers of the system in a