Automatic Image Annotation Using Group Sparsity

Shaoting Zhang¹, Junzhou Huang, Yuchi Huang, Yang Yu, Hongsheng Li and Dimitris N. Metaxas³
¹CBIM, Rutgers University, ²IDEA Lab, Lehigh University

Introduction

Goal: image annotation is to automatically assign relevant text keywords to any given image, reflecting its content.

Previous methods: topic models, mixture models, discriminative models, nearest neighbor based methods, etc.

Limitations: features are often preselected. Feature properties are well investigated in this application.

Our method: feature selection, and use clustering prior and sparsity prior to guide the selection.

Annotation framework

Assume we have similar and dissimilar image pairs from training data, denoted as (P1, P2).

Denote the feature matrix of P1 and P2 as F_P1 and F_P2 respectively. Then the difference matrix can be represented as X = | F_P1 - F_P2 |.

Denote the similarity target as 1 (similar) or 0 (dissimilar), and the similarity vector as Y. We want to find weights w which can minimize Xw = Y in least square sense.

Given any testing data, we first compute the its similarities with all training data. Then select highest values as the best matches. Finally the keywords of these best matches are transferred to the testing image according to keywords’ local frequency.

When some entries of the weight vector equal to zero, we can prune corresponding columns in X, i.e. feature selection.

Regularization methods

- No regularization: analytical solution; not stable.
- L^2 regularization: robust, solvable \( \langle X^T X + \lambda I \rangle^{-1} X^T Y \), no sparsity.
- L^1 regularization: sparsity prior; convex; basis pursuit, Grafting, Shooting, etc.
- Group Sparsity: L^2 inside the same group, L^1 for different groups; projected-gradient.

Obtain Image Pairs

Traditional method solely relies on keyword similarity.

We use EM like method and feedback information to improve the quality of image pairs.

In E step, we first select pairs according to feature distance, then prune them according to keyword similarity or category information.

In M step, we compute weights w by solving regularization problem. Repeat this procedure.

Experiments

Evaluation of Features

Evaluation of Image Pairs

Evaluation of Regularization Methods

References