Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Boosting multi-features with prior knowledge for mini unmanned helicopter landmark detection*

Qin-yuan REN†, Ping LI, Bo HAN
(*The State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China*)
†E-mail: qyren@iipc.zju.edu.cn

Received Mar. 14, 2007; revision accepted July 5, 2007; published online Jan. 11, 2008

Abstract: Without sufficient real training data, the data driven classification algorithms based on boosting method cannot solely be utilized to applications such as the mini unmanned helicopter landmark image detection. In this paper, we propose an approach which uses a boosting algorithm with the prior knowledge for the mini unmanned helicopter landmark image detection. The stage forward stagewise additive model of boosting is analyzed, and the approach how to combine it with the prior knowledge model is presented. The approach is then applied to landmark image detection, where the multi-features are boosted to solve a series of problems, such as rotation, noises affected, etc. Results of real flight experiments demonstrate that for small training examples the boosted learning system using prior knowledge is dramatically better than the one driven by data only.

Key words: Boosting, Prior knowledge model, Landmark detection
doi:10.1631/jzus.A071131 **Document code:** A **CLC number:** TP75

INTRODUCTION

Recently, the mini unmanned helicopters, which have been applied to various areas such as search and rescue mission, surveillance and military applications, are attracting a great deal of attention. Especially due to their small size and various fly modes, they can move in more complex environments than the normal fixed aircraft. The landmark detection system is a very important part of the mini unmanned helicopter navigation system. It detects landmark, and supplies location information to the navigation computer. Cameras are usually utilized in the landmark detection system since they are small in size and can provide a rich source of information about the immediate surroundings. In this paper, we focus on the landmark image detection.

The landmark image detection for the mini unmanned helicopter has been an active topic of research in the past few years. Amidi *et al*.(1998) presented a real-time computer vision system based on the circumrotate-pattern recognition algorithm for searching danger stuff landmarks, but the system performance is utterly dependent on the hardware. A special real-time landmark detection problem is discussed by Sharp *et al*.(2001), but the approach cannot be generalized to other landmarks. An algorithm based on Hu invariant moments was used in detection landing pad by Saripalli *et al*.(2002). The three lower-order invariants given in terms of moments are scale, rotation and translation, however, these moment invariants do not carry sufficient information for complex landmark detection.

To build a robust landmark image detection system for mini unmanned helicopter navigation, such as vision-based autonomous landing and agricultural surveillance, several technique challenges will be faced:

(1) In a 3D environment, the same landmark displays different rotation poses from different viewpoints. Therefore, the rotation landmarks need to be detected.

* Project (No. 2006AA10Z204) supported by the National Hi-Tech Research and Development Program (863) of China

(2) The detection approach should be easily generalized to different landmarks.

(3) Real-time processing is required.

In this paper, the landmark detection is considered as a two-class classification problem. Numerous machine learning algorithms are applied and the data driven approaches have gained prevalence. Popular data-driven classification approaches include decision tree, Bayesian classifier, support vector machine (SVM), etc. However, it is often difficult or inefficient to directly apply them to vision applications due to many challenges, such as curse of dimensionality, storage and computation cost, etc. (Zhou *et al*., 2005). Considering these challenges, we propose a classification algorithm based on boosting method, which is fast, simple and easy to program. Boosting is one of the most successful and practical methods that has recently come from the machine learning community (Schapire *et al*., 1996; Harris and Corinna, 1996; Freund and Schapire, 1997). After the theoretic connection between Adaboost and forward stagewise additive model was discovered, series of boosting algorithms were presented (Friedman *et al*., 2000; Schwenk and Bengio, 2000; Bühlmann and Yu, 2003). We will briefly review them in Section 2.

In the mini unmanned helicopter landmark image detection, due to the challenges posed by the flying style, it is difficult to collect sufficient real and various rotation poses landmark images (as training data) acquired from aerial vehicles. However, like many machine learning methods, boosting is entirely data driven. It implies that the boosting cannot be directly utilized to this kind of application. In some particular applications, such as the call classification, to compensate for the lack of training data, the prior knowledge had been tried in (Schapire *et al*., 2005). Prior knowledge can be acquired from several sources, e.g., human experiences, expert knowledge, and manuals. In this paper, it is discussed how to use human crafted knowledge to compensate for the lack of data in building robust image classifiers. We aim to use the prior knowledge to create a crude classification model (template). Then, the model is refined by the statistics of the small set of training data. The details will be described in Section 3.

Section 4 introduces boosting multi-features with the prior knowledge applied in the landmark detection. In addition to the Hu invariant moments, a large set of edge features called Haar-like features are utilized. These features, which provide the complement information from the viewpoints of geometric moments and edge, are combined with the special boosting algorithm presented in Section 3.

In Section 5, experiments are carried out to verify the effectiveness of the approach by detecting three different landmarks in a mini unmanned helicopter landmark detection system. Finally, Section 6 concludes the paper.

BOOSTING ALGORITHM

Machine learning studies the automatic learning techniques to make accurate predictions based on the past observation. The past observation information can be denoted as a set of training examples (x_1,y_1), …, (x_m,y_m). Each x_i is called an instance. In this paper, each x_i will generally be the feature of an image, and each y_i is the label and it indicates the landmark types. For simplicity, we assume that there are only two classes, −1 (no landmark) and +1 (landmark). All training examples (x,y) are selected independently from some distribution D on the $X\times Y$ which is the space of all possible instances and all labels. The goal is to generate a rule that makes the most accurate predictions possible on the new test examples.

Boosting is based on the observation that finding many rough rules of thumb can be a lot easier than finding a single, highly accurate prediction rule. It can be considered as the forward stagewise additive model:

$$f(x)=\sum_{m=1}^{M}\beta_m b_m(x;\gamma_m), \tag{1}$$

where β_m (m=1, 2, …, M) is the expansion coefficient, and $b_m(x;\gamma_m)$ is the simple function (weak learner). Boosting is carried out to estimate the function F: $\mathbb{R}^d\rightarrow\mathbb{R}$, minimizing an expected loss

$$E[C(Y,F(X))],\quad C(\cdot,\cdot):\mathbb{R}\times\mathbb{R}\rightarrow\mathbb{R}^+ \tag{2}$$

based on the training data (x_i, y_i) (i=1, …, n).

Corresponding to different loss functions C, different boosting algorithms are presented. The most prominent examples are:

$$C(y,f) \overset{y\in\{-1,1\}}{=} \begin{cases} \exp(yf) \rightarrow \text{AdaBoost algorithm}, \\ \log_2(1+\mathrm{e}^{-2yf}) \rightarrow \text{LogitBoost algorithm}, \\ (y-f)^2/2 \rightarrow L_2\text{Boost algorithm}. \end{cases} \tag{3}$$

The minimizers of Eq.(2) are

$$F(x)=\begin{cases} \dfrac{1}{2}\log\left(\dfrac{P[Y=1\mid X=x]}{P[Y=-1\mid X=x]}\right), \\ \qquad \text{for AdaBoost and LogitBoost loss}, \\ E[Y\mid X=x]=P[Y=1\mid X=x]- \\ \quad P[Y=-1\mid X=x], \qquad \text{for } L_2\text{Boost loss}, \end{cases} \tag{4}$$

where the conditional probability $P[Y=1|X=x]$, which is dependent of D, can be estimated as:

$$\sigma(f(x))=\begin{cases} P[Y=1\mid X=x]=[1+\mathrm{e}^{-2f(x)}]^{-1}, \\ \quad \text{for AdaBoost and LogitBoost loss}, \\ P[Y=1\mid X=x]=[1+f(x)]/2, \\ \quad \text{for Boost loss}. \end{cases} \tag{5}$$

$F(\cdot)$ and $\sigma(\cdot)$ are estimated by applying greedy search from data via a minimization of the empirical risk (EMR)

$$n^{-1}\sum_{i=1}^{n}(Y_i,F(X_i)). \tag{6}$$

The basic idea of boosting is to build a highly accurate classifier by combining many weak learners $b(x,\gamma_m)$. But, if these learners are already complex, the boosting is prone to overfitting. The other problem must be concerned with is the convergence of boosting, about which many researchers have discussed. The detail can be found in (Schapire, 1990; Schapire *et al*., 1996; Friedman *et al*., 2000; Bühlmann and Yu, 2003).

INCORPORATING PRIOR KNOWLEDGE

This section describes how to use the prior knowledge to modify the boosting algorithm. In a two-classification problem, if the distribution of the D can be estimated, the accurate classifier can be easily built. In fact, the conditional probability $\sigma(\cdot)=\Pr[Y=1|X=x]$ is dependent of the D. So, our task is to build $\sigma'(\cdot)$ which takes advantage of the prior knowledge, and fuse it with the $\sigma(\cdot)$.

$\sigma'(\cdot)$ is built based on difference backgrounds. In this section, we will focus on how to fuse $\sigma(\cdot)$ and $\sigma'(\cdot)$ by using Kullback-Leibler divergence, which has been studied intensively in statistics, communication and information theory, and has been considered as providing distance between different statistical distribution models (Csiszar, 1975). The Kullback-Leibler divergence between $\sigma'(\cdot)$ and $\sigma(\cdot)$ defined on D is given by

$$R(\sigma'\parallel\sigma)=\sigma'\log\frac{\sigma'}{\sigma}+(1-\sigma')\log\frac{1-\sigma'}{1-\sigma}. \tag{7}$$

Put Eq.(7) into the boosting framework, then the loss function can be modified as

$$C'(y,f)=C(y,f)+\alpha R(\sigma'(f)\parallel\sigma(f)), \tag{8}$$

where α is a parameter to adjust the relative importance between $C(\cdot)$ and $R(\cdot)$.

The last thing is to add a 0th weak learner b_0 into Eq.(1). b_0 is a rough model according to the prior knowledge:

$$b_0(x,\gamma_0)=\begin{cases} \dfrac{1}{2}\dfrac{\sigma'}{1-\sigma'}, & \text{in AdaBoost \& LogitBoost}, \\ 2\sigma'-1, & \text{in } L_2\text{Boost}. \end{cases} \tag{9}$$

LANDMARK IMAGE DETECTION

In this section, we shall introduce how to combine two different types of features by using the special boosting algorithm described in the previous section. The landmark images from video stream captured by the camera fixed on a moving mini unmanned helicopter may vary in orientation. Meanwhile, for the complex outdoor environment, landmark images are easily affected by various noises. To solve these problems, multi-features are employed.

The Hu invariant moments are scale, rotation and translation, which are highly suitable for rotation object image detection such as the simple landing pad detection. However, due to the small moments features set, they do not carry sufficient information and are easily affected by noises. On the other hand, Haar-like features as a large set of edge features can carry sufficient information for complex object recognition (Viola and Jones, 2001; Lienhart and Maydt, 2002). But, they are sensitive to the rotation information. In our approach, in addition to Hu invariant moments, we employ Haar-like features to complement the information.

Feature selection

The Hu invariant moment is one of the geometric features. The (p+q)th order moment of an image $I(x,y)$ is given by

$$m_{pq}=\sum_i\sum_j i^p j^q I(i,j), \tag{10}$$

where i, j correspond to the coordinate axes x, y respectively. The central moments of an object are defined as

$$\mu_{pq}=\sum_i\sum_j (i-\bar{x})^p (j-\bar{y})^q I(i,j), \tag{11}$$

where $\bar{x}=m_{10}/m_{00}$, $\bar{y}=m_{01}/m_{00}$, p, q=0, 1, 2, ... The normalized central moments are defined as

$$\eta_{pq}=\mu_{pq}/\mu_{00}^{\beta}, \tag{12}$$

where β=1+(p+q)/2. The three lower-order Hu invariant moments are defined as

$$\phi_1=\eta_{20}+\eta_{02}, \tag{13}$$

$$\phi_2=(\eta_{20}-\eta_{01})^2+4\eta_{11}^2, \tag{14}$$

$$\phi_3=(\eta_{30}-3\eta_{12})^2+(3\eta_{21}-\eta_{03})^2. \tag{15}$$

In order to reduce the computational complexity, only three lower-order moments are used in our landmark detection system.

To complement the information of Hu invariant moments, Haar-like features are also used. The 14 feature prototypes are shown in Fig.1, and each feature prototype is scaled in x and y directions by an integer factor. Though the set of Haar-like features is large enough to present all details of landmarks, it is too expensive to evaluate all the features. Therefore this paper uses a selected scheme described by Zhou *et al.*(2005).

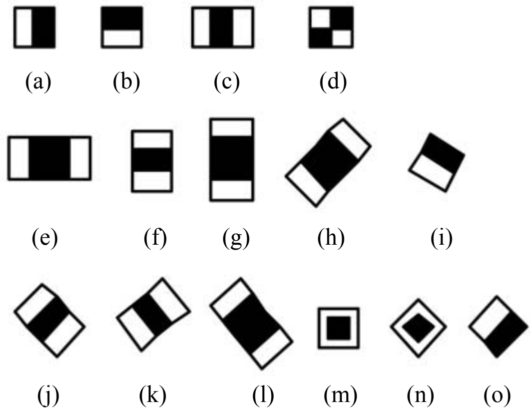

Fig.1 Feature prototypes (a)~(d) were introduced by Voila (2001) and (e)~(o) were introduced by Lienhart (2002). Black areas and white areas have negative and positive weights, respectively

Weak learner

How to construct the weak learners is a fundamental problem in designing a classification algorithm based on boosting method. Following (Voila and Jones, 2001; Lienhart and Maydt, 2002), the weak function set B is constructed by 1D decision stumpsas. In Eq.(1), the weak learner $b(x)$ is associated with a feature $f'(x)$, which is selected from Haar-like features and Hu invariant moments, a decision threshold θ, and a parity direction indicator p. $b(x)$ takes a binary value of either +1 or −1:

$$b(x)=\begin{cases}+1, & \text{if } pf'(x)\geq p\theta,\\ -1, & \text{otherwise.}\end{cases} \tag{16}$$

Prior knowledge model

The prior knowledge model ($\sigma'(\cdot)$=Pr′(Y=1|X=x)) describes the probability of any example belonging to any class according to the prior knowledge. In terms of our observation and simulation experiments, we find the detecting image containing the three Hu invariant moments of the landmark image highly indicates that it is the landmark image. Thereby, in this paper, the prior knowledge can be described as:

“If the image contains the Hu invariant moments of the landmark image, then this image is probably the landmark image, and the probability is p'.”

The Prior knowledge model can be built as

$$\sigma'(\cdot) = \Pr'(Y=1 \mid X=x) = \begin{cases} p'/3, & \text{if } x \text{ is a Hu invariant moment,} \\ (1-p')/(k-3), & \text{otherwise,} \end{cases} \quad (17)$$

where k is the number of used features, p' can be roughly estimated by experience. If the Hu invariant moments of the landmark image are not present, $\sigma'(\cdot)=1/k$.

EXPERIMENTS

Experimental setup

To evaluate the system performance, a mini unmanned helicopter is used as the experimental platform. A control box, which includes a computer system and many sensors, is slung beneath the helicopter fuselage. The box is fixed near the center of mass to offset the helicopter center of inertia as little as possible. The helicopter has a payload of about 2 kg and the rotor diameter is 1.2 m. The maximal fly speed is 80 km/h. On a 1.7 G PC104 embedded computer, the landmark images are detected from Logitech Quick Cam Pro 4000 CCD camera with 1.3 million pixels.

Fig.2 Fully equipped mini unmanned helicopter in flight for landmark detection

Fig.3 shows the experimental landmarks. Fig.3a is the landmark of the helicopter landing pad. Sharp *et al.*(2001) detects this landmark to navigate a mini unmanned helicopter landing. Fig.3b is usually used in the International Aerial Robot Competition. Fig.3c is used in the second China Aerial Robot Competition. Each landmark is printed on a 1 m×1 m plastic board.

As shown in this paper, the approach introduced in the previous section adapts to various boosting algorithms. Different boosting algorithms are applied to different applications, the comparison of them has

(a) (b) (c)

Fig.3 The landmarks of mini unmanned helicopter landmark detection

been described by (Friedman *et al.*, 2000; Bühlmann and Yu, 2003). To simplify programming, the boosting algorithm we used is Adaboost algorithm, which is available in many software development kits. For each landmark, we trained boosting with or without prior knowledge on the 30, 50, 100, 150, …, 1000 examples. The parameters of prior knowledge model p' are 0.9, 0.7, and 0.7 for landmarks (a), (b), and (c), respectively.

Results

Figs.4a~4c show the results of the experiments. The geometrical characteristic of landmark (a) (Fig.3a) is simple. If the experimental condition is well, it is easy to be detected only by using Hu invariant moments (Saripalli *et al.*, 2002). Fig.4a shows that the prior knowledge model is similar to the ultimate detection model, and the machine learning algorithm improves the robustness of the detector described by Saripalli *et al.*(2002).

Since the background of landmark (b) (Fig.3b) is white, parts of the landmarks are easily occluded by noises. Haar-like features are redundant, they can supply enough information to the detector when the three Hu invariant moments cannot be detected due to the noises. Fig.4b shows that the curve driven by data and knowledge tends to be smooth after using 300 training examples.

Detection of landmark (c) (Fig.3c) is difficult due to its complicated geometrical characteristic. Fig.4c shows when the error rate drops to 35%, the approach described in this paper uses 350 training examples, while the "data driven only" approach uses more than 700 training examples.

According to the results shown in Fig.4, for small training examples, the performance of the system using prior knowledge is dramatically better than that driven by data only. When the number of training instances increases, their performance tends to be similar.

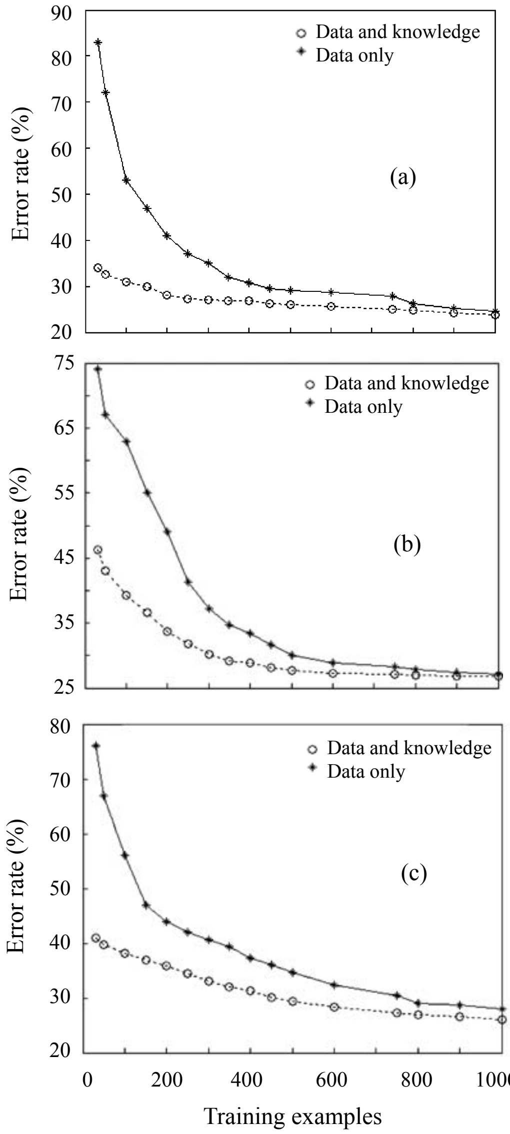

Fig.4 Comparison of test error rate using prior knowledge and data separately or together on landmark detection. (a), (b), (c) are respectively the detection results of landmarks (a), (b), (c) in Fig.3

CONCLUSION

This paper presents a simple approach that incorporates prior knowledge into boosting algorithm for the mini unmanned helicopter landmark detection. The stage forward stagewise additive model of boosting was analyzed. Based on the analysis, the loss function was modified to incorporate the prior knowledge by using Kullback-Leibler divergence. Then, the approach applied in the landmark image detection is described. Meanwhile, the multi-features are boosted to solve a series of problems in the landmark detection, such as rotation, noises affected, etc. The approach effectiveness is verified by experiments.

References

Amidi, O., Kanade, T., Miller, J.R., 1998. Vision-based Autonomous Helicopter Research at Carnegie Mellon Robotics Institute 1991-1997. American Helicopter Society Int. Conf., p.375-386.

Bühlmann, P., Yu, B., 2003. Boosting with the L_2-loss: regression and classification. *J. Am. Stat. Assoc.*, **98**:324-338. [doi:10.1198/016214503000125]

Csiszar, I., 1975. I-divergence geometry of probability distributions and minimization problems. *Ann. Probab.*, **3**:146-158. [doi:10.1214/aop/1176996454]

Freund, Y., Schapire, R., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Computer Syst. Sci.*, **55**(1):119-139. [doi:10.1006/jcss.1997.1504]

Friedman, J.H., Hastie, T., Tibshirani, R., 2000. Additive logistic regression: a statistical view of boosting. *Ann. Stat.*, **28**:337-374. [doi:10.1214/aos/1016218223]

Harris, D., Corinna, C., 1996. Boosting decision trees. *Adv. Neural Inf. Processing Syst.*, **8**:479-485.

Lienhart, R., Maydt, J., 2002. An Extended Set of Haar-like Features for Rapid Object Detection. Proc. IEEE Int. Conf. on Image Processing, **1**:900-903. [doi:10.1109/ICIP.2002.1038171]

Saripalli, S., Montgomery, J.F., Sukhatme, G.S., 2002. Vision-based Autonomous Landing of an Unmanned Aerial Vehicle. IEEE Int. Conf. on Robotics and Automation, **3**:2799-2804.

Schapire, R.E., 1990. The strength of weak learnability. *Matching Learning*, **5**(2):197-227. [doi:10.1023/A:1022648800760]

Schapire, F., Freund, Y., Schapire, R., 1996. Experiments with a New Boosting Algorithm. Proc. 13th Int. Conf., p.148-156.

Schapire, R.E., Rochery, M., Rahim, M., Gupta, N., 2005. Boosting with prior knowledge for call classification. *IEEE Trans. on Speech and Audio Processing*, **13**:174-181. [doi:10.1109/TSA.2004.840937]

Schwenk, H., Bengio, Y., 2000. Boosting neural networks. *Neural Computation*, **12**(8):1869-1887. [doi:10.1162/089976600300015178]

Sharp, C.S., Shakernia, O., Sastry, S.S., 2001. Vision System for Landing an Unmanned Aerial Vehicle, Robotics and Automation. IEEE Int. Conf. on Image Processing, **2**:1720-1727.

Viola, P., Jones, M., 2001. Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition*, **1**:511-518.

Zhou, S.K., Georgescu, B., Zhou, X.S., Comaniciu, D., 2005. Image Based Regression Using Boosting Method. Proc. Tenth IEEE Int. Conf. on Computer Vision, **1**:541-548. [doi:10.1109/ICCV.2005.117]