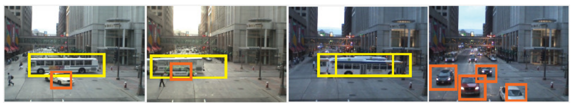
Incremental Multiple Kernel Learning for Object Detection

Aniruddha Kembhavi, Behjat Siddiquie, Scott McCloskey, Roland Miezianko and Larry S. Davis

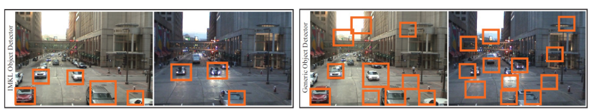
University of Maryland 1856

Motivation

- Components of a visual categorization system:
  - Representative training dataset
  - Efficient and effective feature extraction methods
  - Powerful classifier
- Obtaining a generic training dataset is relatively easy
- Obtaining a scene specific training dataset for a given application is harder
  - Fair amount of manual labor required for every new scene



- Scene specific characteristics of a traffic intersection:
  - Camera location and typical vehicle paths restrict observed poses
  - Camera location restricts the negative class (background)
  - Images of vehicles and background change over time
    - Changing illumination conditions
    - Shadows cast by the buildings



- Our Incremental Multiple Kernel Learning (IMKL) based approach initializes with a generically obtained training database
- It tunes itself automatically towards the classification task
  - Updates the training dataset, tailoring it towards the scene
  - Updates the weights used to combine multiple information sources
  - Tunes the classifier in an online fashion
- Ability to remove training examples over time
  - Useful when dealing with changing illumination conditions
- IMKL approach is a fusion of:
  - Multiple Kernel Learning (MKL) †
  - Incremental Support Vector Machine (ISVM) *

† A. Rakotomamonjy, F.R. Bach, S. Canu and Y. Grandvalet
More efficiency in multiple kernel learning. ICML 2007
* G. Cauwenberghs and T. Poggio.
Incremental and decremental support vector machine learning.
NIPS, 2000

IMKL Algorithm

IMKL Optimization Problem

$$\begin{aligned} \min \quad & \sum_k \frac{1}{d_k} w_k w_k^T + C \sum_i \xi_i \\ \text{such that} \quad & y_i \sum_k \phi_k(x_i) + y_i b \geq 1 - \xi_i \ \ \forall i \\ & \xi_i \geq 0 \ \ \forall i, \ \ d_k \geq 0 \ \forall k, \ \ \sum_k d_k = 1 \end{aligned}$$

→

KKT conditions

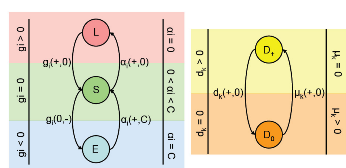
$$\begin{aligned} g_i = \sum_j \sum_k d_k \alpha_j Q_{ij}^k + y_i b - 1 = 0 \\ \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j Q_{ij}^k + \mu_k - \lambda = 0 \\ \sum_i \alpha_i y_i = 0 \\ \mu_k d_k = 0 \\ \sum_k d_k = 1 \end{aligned}$$

- Optimization problem is convex
  - KKT conditions and *necessary* and *sufficient*
- When a new point x_{new} is added, we need to calculate its Lagrange multiplier α_{new}:
  - Bounded by 0 and C
  - Begin with 0 and keep incrementing till solution is reached
  - Every time we increment α_{new}, we must update the remaining Lagrange multipliers, kernel weights and bias to maintain the KKT conditions
  - These changes are given by the differential forms of the KKT conditions

Differential Form of KKT conditions

$$\begin{aligned} & \sum_j \alpha_j \sum_k \Delta d_k Q_{ij}^k + \sum_j \Delta\alpha_j \sum_k Q_{ij}^k \\ & + \sum_j \sum_k \Delta d_k \Delta\alpha_j Q_{ij}^k + y_i \Delta b = 0, \ \forall i \in S, \forall j \in \{S, E, L, q\} \\ & \sum_i \sum_j \Delta\alpha_i \alpha_j Q_{ij}^k + \frac{1}{2} \sum_i \sum_j \Delta\alpha_i \Delta\alpha_j Q_{ij}^k \\ & + \Delta\mu_k - \Delta\lambda = 0, \qquad \forall k \in K \\ & \sum_i \alpha_i y_i = 0 \forall i \in \{S, E, L, q\}, \ \ \sum_k \Delta d_k = 0 \\ & \Delta\mu_k d_k + \mu_k \Delta d_k + \Delta\mu_k \Delta d_k = 0, \qquad \forall k \in K \end{aligned}$$

- Differential equations hold when α_{new} is small enough to ensure that there is no change in set membership
  - When set membership changes, equations are updated

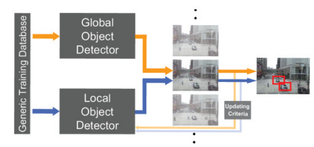


- Termination Conditions:

$g_{new} > 0$	at $\alpha_{new} = 0$	Correctly classified (set L)
$g_{new} = 0$	before $\alpha_{new} = C$	Support Vector (set S)
$g_{new} < 0$	at $\alpha_{new} = C$	Wrong side of the margin (set E)

Object Detection Framework

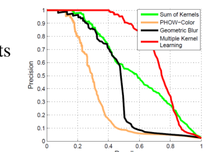
GLOBAL detector	LOCAL detector
• Initialized with a generic training dataset • Never updated • Helps classify atypical objects • Prevents overfitting	• Initialized with a generic training dataset • Updated using IMKL • Tunes itself towards recently observed image patches



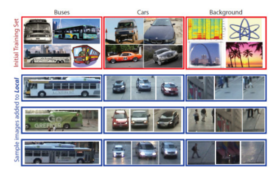
Background image updating

Experiments

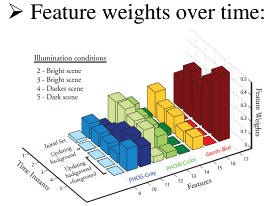
- 17 kernels used:
  - Pyramidal Histogram of Oriented Gradients
  - Pyramidal Histogram of Visual Words
    - Gray scale and color variants
  - Geometric Blur



- *Local* Dataset Snapshots:

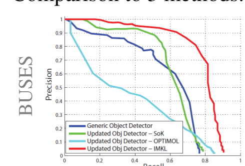


- Feature weights over time:

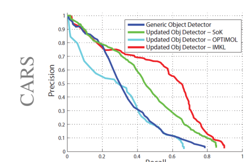


- Comparison to 3 methods:

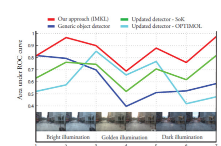
BUSES



CARS



- Performance over time:



- Efficiency comparison: