

EARLY TERMINATING ALGORITHMS FOR ADABOOST BASED DETECTORS

Arif Mahmood and Sohaib Khan

Department of Computer Science, School of Science and Engineering, LUMS

ABSTRACT

In this paper we propose an early termination algorithm for speeding up the detection phase of the Adaboost based detectors. In the basic algorithm, at a specific search location, the AdaBoost ensemble response is computed as monotonic decreasing function of weak learners. As more weak learners are evaluated, the response either decreases or remains the same. As soon as the current response becomes lower than the AdaBoost global threshold, remaining computations may be skipped without any loss of accuracy. We further extend the basic algorithm by integrating it with the Non Maxima Suppression (NMS) process. Any candidate location may be discarded, as soon as its current response becomes lower than another candidate location, within the same non-maxima suppression window.

In our experiments, our proposed algorithm has been found to be an order of magnitude faster than the traditionally used AdaBoost detector, for the application of edge-corner detection. Speedup comparisons are also done with other three well known edge corner detectors. The early terminated AdaBoost detector has been found to be significantly faster than all three of these detectors.

Index Terms— Early Termination Algorithms, AdaBoost, Object Detection, Edge-Corner Detection

1. INTRODUCTION

After the seminal work of Viola and Jones on real time face detection using AdaBoost algorithm [1], [2], the face detection problem has been well explored by many other researchers as well [3], [4], [5], [6], [7]. In all of these techniques, fast speedup have been obtained by exploiting the fixed pattern of a human face. For example, the most extensively used principle is: if there are no eyes, there is no face. Unfortunately, such rules cannot be made for objects which do not possess a fixed orientation or a fixed pattern. The early termination algorithm proposed in this paper is generic and applicable to the detection of any type of objects.

In the basic algorithm, each candidate location is initialized with the total weight of the trained ensemble. If a weak learner classifies the current location as a non-object, the weight of that learner is subtracted from the current total weight. As more learners are processed, the weight of the

candidate location monotonically decreases, and as soon as the current weight becomes less than the AdaBoost global threshold, that location can never become a positive instance, therefore further calculations may be skipped and the location may be discarded.

In order to suppress multiple responses to the same object, only local maxima in each locality has to be retained, while the local non-maxima candidates have to be suppressed to zero, a process known as Non-Maxima-Suppression (NMS). We reduce the computations at local-non-maxima candidate locations by developing the *Early Non-Maxima Suppression* (ENMS) algorithm. In ENMS algorithm, we partially compute the AdaBoost detector response at all candidate locations. In each local NMS window, we choose the candidate location with the best partial result, and compute the final detector response at that location. If this final response is larger than the AdaBoost classification threshold, then for the remaining candidate locations in that NMS window, the early termination threshold is raised to the final value of the local maxima. That is, in a specific NMS window, a candidate location will be discarded as soon as the detector response falls below the local maxima or below the AdaBoost classification threshold, whichever is larger. The ENMS algorithm is helpful in reducing the redundant computations done at local non-maxima candidate locations.

The proposed early termination algorithm is incorporated within our previous implementation of the AdaBoost based edge-corner detector [8]. The quality of the detected edge-corners has remained exactly the same, while the speedup over the original algorithm is more than an order of magnitude. We have also compared the quality and speedup of the edge-corners detected by the Adaboost detector with three other detectors including KLT detector, Harris detector [9] and Xaio's detector [10]. We find that the edge-corners detected by AdaBoost detector are of comparable quality as the KLT, Harris and Xaio detectors while the execution time speedup is 2.72 times over KLT, 10.52 times over Harris and 44.88 times over Xaio's detector.

2. RELATED WORK

The details of the AdaBoost algorithm may be found in any text on machine learning and the details of edge-corner detection using AdaBoost algorithm may be found in our earlier

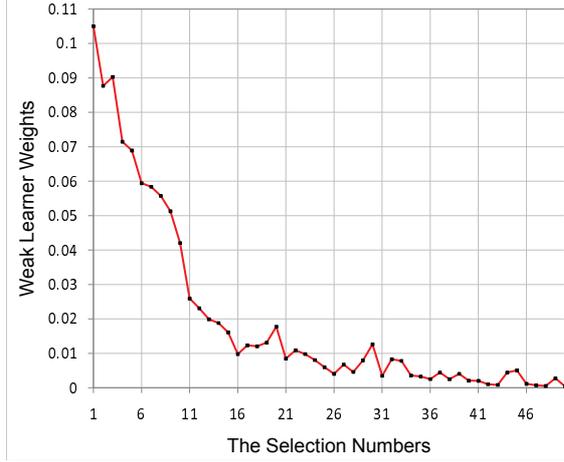


Fig. 1. The weights of weak learners are not always in decreasing order with respect to the selection number. Therefore, after training phase, the ensemble should be sorted in decreasing order of weights.

work [8]. For completeness, the detection phase of AdaBoost algorithm is briefly described as used in [2].

Suppose the trained ensemble of weak learners consists of m learners, $\{f_1, f_2, f_3, \dots, f_m\}$, ordered in the descending order of weights: $\{\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \dots \geq \alpha_m\}$ (Figure 1). At a specific candidate location r_{i_o, j_o} , where (i_o, j_o) are the coordinates of first pixel of the search window, AdaBoost detector response is given by:

$$\Lambda(r_{i_o, j_o}) = \sum_{k=1}^m \alpha_k l_k(r_{i_o, j_o}), \quad (1)$$

where $l_k(r_{i_o, j_o})$ is the label of r_{i_o, j_o} as predicted by the learner f_k . $l_k(r_{i_o, j_o})$ may have only two values:

$$l_k(r_{i_o, j_o}) = \begin{cases} 1 & \text{if prediction is Object,} \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

After evaluating $\Lambda(r_{i_o, j_o})$ in the whole search space, the labeling process starts: the search locations where $\Lambda(r_{i_o, j_o})$ is larger than the AdaBoost global threshold \mathcal{G}_t , are labeled as objects, while the remaining locations are labeled as non-objects.

$$L_m(r_{i_o, j_o}) = \begin{cases} 1 & \text{if } \Lambda(r_{i_o, j_o}) \geq \mathcal{G}_t, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where $L_m(\cdot)$ is the final label of a search location. The AdaBoost global threshold \mathcal{G}_t , is defined as:

$$\mathcal{G}_t = T_\alpha \sum_{k=1}^m \alpha_k, \quad (4)$$

where $1.0 \geq T_\alpha \geq 0.0$.

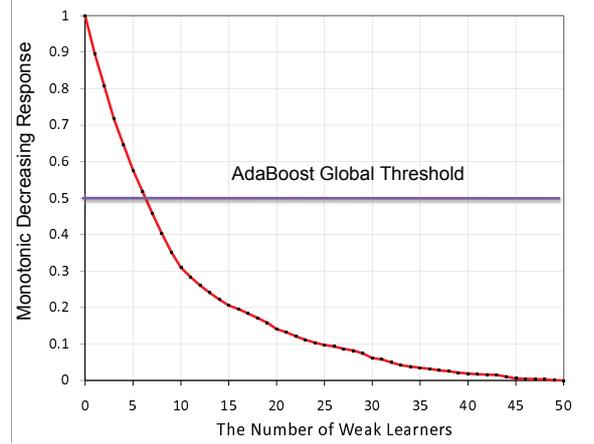


Fig. 2. Monotonic decreasing ensemble response at a non-object location. As more and more weak learners are evaluated, the response decreases accordingly. Once the response is below the AdaBoost global threshold, it can never move up.

Once labeling process is complete, the Non Maxima Suppression (NMS) process has been followed to suppress multiple responses to the same object.

3. ADABOOST GLOBAL THRESHOLD BASED EARLY TERMINATION ALGORITHM

In our proposed algorithm, the current candidate search location is initially assigned the maximum possible AdaBoost detector response, w_m :

$$w_m = \sum_{k=1}^m \alpha_k, \quad (5)$$

then starting with the weak learner f_1 , with maximum weight α_1 , in the trained ensemble, we keep on evaluating learners in the order of decreasing weights: $\{\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \dots \geq \alpha_m\}$. If a learner predicts the current search location as object, we take no action; however if the predicted label is 0, we subtract the weight of that learner from the current value of response. Therefore, the detector response, after processing $i < m$ learners is given by:

$$\Lambda^i(r_{i_o, j_o}) = w_m - \sum_{k=1}^i \alpha_k (1 - L_k(r_{i_o, j_o})), \quad (6)$$

In this form, the AdaBoost detector response has become monotonic decreasing function over the number of processed learners. After processing each learner, either the response remains same or decreases (Figure 2). As soon as the current response, $\Lambda^i(r_{i_o, j_o})$, falls below the global threshold \mathcal{G}_t , computation of the remaining learners becomes redundant and may be skipped without any loss of accuracy.



Fig. 3. Image dataset consisting of still images of varying details, is used for detection speedup comparison. The size of each image is 2304×3072 pixels.

Since most of the ensemble weight is generally concentrated in the first few learners, for non-object locations the detector response rapidly decreases to less than global threshold. Therefore the average number of learners to be evaluated at any location reduces to a very small number, rendering the detection speed significantly faster, without any loss of accuracy.

4. EARLY NON-MAXIMA SUPPRESSION ALGORITHM

Non Maxima Suppression (NMS) process has been commonly used to suppress multiple detections corresponding to the same real world object. Assuming that the detector response surface is smooth and considering an NMS-window of appropriate size around the current search location, the non-maxima suppression may be described as: if the detector response at current location is not maximum within the NMS window, current location will be labeled as non-object, otherwise it will remain labeled as object. That is, the label of current location r_{i_o, j_o} is given by:

$$L_m(r_{i_o, j_o}) = \begin{cases} 0 & \text{if } \Lambda(r_{i_o, j_o}) \leq \Lambda(r_{i'_o, j'_o}), \\ 1 & \text{otherwise,} \end{cases} \quad (7)$$

where $\Lambda(r_{i_o, j_o})$ is the detector response at the current location and $\Lambda(r_{i'_o, j'_o})$ is the maximum detector response at any other location within the same NMS-window.

The early termination algorithm discussed in the last section may be integrated with NMS process to further reduce redundant computations. If in a locality, the local maxima $\Lambda(r_{i'_o, j'_o})$ is significantly higher than the global threshold, \mathcal{G}_t , then all search locations in that locality having response less than $\Lambda(r_{i'_o, j'_o})$ are non-object locations. Therefore computations at the current location will stop as soon as the detector

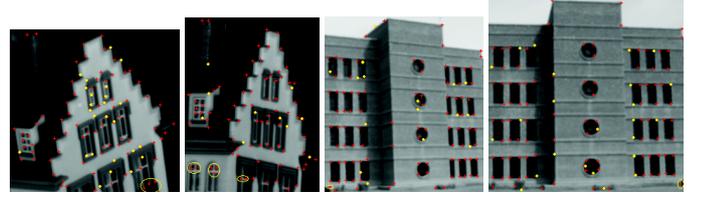


Fig. 4. View Invariance of early terminated AdaBoost detector: (a)-(b) Two views of LUMS library building (c)-(d) two views from hotel sequence. Red crosses show the AdaBoost detections, yellow dots show missing detections and yellow circles show the detections outside the viewing area.

response falls below the local maxima, $\Lambda(r_{i'_o, j'_o})$.

$$L_m(r_{i_o, j_o}) = \begin{cases} 0 & \text{if } \Lambda^i(r_{i_o, j_o}) \leq \max(\Lambda(r_{i'_o, j'_o}), \mathcal{G}_t) \\ 1 & \text{otherwise,} \end{cases} \quad (8)$$

where $\Lambda^i(r_{i_o, j_o})$ is the detector response at current location for $i < m$ learners.

In order to find local maxima, we compute AdaBoost detector response over all search locations for first p learners such that the sum of weight of these first p learners, w_p , satisfies the following bound:

$$w_p \geq (1 - T_\alpha) \sum_{k=1}^m \alpha_k, \quad (9)$$

which means that $w_e - w_p \leq \mathcal{G}_t$, the global threshold. Partial response over p learners is given by:

$$\Lambda^p(r_{i_o, j_o}) = w_e - \sum_{k=1}^p \alpha_k (1 - L_k(r_{i_o, j_o})). \quad (10)$$

Search locations where $\Lambda^p(r_{i_o, j_o}) \leq \mathcal{G}_{th}$ are labeled as 0, while the search locations where $\Lambda^p(r_{i_o, j_o}) \geq \mathcal{G}_{th}$ are still undecided.

$$L_m(r_{i_o, j_o}) = \begin{cases} 0 & \text{if } \Lambda^p(r_{i_o, j_o}) \leq \mathcal{G}_t, \\ u & \text{otherwise,} \end{cases} \quad (11)$$

where u means label is yet undecided. Upon these undecided locations, ENMS algorithm is implemented as follows: if partial response at current search location, $r_{i'_o, j'_o}$, is larger than the partial response at all search locations within the current NMS window, calculate the complete response over m learners at the current search location:

$$\Lambda^m(r_{i'_o, j'_o}) = \Lambda^p(r_{i'_o, j'_o}) - \sum_{k=p+1}^m \alpha_k (1 - L_k(r_{i'_o, j'_o})). \quad (12)$$

If $\Lambda^m(r_{i'_o, j'_o}) \geq \mathcal{G}_t$, all remaining search locations in the current NMS window having partial response less than

Table 1. Execution time (sec) of AdaBoost, Early-terminated AdaBoost, KLT, Harris, and Xaio edge-corner detectors.

Img ID	AdaBoost	EAdaBoost	KLT	Hrs	Xaio
1	48.17	1.73	5.86	22.57	34.44
2	48.12	1.46	5.89	22.86	9.82
3	48.76	2.76	5.89	22.91	244.15
4	48.83	3.96	6.02	22.93	300.16
5	48.22	1.34	5.83	22.96	5.366
6	48.19	1.84	6.00	22.93	60.20
7	48.23	1.77	5.81	22.47	54.60
8	48.35	2.27	5.95	22.90	130.24
9	48.03	1.84	5.81	22.87	48.81
10	48.15	2.69	5.91	22.86	86.23
Mean	48.30	2.16	5.90	22.83	97.40

$\Lambda^m(r_{i'_o, j'_o})$ will be labeled as non-objects:

$$L_m(r_{i_o, j_o}) = \begin{cases} 0 & \text{if } \Lambda^p(r_{i_o, j_o}) \leq \Lambda^m(r_{i'_o, j'_o}), \\ u & \text{otherwise,} \end{cases} \quad (13)$$

where u means label is yet undecided. At each of these undecided locations, further learners are evaluated until that location is labeled as non-object or final response is computed. In any locality, as soon as a maxima larger than the previous known maxima is found, the previous best location is labeled as non-object. When all locations are exhausted, the last undecided location in each locality will be labeled as object.

5. EXPERIMENTS AND RESULTS

The speedup generated by the proposed early termination algorithm is compared with the previous AdaBoost implementation [8] as well as KLT, Harris and Xaio's detectors. The speedup comparison is done on a dataset of ten images shown in Figure 1, each of size 2304×3072 pixels, having varying levels of details and different types of contents. Therefore the number of detected edge corners also largely varies across images. Processing in Xaio and EAdaBoost detectors increases if the number of detected edge-corners increase, and therefore these algorithms take more execution time, while in AdaBoost, Harris and KLT detectors, the processing remains almost same regardless the number of detected edge corners. The thresholds for each algorithm are set such that the number of detected edge corners remain approximately same for a specific image. In case of EAdaBoost detector, number of learners initially evaluated upon all search locations are $(p + 3)$, where p was computed from the global threshold. Three extra learners were evaluated so that the number of locations requiring non-maxima suppression should be as minimum as possible.

The speedup comparison is done on HP Pavilion Notebook PC with Intel Core 2 Duo CPU 2.0 GHz and 2GB RAM. On the average, the early terminated AdaBoost detector 2.72

times faster than KLT detector, 10.52 times faster than Harris detector, is 22.36 times faster than the traditional AdaBoost detector and 44.89 times faster than Xaio's detector (Table 1).

Invariance comparison of EAdaBoost based detector with the three other detectors is made in the presence of view changes, scale changes, blurring noise, additive Gaussian noise and rotation (Figure 2). The quality of AdaBoost detector has been found to be comparable with the other detectors. Note that the quality of AdaBoost detector with and without early termination has remained exactly same.

6. CONCLUSION

Early termination algorithms to speedup the AdaBoost based object detection has been proposed. The proposed algorithms are incorporated within AdaBoost based edge-corner detector and the speedup obtained over basic implementation is larger than one order of magnitude. The early terminated AdaBoost detector is also compared with some other commonly used detectors and we find that this detector outperforms these detectors by a significant margin.

7. REFERENCES

- [1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE CVPR*, 2001.
- [2] P. Viola and M. Jones, "Robust real time face detection," *IJCV*, 2003.
- [3] R. Vincenzo and U. Lisa, "An improvement of adaboost for face-detection with motion and color information," *ICIAP*, 2007.
- [4] C. Huang, B. WU, Haizhou, and Shihong, "Omnidirectional face detection based on real adaboost," *IEEE ICIP*, 2004.
- [5] D. Cristinacce and T. Cootes, "Facial feature detection using adaboost with shape constraints," *BMVC*, 2003.
- [6] Bo Wu, Haizhou, C. Huang, and S. Lao, "Fast rotation invariant multi-view face detection based on real adaboost," *IEEE FG*, 2004.
- [7] D. Cristinacce and T. Cootes, "Facial feature detection using AdaBoost with shape constraints," *BMVC*, 2003.
- [8] A. Mahmood, "Structure-less object detection using adaboost algorithm," in *Int. Conf. Machine Vision*, December 2007.
- [9] C. Harris and M. Stephens, "A combined corner and edge detector," in *4th ALVEY vis. conf., Univ. of Manch., England.*, 1988.
- [10] J. Xiao and M. Shah, "Two-frame wide baseline matching," *IEEE CVPR*, 2003.