BOOSTED MULTI-CLASS OBJECT DETECTION WITH PARALLEL HARDWARE IMPLEMENTATION FOR REAL-TIME APPLICATIONS

Yao-Tsung Yang[2]*,and Ching-Te Chiu*[1]

Department of Computer Science, National Tsing Hua University, Taiwan ,R.O.C[1]
Institute of Communications Engineering, National Tsing Hua University, Taiwan ,R.O.C[2]
Email:ctchiu@cs.nthu.edu.tw[1]

ABSTRACT

Real-time multi-class object detection becomes popular for various applications such as vehicle vision systems, computer vision and image processing. Boosted cascades achieve fast and reliable object detection for one object class, but require parallel usage of multiple cascades for multi-class detection. The multi-class capable cascade splits the root-cascade into sub-cascades iteratively until each sub-cascade contains one class. That requires a huge number of classifiers in the generated hierarchy of interlinked cascades. In this paper, we propose a boosted multi-class object cascade that only splits one class object from the upper-level-cascade when building the sub-cascades. Since only once class object is split so we can reduce the number of classifiers in each stage. From the simulation results, the boosted multi-class object detection can reduce 46% weak classifiers compared to the multi-class capable cascade for the MIT CBCL database. The proposed method achieves high detection rate(95.54%) and low false positive rate(1.94%). We implement our proposed algorithm with a parallel architecture to accelerate the detection operation using TSMC 90nm CMOS technology. The implementation results show that the design achieves an operation frequency of 100MHz of processing images of 30 fps with size 160×120.

Index Terms— Multi-class Object Detection, weak classifiers, Boosted Cascade

1. INTRODUCTION

Object detection is an important operation for many applications. For example, with the development of automotive electronic technology, detecting various objects and warning the driver of potential dangers becomes one feasible automotive tasks. In automotive systems, object detection has to be both, reliable and fast [1].

Object detection in machine learning generally is divided into two categories: the generative and discriminative model. Generative classifiers learn a model of the joint probability, $p(x, y)$, of the inputs x and the label y. The Bayes rules are used to calculate the predictions $p(y|x)$ and the most likely label y. Discriminative classifiers model the posterior $p(y|x)$ directly, or learn a direct map for inputs x to the class labels (e.g., AdaBoost, SVM).

The probability boosting tree [2] [3], the multi-class capable cascade [1], and boosted cascade [4] [5] belong to the discriminative model for multi-class object detections. However, one of the main problems for multi-class object detection is the huge number of weak classifiers required during detection process. For the Probability Boosting Tree (PBT) [2] [3], the PBT automatically constructs a tree in which each node combines a number of weak classifiers into a strong classifier (conditional posterior probability) in the learning stage. In the testing stage, the conditional probability is computed at each tree node based on the learned classifier, which guides the probability propagation in its sub-trees. The top node of the tree therefore outputs the overall posterior probability by integrating the probabilities gathered from its sub-trees. However, huge amount of weak classifiers in PBT are needed, so it requires large memory during implementation.

The multi-class capable cascade (MCCC) method [1] discriminate multiple classes against background in the first stages of the cascade. each classifier is trained with Adaboost to select the shared features. The final multi-class capable cascade consists of a hierarchy of interlinked cascades [1]. The root-cascade is split into sub-cascades that can be responsible for one or more object classes. Splitting is repeated until each class is represented by its own sub-cascade. A large number of weak classifiers and stages are needed due to parallelism.

The boosted cascade algorithm is the extension of the AdaBoost algorithm that selects a small number of critical visual features from a larger set and yields extremely efficient classifiers [4] [5]. The boosted cascade algorithm, minimizing the number of weak classifier and achieving high detection rate and low false positive rate, compared to above other two algorithms. However, the traditional boosted cascade is used for a single class object detector.

In this paper, we combine the concept of the boosted cascade and multi-class capable cascade to achieve real-time multi-class object detection with reduced number of classi-

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

fiers. With our proposed boosted multi-class object cascade, we reduce the number of weak classifiers by 46% in MIT CBCL database. The implemented boosted multi-class object cascade chip achieves an operation frequency of 100MHz when processing images of 30 fps with size 160 × 120 with core area of 1.21 mm^2.

The rest of this paper is organized as follows. Section II introduces the boosted multi-class object cascade detection. In Section III, we present the performance evaluation and memory cost analysis. Section IV presents the hardware architecture and implementation results. Section V gives a brief conclusion.

2. BOOSTED MULTI-CLASS OBJECT CASCADE (BMCOC)

2.1. Overall Flow

To detect multi-class object, multiple instances of different cascades are necessary. This increases computation time with every new requested object class. The fundamental idea behind the multi-class capable cascade [1] is to discriminate multiple classes against background in the first stages of the cascade. Fig. 1 (a) shows the structures of the MCCC.The final multiple object capable cascade consists of a hierarchy of interlinked cascades. After jointly examining the stages, the root-cascade is split into sub-cascades. Each sub-cascade can be responsible for one or more object classes. To separate each object class from all others, splitting is repeated until each class is represented by its own sub-cascade.

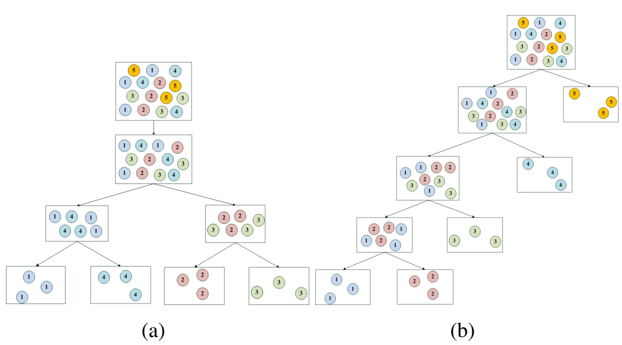


Fig. 1. (a) MCCC structure and (b) Boosted Multi-class Object Cascade Structure.

We propose a boosted multi-class object cascade that only splits one class object from the upper-level-cascade when building the sub-cascades. Since only once class object is split so we can reduce the number of classifiers in each stage. The boosted multi-class object cascade (BMCOC) is shown in Fig. 1(b). The boosted cascade in all nodes that can easily remove negative samples to achieve high accuracy.

In the training stage, AdaBoost is used to select the weak classifiers for each stage. Similar to the MCCC, we discriminate multiple classes against background in the first stages of the cascade. Using positive samples of all object classes in the first stage of the cascade allows AdaBoost to extract representative weak learners for all classes at once. The chosen weak learners share common features for discrimination against the background. After removing the negative background images in the root-cascade, we extract one class object in the next stage cascade. The positive samples of one sub-cascade are explicitly used as negative samples for all other sub-cascades.

In the testing stage, we define a search window with different size and aspect ratio. Then the window scans through a testing image with a moving step size. Fig. 2 shows the algorithm of the BMCOC. Harr-like feature of the sub-window is extracted based on the learned weak classifiers at each cascade and the strong classifier is calculated from a few weak classifiers. The hypothesis is calculated and used for detection judgement. At the root-cascade, all negative images are removed as much as possible. At the next stage, a hypothesis is calculated and formed to detect one class object. If a hypothesis reaches a split, this hypothesis is used to detect one class object. If the result does not belong to the class object, then the process moves on the next sub-cascade for other class objects until it reaches the end of the cascades.

- ➢ Integral image and Harr-like feature extraction on the sub-window of the test image
- ➢ Calculate weak classifiers ($h_{i,j}(v_k)$) and the strong classifier ($H_T(v_k)$) on the node and class estimation

for class $= 1: K-1$
  for $s = 1$: stage_number
    T = # of weak classifiers in sth stage

$$h_{i,j}(v_m) = \begin{cases} 1, \text{if } f_j(v_m) < \theta_{i,j} \text{ and } p_{i,j} = 1, \\ \text{ or } f_j(v_m) \geq \theta_{i,j} \text{ and } p_{i,j} = 0 \\ 0, \text{otherwise} \end{cases}$$

where $\theta_{i,j}$ and $p_{i,j}$ is the j^{th} feature value and polarity of the image i
where $f_j(v_m)$ is the j^{th} feature value of the image m

$$H_T(v_m) = \begin{cases} 1, \text{if } \sum_{t=1}^{T} \alpha_t h_t(v_m) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0, \text{otherwise} \end{cases}$$

```
end
  if class == K − 1
    if H_T(v_m) == 1
      estimateclass(v_m) = K − class
      break
    end
  else
    if H_T(v_m) == 0
      estimateclass(v_m) = K − class
      break
    end
  end
end
```

Fig. 2. The boosted multi-class object cascade detection algorithm.

2.2. Performance Evaluation

Here we use MATLAB as the software simulation tool. In the training procedure, the rectangle feature set and a training set, including 4140 car samples, 4161 pedestrian samples, 4575 traffic sign samples, and 15000 negative samples, are used to train multiple object detection. In the testing procedure, the databases, including MIT CBCL car database, MIT CBCL pedestrian database, and GTSRB traffic sign database, are used to compare detection rate and false positive rate. The MIT CBCL car database has 516 images with size of

128×128. The number of images in the MIT CBCL pedestrian database, and the GTSRB traffic sign database are 924 of size 64×128 and 1000 with scalable size.

According to the MCCC algorithm, the single stage cascade removes negative objects in the root node and the parallel cascades are used to perform multiple object detection in following nodes. The total number of weak classifiers is 690 with 25 stages for MIT CBCL car database. Table 1 is the comparison of MCCC and our proposed BMCOC. The comparisons in Table 1 include number of weak classifier, number of stage, detection rate and false positive in different databases. Our boosted structure reduces the number of weak classifiers with fewer stages and has higher detection rate and lower false positive rate in different databases.

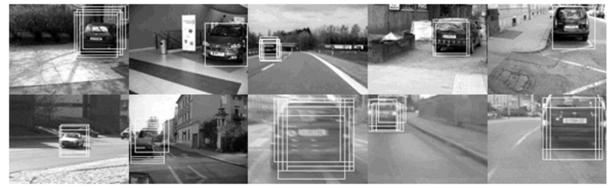
Table 1. The comparison between MCCC and BMCOC.

	MCCC	BMCOC
Number of weak classifiers	690	369
Number of stage	25	15
Detection rate (Car)	94.38% (487/516)	95.54% (493/516)
False Positive rate (Car)	3.1% (16/516)	1.94% (10/516)
Detection rate (Pedestrian)	91.88% (849/924)	93.72% (866/924)
False Positive rate (Pedestrian)	7.14% (66/924)	5.52% (51/924)
Detection rate (Sign)	97.2% (972/1000)	98.3% (983/1000)
False Positive rate (Sign)	2.3% (23/1000)	1.7% (17/1000)

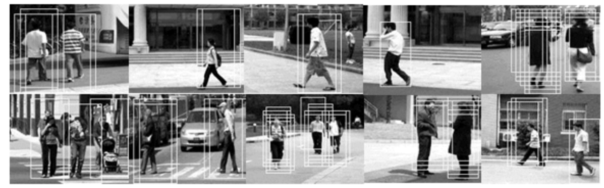
In the MCCC algorithm, it reduces computation time and achieves multiple object detection; however, a large number of weak classifiers and stages is needed. Our proposed boosted multi-class object cascade structure reduces the number of weak classifiers to 369 for the car database and decreases the stage number to 15. Our proposed BMCOC has higher detection rate and lower false positive rate than the MCCC in different databases.

PASCAL car database [6] and Penn-Fudan pedestrian database [6] are also used to test the boosted multi-class object cascade. Fig. 3 shows the simulation results and the detected objects are marked as white blocks. The detection rate of this system achieves 93.75% with 6.25% false positive rate for the car database and 93.27% detection rate with 5.64% false positive rate with pedestrian database.

The memory cost is one of the most important issue because it influences the area, power, and speed. We compare the memory cost of multiple object detection methods in Table 2. The difference between the MCCC and the BMCOC is the number of weak classifier. The memory size of the weak classifier is calculated by the number of feature information bit $\times$ the number of weak classifier. The number of feature information bit is set to 41 in our simulation case. Our proposed method reduced the memory buffer to 13.161×10^3 bits compared to the original MCCC method.



(a)



(b)

Fig. 3. The simulation result (a) PASCAL car database and (b) Penn-Fudan pedestrian database.

Table 2. The memory cost comparison with image size of 160x120)

Memory type	MCCC	BMCOC
Frame buffer	153.6×10^3 bits	153.6×10^3 bits
Integral images	4096 bits	4096 bits
Total weak classifiers	28.29×10^3 bits	15.129×10^3 bits
Total memory	185.986×10^3 bits	172.825×10^3 bits

3. HARDWARE IMPLEMENTATION

The overall architecture of our proposed boosted object detection process is shown in Fig. 4. There are two main units in the system, the integral image processing unit and the cascade detection unit. The integral image processing unit stores a sub-window (16×16) of the input image in the frame buffer.

The integral image calculator then computes the integral values of the sub-window image and stores the results in the integral buffer. The cascade detection processing unit stores the weak classifiers in ROMs. The Haar-like feature extractor extracts reference points based on features stored in the weak classifiers, and the cascade feature detector then updates the feature value and compares it with the expected values stored in the weak classifiers until it detects a row of sub-windows. After detecting all sub-windows, the image is down-sampled to a smaller size and scanned from the beginning. The processes are repeated until the down-sampled image is smaller than 16×16 [10] [7].

The main challenges of this architecture are the large storage requirements for the number of weak classifiers, and the processing time required for feature access and detection. We exploit the boosted multi-class object cascade to reduce the number of weak classifiers and exploit the parallel processing structure to increase processing speed for the cascade detection. We add a classifier selection signal to achieve multiple object detection in cascade detection unit.

Table 3. Comparison of different boosted cascade detection system.

Design reference	Technology (nm)	Freq (MHz)	Total area(mm^2)	Speed (fps)	Frame size	Detection rate	Database
This work	TSMC 90	100	2.4336	30	160×120	93%	Car
[7]	TSMC 180	83.3	5.0625	190	160×120	91%	Face
[8]	CMOS 90	54	2.1	8	320×240	81%	Face
[9]	TSMC 65	800	88	133	320×240	95%	Face

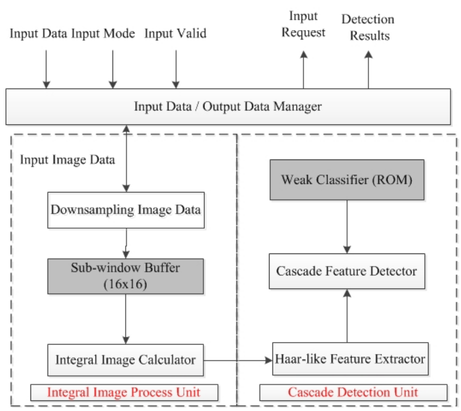


Fig. 4. The boosted multi-class object cascade detection architecture.

The object detection architecture is synthesized and runs at the clock rate of 100 MHz. Three ROMs are used for the storage of the weak classifiers, and the sizes for each on-chip ROM is 128×41 bits.

The proposed object detection architecture executes the auto place and route with soc encounter under TSMC 90nm CMOS technology and the results are summarized in Table 4. The core size is $1.21mm^2$. The on-chip memory is 1968 bytes for three synchronous ROM files. This design can operate at 100 MHz and process 160×120 gray scale images at speeds of up to 30 fps with a higher detection rate (93%)than [8]. Fig. 5 shows the layout of the proposed design with staggered IO pads.

Table 4. The specification table.

Name	Design and Implementation of Boosted Multi-class Object Detection
Technology	TSMC 90 nm
Chip Size (Core Only)	$1.1 \times 1.1mm^2$
Chip Size (With Bond)	$1.56 \times 1.56mm^2$
Gate Count	395,967
On-chip Memory	$41bits \times 128 \times 3 = 1968$ bytes Synchronous ROM Files
Speed	100 MHz @ 160×120
Power	37.3869 mW
Test Coverage	95.3%
Number of Input Ports	82 ($PDIDGZ\_33$)
Number of Output Ports	18 ($PDO02CDG\_33$)
Number of Total Ports	100
Package	CQFP100

Table 3 compares the performance of the object detection system and other boosted cascade detection designs in databases of different object types. We achieve higher detection rate and operation frequency compared to [8] in the 90nm technology.

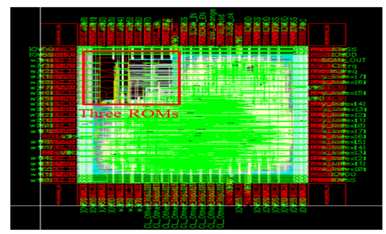


Fig. 5. Layout of the design; the chip core size is $1.1 \times 1.1mm^2$.

4. CONCLUSION

In this paper, we we propose a boosted multi-class object cascade that only splits one class object from the upper-level-cascade when building the sub-cascades. Since only once class object is split so we can reduce the number of classifiers in each stage. We implement the multiple object boosted cascade detection in TSMC 90nm technology. The experimental results show that the BMCOC achieves high detection rate (95.54%) and low false positive rate (1.94%) in MIT CBCL car database. This design can achieve operation frequency of 100MHz, processing images of 30 fps with size 160×120, and with core area of 1.21 mm^2. Finally, the multiple object boosted cascade achieves higher detection rate (93%) and operation frequency (100 MHz) compared to [8].

5. REFERENCES

[1] I. Kallenbach, R. Schweiger, G. Palm, and O. Lohlein, “Multi-class object detection in vision systems using a hierarchy of cascaded classifiers,” Intelligent Vehicles Symposium, pp.3 83-387, Tokyo, Japan, 2006.

[2] Z. Tu, “Probabilistic boosting-tree: learning discriminative models for classification, recognition, and clustering,” IEEE International Conference on Computer Vision, vol. 2, pp. 1589–1596 Vol. 2, 2005.

[3] Z. Tu, K. Narr, P. Dollar, I. Dinov, P. Thompson, and A. Toga, "Brain anatomical structure segmentation by hybrid discriminative/generative models," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 4, pp. 495–508, 2008.

[4] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Confernce on Computer Vision and Pattern Recognition, vol. 1, pp. I.511-518, 2001.

[5] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119 – 139, 1997.

[6] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, Jun. 2010.

[7] C.-R. Chen, W.-S. Wong, and C.-T. Chiu, "A 0.64 mm real-time cascade face detection design based on reduced two-field extraction," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 19, no. 11, pp. 1937–1948, 2011.

[8] Y. Hanai, Y. Hori, J. Nishimura, and T. Kuroda, "A versatile recognition processor employing haar-like feature and cascaded classifier," International Conference on Solid State and Circuit Conference, pp. 148–149, 149a, 2009.

[9] C. Kyrkou and T. Theocharides, "A flexible parallel hardware architecture for adaboost-based real-time object detection," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 19, no. 6, pp. 1034–1047, 2011.

[10] W.-S. Wong, C.-R. Chen, and C.-T. Chiu, "A 100mhz hardware-efficient boost cascaded face detection design," International Confernce on Image Processing, pp. 3237–3240, 2009.