A ROBUST OBJECT DETECTION APPROACH USING BOOSTED ANISOTROPIC MULTIRESOLUTION ANALYSIS

Rashid Minhas, Abdul Adeel Mohammed, Q.M. Jonathan Wu, Maher A. Sid-Ahmed

Department of Electrical Engineering,
University of Windsor, ON, N9B3P4, Canada

ABSTRACT

In an unconstrained environment, adaptive classifiers produce improved recognition. Fast discrete curvelet transform has recently gained attention due to its ability to capture singularities along curves far away from smooth regions. Therefore, curvelet coefficients contain enhanced representation of image details at different scales and orientations. We propose a new approach for object class detection based on *curvelet feature subspace* obtained using Kernel PCA (KPCA) and learned using AdaBoost scheme [1]. The first contribution of current paper is a unique representation of an image called *curvelet feature subspace* that preserves global structure, and supports reliable detection of singularities along curves which play a considerably important role in recognition. Second contribution of our proposed method is an adaptive selection of features obtained using anisotropic style multiresolution analysis for robust object detection of varied inter-class, and intra-class attributes. Our proposed method achieved better detection rate compared to *state-of-the-art* schemes.

Index Terms— Multiresolution analysis, AdaBoost, KPCA, Object detection, Supervised learning

1. INTRODUCTION

Research community has recently observed considerable interest in automated object detection due to various applications in law enforcement, military, industry and content based image retrieval. The classification of visual objects in an unconstrained environment is a difficult task due to varying illumination conditions, pose variations and noisy measurements (see Fig. 1). To classify complex data samples, most classification techniques use fixed set of features. Practically, fixed features cannot always be detected due to view point changes, and instability of feature descriptors.

Traditional visual classification techniques extract features using machine learning approaches and generate a feature model for classification. Higher data dimension in classification represents a bottleneck towards real time performance. Recently, classification techniques based on partial / global information [6-10, 13, 5] and cascaded/hierarchical style schemes [11-12] have been proposed. Partial information based classification schemes represent object of interest using important local features and their mutual geometrical information, whereas global information based frameworks use dimension reduction methodology to minimize computational complexity without sacrificing important information. In cascaded style classification, the cascade can be considered as a target oriented mechanism which statistically guarantees that regions discarded at earlier stage are unlikely to contain object of interest. Our proposed method uses global structure information of an image using curvelet feature subspace which is a unique and sparser representation for singularities along curves. Sparse representation of singularities plays considerably important role in recognition and classification tasks.

The remainder of this paper consists of six sections. Sections 2-4 discuss theory of anisotropic multiresolution analysis, kernel PCA (KPCA) and AdaBoost algorithm. Section 5 contains proposed algorithm with results and conclusion explained in section 6-7 respectively.

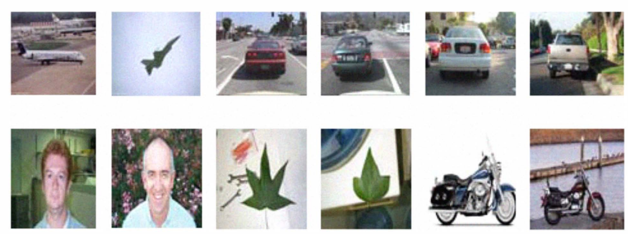

Fig. 1: Different images (from datasets used in our experiments) with different illumination, background, and changing view point.

2. FAST DISCRETE CURVELET TRANSFORM

Fourier series decomposes a periodic function into a sum of simple oscillating functions, namely *sines* and *cosines*. In a Fourier series sparsity is destroyed due to discontinuities (Gibbs Phenomenon) and it requires a large number of terms to reconstruct a discontinuity precisely. Development of new mathematical and computational tools based on multiresolution analysis is a novel concept to overcome limitations of Fourier series.

Table 1: FDCT via Wrapping

1. Apply 2D FFT and obtain Fourier samples $\hat{f}[n_1, n_2]$, $-n/2 \le n_1,\ n_2 < n/2$.
2. For each scale j and angle ℓ, form the product $\tilde{U}_{j,\ell}[n_1, n_2]\hat{f}[n_1, n_2]$.
3. Wrap this product around the origin and obtain $\hat{f}_{j,\ell}[n_1, n_2] = W(\tilde{U}_{j,\ell}\hat{f})[n_1, n_2]$, where the range n_1 and n_2 is now $0 \le n_1 < L_{1,j}$ and $0 \le n_2 < L_{2,j}$.
4. Apply the inverse 2D FFT to each $\hat{f}_{j,\ell}$, hence collecting the discrete coefficients.

Many fields of contemporary science and technology benefit from multiscale, multiresolution analysis tools for maximum

978-1-61284-857-0/11/$26.00 ©2011 IEEE

throughput, efficient resource utilization and accurate computations. Multiresolution tools render robust behavior to study information content of images and signals in the presence of noise and uncertainty. Wavelet transform is a well known multiresolution analysis tool capable of conveying accurate temporal and spatial information. Wavelet transform has been profusely used to address problems in data compression, pattern recognition, image reconstruction, and computer vision. Wavelets have better ability to represent objects with point singularities in 1D and 2D space but fail to deal with singularities along curves in 2D. Discontinuities in 2D are spatially distributed which leads to extensive interaction between discontinuities and many terms of wavelet expansion. Therefore, wavelet representation does not offer sufficient sparseness for image analysis. Following the introduction of wavelet transform, research community has witnessed intense efforts for development of contourlets [15], ridgelets [16]. These tools have better directional and decomposition capabilities than wavelets.

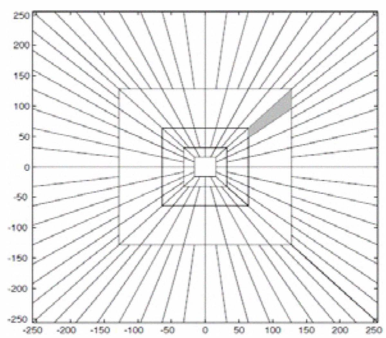

Fig. 2: Digital tiling; localized FFT with sheared wedges [4]

Fast discrete curvelet transform (FDCT) [4] is a recent addition to the family of multiresolution analysis tool that is designed and targeted to represent smooth objects with discontinuity along a general curve. Curvelet transform overcomes limitations of existing multiresolution analysis schemes and offers improved directional capacity to represent edges and other singularities along curves. Curvelet transform is a multiscale non-standard pyramid with numerous directions and positions at each length and scale. Curvelets outperform wavelets in situations that require optimal sparse representation of objects with edges, representation of wave propagators, image reconstruction with missing data etc. Curvelets have useful geometric features that set them apart from wavelets [4]. Two new algorithms have been proposed in [4] to improve previous implementations of FDCT as follows

1. *Curvelet via USFFT*
2. *Curvelet via wrapping*

New implementations of FDCT are ideal for deployment in large-scale scientific applications due to their low computational complexity. We used FDCT via wrapping, detailed in table 1, for our proposed scheme. For digital coronization, figure 2 illustrates digital tiling where $\tilde{U}_{j,\ell}$ smoothly localizes the Fourier transform near the sheared wedges obeying parabolic scaling.

3. KERNEL PRINCIPAL COMPONENT ANALYSIS

Traditionally, principal component analysis (PCA) is a powerful technique for extracting structure information from higher dimension data. PCA is an orthogonal transformation of the coordinate system and is evaluated by diagonalizing the covariance matrix. If $x_i \in R^N$, $i = 1,2,3......m$, represent a set of feature vectors which are centered with zero mean, their covariance matrix is evaluated as:

$$C = \frac{1}{m}\sum_{j=1}^{m} x_i x_j^T .$$

Eigen value equation, $\lambda v = Cv$ is solved where v is eigenvector matrix. To obtain a data with N dimensions, eigenvectors corresponding to the N largest eigenvalues are selected as basis vectors of the lower dimension subspace.

KPCA [17] is an important technique for dimensionality reduction with improved ability to encode structural information, and to capture nonlinear image features. KPCA is a generalization of PCA to compute the principal components of a feature space that is nonlinearly related to an input space. Feature space variables are obtained by higher order correlations between input variables. KPCA acts as a nonlinear feature extractor by mapping input space to a higher dimension feature space through a nonlinear map where the data is linearly separable. Mapping achieved using the kernel trick solves the problem of nonlinear distribution of low level image features and dimensionality reduction. Data is transformed form a lower dimension space to a higher dimension using the mapping function $\phi : \mathrm{R}^{\mathrm{N}} \rightarrow \mathrm{F}$, and linear PCA is performed on F. The covariance matrix in the new domain is calculated using following equation:

$$\overline{C} = \frac{1}{m}\sum_{j=1}^{m} \phi(x_i)\ \phi(x_j)^T . \qquad (1)$$

The problem is reduced to an Eigenvalue equation as in PCA and is solved using the identity $\lambda v = \overline{C}v$. As mentioned earlier the nonlinear map φ is not computed explicitly and is evaluated using the kernel function $K(x_i,x_j)=(\varphi(x_i),\varphi(x_j))$. The kernel function implicitly computes the dot product of vectors x_i and x_j in the higher dimension space. Kernels are considered as functions measuring similarity between instances. The kernel value is high if two samples are similar and zero if they are distant. Some of the commonly used kernel functions are Gaussian, polynomial and sigmoid kernels. Pair wise similarity amongst input samples is captured in a Gram matrix K and each entry of the matrix $K_{i,j}$ is calculated using the predefined kernel function $K(x_i,x_j)$. Eigenvalue equation in terms of Gram matrix is written as:

$$m\lambda\beta = K\beta \qquad (2)$$

K represents a positive semi definite symmetric matrix and contains a set of Eigenvectors which span the entire space. β denotes the column vector with entries $\beta_1, \beta_2 ....\beta_m$. Since the eigenvalue equation is solved for β instead of eigenvector V_i of the kernel PCA, the entries of β are normalized in order to ensure that the eigen values of kernel PCA have a unit norm in the feature space. After normalization the eigenvector matrix of kernel PCA is computed as $V = D\beta$ where $D = [\phi(x_1)\,\phi(x_2)\,......\,\phi(x_m)]$ is a data matrix in the feature space.

4. ADAPTIVE BOOSTING ALGORITHM

AdaBoost facilitates powerful incremental learning approach for classification. AdaBoost represent an ensemble learning approach using a collection of weak learners trained in an iterative fashion, where each weak learner is selected based on its classification accuracy on the training set.

Table 2: The adaptive boosting algorithm

INPUT: Sequence of N labeled training examples $\{(x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N)\}$ where y_i represents label of example x_i.

- Distribution *D* over the *N* training images
- Weak learning algorithm **WeakLearn**
- Integer T representing maximum number of iterations
- **Initialize** the weight vector: $\varphi_{1,i}=D(i)$ for $i=1,\ldots,N$.
- **Do for** $t=1,2,\ldots,T$

1. $$p_{t,i} = \frac{\phi_{t,i}}{\sum_{i=1}^{N} \phi_{t,i}}$$
2. Call **WeakLearn,** providing it with the distribution $p_{t,i}$; get back a hypothesis $h_{t,i}(x) = X \rightarrow [0,1]$
3. Calculate the error of $h_{t,i}$:
$$\varepsilon_{t,i} = \sum p_{t,i} \left| h_{t,i}(x) - y_i \right|$$
4. $$\beta_{t,i} = \frac{\varepsilon_{t,i}}{1 - \varepsilon_{t,i}}$$
5. Set the new eight vectors to be
$$\phi_{t+1,i} = \phi_{t,i} \beta_{t,i}^{1-\left|h_{t,i}(x) - y_i\right|}$$

OUTPUT: Classifier - $h_f(x)$

$$h_f(x) = \begin{cases} 1 & \text{if } \sum_{t=1}^{T} (\log 1/\beta_t) h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \log 1/\beta_t \\ 0 & \text{otherwise} \end{cases}$$

AdaBoost removes inherent problems in supervised learning such as overtraining, high error rate and computational cost, and tenders more emphasis on data that is hard to classify. The main idea of boosting is to combine several weak learners to form an *ensemble,* where each weak learner performs slightly better than a random guess. An *ensemble* is formed in a fashion that the performance of individual ensemble member is improved, i.e., *boosted.* Suppose we have a set of hypotheses h_1, h_2, ..., h_n; combined *ensemble* hypothesis takes the form

$$h_f(x) = \sum_{i=1}^{n} \varphi_i h_i(x)$$

where φ_i denotes the weight for individual *ensemble* members. The idea of boosting finds its roots back to PAC learning algorithm [3]. Main steps involved in AdaBoost method are presented in Table 2.

5. PROPOSED METHOD

Our proposed method enables us to exploit 1) FDCT subspace selected through KPCA that offers sparse and better representations of singularities along curves, 2) highly discriminative features used by AdaBoost for classification. The proposed method is also extendable to multi-classification problems with minor modifications in the boosting scheme.

In this work curvelet based features of objects are extracted using FDCT via wrapping technique. The coarse level coefficients of FDCT are selected for object representation and their dimension is reduced using kernel PCA. Approximate coefficients are selected since they contain overall structural information of an image instead of high frequency content only which does not greatly help to improve accuracy.

Images from each dataset are converted into gray level image with 256 gray levels. The only pre-processing steps performed on our input images are i) conversion from RGB to gray level format, ii) image size reduction to an equal size. We do not perform any further operations that may lead to image degradation(s). Next, we randomly divide image database into two sets namely training, and testing sets. Recently, research community has observed dimensionality reduction techniques being applied on data to be classified for real-time, accurate and efficient processing. All images of the datasets are resized to the same dimension, i.e., Row x Col before further processing. Such image resizes support assembly of equal sized curvelet coefficients and feature vector extraction with identical level of global content.

Let *Training*, *Test* represent sets of *n* training and *m* testing images respectively; where individual training and test images are represented by $Training_i$ and $Test_j$.

Table 3: Our proposed algorithm's different steps

INPUT: Randomly divide image dataset into two subsets $Training_i$ and $Test_j$ where i={1,2,…,N} and j={1,2,…,M} representing training and test images of equal size i.e. *RowxCol*.

OUTPUT: Classifier - f(x)

1. Compute the curvelet transform of all images and extract coarse level feature sets. (Refer to [4] for more details)
2. Vectorize coarse level feature sets into 1D vector.
3. Compute the kernel matrix $K_{Training}$ and $K_{Testing}$ where each entry of the matrix is computed using Gaussian kernel function mentioned in table - 1.
4. Solve eigenvalue equations:
$$u \Lambda A_{Training} = K_{Ttraining} A_{Training}$$
$$u \Lambda A_{Testing} = K_{Testing} A_{Testing}$$
where Λ, A_x are eigenvalue and eigenvector matrices respectively.
5. Obtain kernel PCA based feature vectors, f_p, by computing principal component projections of each image into non-linear subspace using A_x.
6. Train AdaBoost classifier, algorithm in table 2, using KPCA based feature vectors computed in step 5.
7. Classify test images' feature vectors using AdaBoost classifier trained in step 6.

FDCT is computed for training images for a particular scale and orientations. Anisotropic nature of FDCT produces sparse representation due to its variable orientations and least number of required coefficients for better approximation. Ali and Shah [5] proposed the use of gradient Sobel operator to extract features from resized images. Gradient operators are well suited for detection in single direction but cannot precisely track singularities along curves. Hence, we propose the use of curvelet transform for reliable detection of singularities in high dimensional data. Depending upon selected scale and number of orientations, FDCT computes coefficients of different sizes. In our experiment, we selected scale at 2 levels with 8 different orientations for feature computation using training and test image sets.

6. RESULTS AND DISCUSSION

We used Caltech databases for performance comparison of our proposed method with the state-of-the-art schemes proposed in [5, 13, 8, 14]. Our experiments are based on binary classification problem although it can also be extended to multiclass scenario. Cars, airplanes, motorbikes, faces, leaves and background image sets from Caltech are used in our experiments. Images are randomly selected from Caltech background dataset to represent negative images. We used same number of training and test images as mentioned in [5] for justified comparison. Table 4 represents comparison of our proposed method against well established schemes. It is clearly evident that detection rate for airplane and motorbike datasets is maximized using our proposed scheme, whereas for supplementary datasets accuracy achieved is comparable to existing popular methods.

Table 4: Classification accuracy (%age) analysis

Data Set(s)	Our Method	[5]	[13]	[8]	[14]
Caltech Faces	97.9	98	96.4	-	93.5
Caltech Motorbikes	95.3	93.4	92.5	73.9	92.2
Caltech Airplane	93.1	90	90.2	92.7	88.9
Caltech Cars	93.2	96	90.3	97	97
Caltech Leaves	94.5	94.2	-	97.8	-

Number of dimensions play pivotal role in accurate detection and classification. As KPCA encodes global structure of the scene where important information is represented by dimensions corresponding to higher eigenvalues; there is no criterion to select optimal number of dimension for discriminative selection. Number of selected dimensions represents a trade-off between required accuracy and available computational resources. Higher number of dimensions does not guarantee accurate classification at the cost of higher computation and *vice versa*. Figure 3 shows detection accuracy (in %age) against number of dimensions for Caltech airplanes dataset; this validates our intuitive idea that increasing number of dimensions is not an optimal choice to attain improved accuracy. Since a limited number of dimensions in a nonlinear space is discriminative and choosing the number of dimensions in a monotonically increasing fashion leads to saturation without any contribution towards better classification.

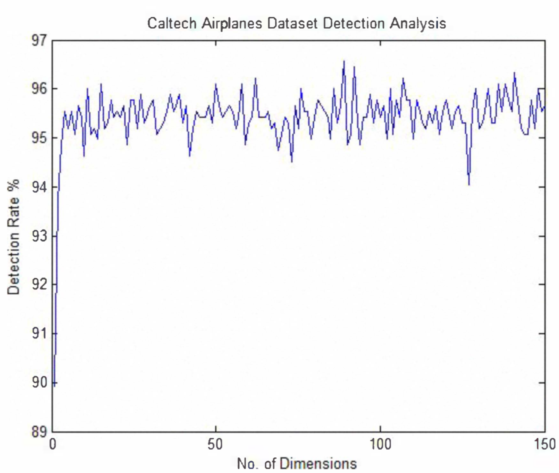

Fig. 3. Dimensionality and detection rate trade-off

7. CONCLUSION

Our proposed scheme combines boosting with nonlinear subspace of anisotropic multiresolution analysis for robust object detection. Anisotropic multiresolution analysis supports sparse representation of singularities along curves that play a key role for improved object detection and classification. A unique image representation is proposed based on nonlinear subspace to minimize computational load and improve detection accuracy. Our proposed method has been tested for a wide variety of image sets and promising results are achieved. The classification obtained using our proposed scheme is capable to handle illumination variations, view point and appearance changes, cluttered background and noisy measurements.

ACKNOWLEDGEMENT This work was supported in part by the Canada Chair Research Program and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

[1] Y. Freud, R.E. Schapire, ‘A decision theoretic generalization of on-line learning and an application to boosting’, Journal of Computer and system sciences 55, 119-139 (1997)

[2] R. Meir, Gunnar Ratsch, ‘An introduction to boosting and leveraging,________.

[3] L.G. Valiant, ‘A theory of learnable’, Artificial intelligence and language processing, Vol 27, 134-142, 1984

[4] Candès E.J., Demanet L., Donoho D.L., Ying.L.: Fast discrete curvelet transforms. Multiscale Model. Simul., pp. 861--899 (2005)

[5] Saad Ali, Mubarak Shah, A supervised learning framework for generic object detection in images, International conference on Computer Vision, China, Oct. 2005

[6] H. Schneiderman, T. Kanade, Object detection using the statistics of parts, International Journal of C.V., pp 151-177, 2004

[7] E. Bart, E. Byvatove, S. Ullman, View-invariant recognition using corresponding object fragments, ECCV, Prague, May 2004

[8] R. Furegus, P. Perona, A. Zisserman, Object class recognition by unsupervised scale-invariant learning, CVPR 2003

[9] A. Opelt , M. Fussenegger, A. Pinz, P. Auer, Weak hypothses and boosting for generic object detection and recognition’, ECCV 2004

[10] A. Opelt, A. Pinz, A. Zisserman, Incremental learning of object detectors using a visual shape alphabet, CVPR 2006

[11] P. Viola, M. Jones, Rapid object detection using a Boosted cascade of simple feature, CVPR 2001

[12]R. Lienhart, A. Kuranov, V. Pisarevisky, Empirical analysis of detection cascades of Boosted classifiers for rapid object detection, MRL technical report 2002.

[13] J. Shotton, A. Blake, R. Cipolla, Contour based learning for object detection, ICCV, Beijing China, October 2005

[14] J. Amores et. al., Boosting contextual information in content-based image retrieval, ACM SIGMM, 2004

[15] Do M.N., Vetterli M.: The contourlet transform: An efficient directional multiresolution image representation. IEEE trans. on image processing, 14(12), pp. 2091--2106 (2005)

[16] Do M.N., Vetterli M.: The finite ridgelet transform for image representation, IEEE trans on image processing. 12(1), pp. 16--28, (2003)

[17] B. Scholkopf et. al.: Nonlinear component analysis as a kernel eigenvalue problem, Neural computation, 1998