Vision-based Wheelchair Navigation using Geometric AdaBoost Learning

Eun Yi Kim

Visual Information Processing Laboratory, Dept. of Internet and Multimedia Engineering, Konkuk University

1204 New Millennium Building, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, 05029

Seoul, South Korea

eykim@konkuk.ac.kr

Abstract

Recent years, This paper proposes a novel training algorithm called geometric AdaBoost learning, which integrates the local appearance models with the explicit shape model. The proposed algorithm employs a two-stage AdaBoost learning algorithm. The first stage learning is performed to learn the local texture model within local image patches and to produce a confidence map. Based on the confidence values, the high discriminative local patches are selected, and then the global context models between them are trained in the later stage using AdaBoost learning. The proposed algorithm is applied to wheelchair navigation, and the results demonstrate that it outperforms the state-of-the-art algorithms with improvements of 23.3% and 49% in terms of accuracy and speed.

Keywords-Light

# I. Introduction

Navigation that recognizes surrounding environments and avoids collisions is essential in a mobile robot. In the literature, algorithms such as the potential field approach, the vector field histogram and the dynamic window approach have been widely used for obstacle avoidance [1]. These algorithms determine viable paths based on the sensor data measured from their local environments. Thus, their performance is greatly dependent on the sensor capabilities and sensitivity to sensor noise. To address this limitation, machine learning algorithms (MLs) have been intensively investigated in recent years. In [2,3], neural networks (NNs) were used to determine the viable paths in mobile robots, where ultrasonic sensors or the fusion of sensors and cameras were used to measure the environmental information. In addition, a support vector machine (SVM) was employed for vision-based wheelchair robot navigation [4]. Through ML, a robot can automatically learn the experience of human beings in finding viable paths in real situations.

In this paper, we propose a new learning algorithm called geometric AdaBoost learning (GAL) that provides more effective and efficient navigation by fusing local appearance models with global context information. In the GAL, the most discriminative local feature set is first selected, and then the features are combined to form more powerful representations based on geometric relationships. To select the features, we use AdaBoost learning because it is known as a powerful tool for feature selection and assembling classifiers. Here, two-stage AdaBoost learning is employed: 1) the first AdaBoost learning is performed to learn a texture classifier on local image patches and to produce a confidence map; 2) based on the confidence values, the more discriminative local patches are selected and the co-occurrence between them is trained using the second AdaBoost learning stage.

The GAL was applied to vision-based wheelchair navigation, and tested indoors and outdoors with complex environments.

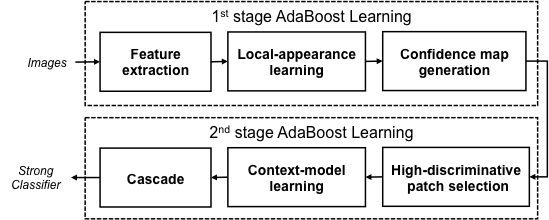


Fig. 1 Overview of the proposed Geometric AdaBoost Algorithm.

# II. Geometric AdaBoost learning

The motivation of this study originated in the basic observation that both local information and global context information should be used for more accurately discriminating a target object from others. Based on this, we developed a new learning algorithm called GAL that selects the simple but highly discriminative features for classification, which can simultaneously ensures efficiency and effectiveness. For this, two stages of AdaBoost learning are used: the first stage is performed for local appearance model learning and confidence map generation, and the second one is for global context model learning. The detail process of GAL is depicted in Fig. 1.

**Feature extraction:** To capture the local texture properties and learn them, the appropriate feature extraction scheme needs to be integrated into AGL. In this study, an occupancy grid map (OGM) was used, which represents the environmental information surrounding a robot, such as the sizes and positions of obstacles. Using online background learning and subtraction, the images sized at 320×240 are transformed into the 32×24 OGMs, where each cell corresponds to one 10×10 block in an input image and has a gray-level value according to the risk level. The detailed process for calculating the OGMs is described in [4].

**Local-appearance learning:** For the proposed method to be robust against scale and occlusion problems, the local image patches with variable sizes are generated from the OGMs. Then, to describe texture properties within a local image patch, we used three Harr-like features (HLFs) that were proposed by Viola and Jones [5], which measure the average intensity differences of the adjacent rectangle regions.

For each local patch x\_i, a weak classifier consists of a feature (f), a threshold (θ) and a polarity (p). When the examples are given, the best classifier for the tth feature is selected using AdaBoost with respect to the weighted error:

|  |  |
| --- | --- |
| . | (1) |

**Confidence map generation:** From all the learned weak classifiers, a confidence map is constructed on a OGM, each cell of which has the confidence value that represents its power in discriminating the target object from others. The value is obtained by aggregating the detection accuracies of all weak classifiers that pass on a cell

|  |  |
| --- | --- |
|  | (2) |

In (2), denotes the number of weak classifiers that pass on , and is the error of its th classifier. The cells with the higher confidences are considered as the regions that have the higher discriminative power to the problem.

**High-discriminative patch selection and context-model learning:** To model context information among local patches, we constructed new HLFs. The feature is composed of a pair-wise or a triplet of local patches that have strong geometric relationships. Then, if all local image patches are used in context modeling, a huge amount of configuration can be generated, which will burden the training process. To reduce the feature pool, only the local patches with the higher confidence values are used. Then, for each local patch x, its confidence is computed by averaging the confidences of all cells that it covers, that is, .

Thereafter, to effectively characterize the geometric relationships among image patches, we modify the geometric constraint of the traditional HLFs. Instead of using image patches that must be adjacent, we use the separated images patches. For example, one image patch is extracted on the left confident zones and others lie on the right zones (see Fig. 3). Then, the output of the proposed features is computed as:

|  |  |
| --- | --- |
|  | (3) |

In (3), x, y and z are some selected image patches, and α, β and γ are the weights to determine the types of context models: for pair-wise models of image patches, and ; for triplet models, and

Even if we use only the highly discriminative patches, there is a huge set of pair-wise and triple models. Among them, the most discriminative features are selected automatically through AdaBoost learning.

**Cascade learning:** With all the selected context features, we construct a strong classifier. Then, to increase detection performance while radically reducing computational time, the strong classifiers learned by the AdaBoost algorithm are cascaded by the method proposed in [5].

Application to vision-based wheelchair robot navigation: In [4], we developed a wheelchair navigation system that provided safe mobility for disabled and elderly people. From real-time image streaming, various types of obstacles were first detected using online background learning, and then the viable paths were determined using SVM-based classifiers. In this experiment, the determination of viable paths was implemented using the proposed GAL algorithm.

Fig. 2 presents some OGMs generated by online background learning, where the more intense a grid cell, the more closely spaced the obstacles. The first two are the results for indoors and the last two are the results for moving obstacles outdoors.

|  |
| --- |
| Macintosh HD:Users:kimeunyi:Desktop:EYKIM:PAPER:Geometric Adaboost:IEE:Figs & Tables:Fig. 2a:스크린샷 2016-09-29 오후 1.12.12.pngMacintosh HD:Users:kimeunyi:Desktop:EYKIM:PAPER:Geometric Adaboost:IEE:Figs & Tables:Fig. 2a:스크린샷 2016-09-29 오후 1.12.40.pngMacintosh HD:Users:kimeunyi:Desktop:EYKIM:PAPER:Geometric Adaboost:IEE:Figs & Tables:Fig. 2a:스크린샷 2016-09-29 오후 1.13.04.pngMacintosh HD:Users:kimeunyi:Desktop:EYKIM:PAPER:Geometric Adaboost:IEE:Figs & Tables:Fig. 2a:스크린샷 2016-09-29 오후 1.13.24.png |

*a*

|  |
| --- |
| Macintosh HD:Users:kimeunyi:Desktop:EYKIM:PAPER:Geometric Adaboost:IEE:Figs & Tables:Fig. 2b:스크린샷 2016-09-29 오후 1.12.27.pngMacintosh HD:Users:kimeunyi:Desktop:EYKIM:PAPER:Geometric Adaboost:IEE:Figs & Tables:Fig. 2b:스크린샷 2016-09-29 오후 1.12.47.pngMacintosh HD:Users:kimeunyi:Desktop:EYKIM:PAPER:Geometric Adaboost:IEE:Figs & Tables:Fig. 2b:스크린샷 2016-09-29 오후 1.13.13.pngMacintosh HD:Users:kimeunyi:Desktop:EYKIM:PAPER:Geometric Adaboost:IEE:Figs & Tables:Fig. 2b:스크린샷 2016-09-29 오후 1.13.32.png |

*b*

Fig. 2 Results of feature extraction

a: Input images

b: Occupancy grid maps generated by online background learning [4].

The wheelchair is controlled by four directions: ‘go-straight’, ‘stop’, ‘turn-left’, and ‘turn-right’. As the standard AdaBoost is designed for binary classification problems, a one-vs-all scheme is adopted in order to directly apply the proposed GAL to a four direction classification: the classification of ‘move’ and ‘stop’ is first performed; then the classification of ‘go-straight’ and ‘turn’ is performed; and finally the classification of ‘turn-left’ and ‘turn-right’ is determined.

For the first round of each classification, the local appearance models within local image patches are learned using AdaBoost, and confidence maps are generated. Fig. 3 is a visualization of the confidence map generated for the classification of ‘move’ and ‘stop’, where only the lower half of the map is shown. The cells appear in red for higher confidence and blue for lower confidence. In most cases, the upper half of an OGM is covered by many static obstacles such as walls and buildings, so it has lower discriminative powers. Thus, it is discarded in the second stage of Adaboost learning for modeling the context features. Among 562 local patches, we selected only 75 patches with higher confidences and used them to model the context features. By our geometric constraints, 105 pair-wise features and 310 triplet features were generated and were used to construct a cascade of strong classifiers.

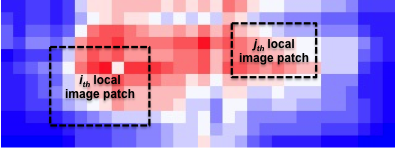


Fig. 3 Confidence map for the decision ‘move’ and ‘stop.’

# III. Experimental results

To assess the validity of the proposed GAL in obstacle avoidance, experiments were performed on images collected both indoors and outdoors with various environments. Thus, 80,000 images were collected over a period of one year at different times. For all images, the ground truths were manually annotated by humans. From the datasets, 1000 items of training data were randomly selected.

For the input image streaming obtained from a camera attached to wheelchair, the OGMs were first produced and then the viable paths were determined using the proposed GAL. Then, to numerically demonstrate the advantages of the proposed algorithm, the results were compared with those of existing methods using NN, SVM and standard AdaBoost. The NN and SVM used the whole of the OGM as input feature vector of the classifiers. In contrast, the standard AdaBoost just used a combination of the local patches within the OGM to construct a strong classifier.

Fig. 4 and Table 1 present a summary of the numerical comparison for the four methods in terms of accuracy and processing time. On average, the four methods had accuracies of 0.86, 0.90, 0.77 and 0.95, respectively. In respect to the processing time, the proposed method was the fastest, followed by the standard AdaBoost and the NN, and finally the SVM. Evidently, the proposed GAL performed best in all environments despite of the fastest processing. In contrast, the standard AdaBoost method, which uses only the local appearance models, had the lowest accuracy, while it guaranteed fast processing. Even though the SVM and NN required more calculations for global context modeling than the AdaBoost, the accuracy was clearly better.

Consequently, the comparison results confirmed that the proposed algorithm can effectively model global context features in a more efficient way and can be successfully applied for real-time obstacle avoidance in a mobile robot.

Fig. 4 Performance comparison for the four methods.

**Table 1: Average frame-processing time in the four methods (ms).**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Classifiers | NN | SVM | AdaBoost | GAL |
| OGM generation | 2.95 | 2.95 | 2.95 | 2.95 |
| Path determination | 0.77 | 2.08 | 0.71 | 0.36 |

# IV. Conclusion

`

This paper presents simple but highly discriminative features for pattern classification by integrating local appearances and the global context together, which can simultaneously ensure effectiveness and efficiency. The proposed method was applied to wheelchair navigation and the results demonstrated its effectiveness.

# References

[1] M. Gillham, G. Howells, “ A dynamic localized adjustable force field method for real-time assistive non-holonomic mobile robotics,” International Journal of Advanced Robotic Systems, 2015, 12, doi: 10.5772/61190

[2] J. Chen, L. Niu, D. Chen, “Research on intelligent wheelchair obstacle avoidance based on AdaBoost,” Applied Mechanics and Materials,2013.312, doi:10.4028/www.scientific.net/AMM.312.685

[3] X. Song, H. Fang, X. Jiao, , Y. Wang, “Autonomous mobile robot navigation using machine learning,” IEEE ICIFS, 2012, pp. 135-140, doi: 10.1109/ICIAFS.2012.6419894

[4] Y. Ji, M. Lee, E. Kim, “ An intelligent wheelchair to enable safe mobility of the disabled people with motor and cognitive impairments,” ECCV 2014 Workshops, 2014, doi: 10.1007/778-3-319-16199-0\_49

[5] P. Viola, M. Jones, “Robust real-time face detection,” International Journal of Computer Vision, 2004, 57(2), pp. 137-154, doi: 10.1023/B:VISI.0000013087.49260.fb