Human Tracking based on Soft Decision Feature and Online Real Boosting

Takayoshi Yamashita[1,2] Hironobu Fujiyoshi[2] Shihong Lao[1] Masato Kawade[1]
[1] OMRON Corporation,
9-1, Kizugawadai, Kizugawa-City KYOTO, JAPAN, 6190283
[2] CHUBU University,
1200, Matsumoto-cho, Kasugai-City, Aichi, JAPAN, 4878501
takayosi@ari.ncl.omron.co.jp

Abstract

Online Boosting is an effective incremental learning method which can update weak classifiers efficiently according to the object being trackedt. It is a promising technique for online object tracking to adapt tothe appearance variations of objects during tracking process. However, proposed online-boosting based tracking methods update and select weak classifiers from fixed the offline learned weak classifiers, which might not be an optimal selection for object appearance variations. In this paper, we propose a new feature adjusting strategy for online boosting called Soft Decision Feature. We combine it with online real AdaBoost to achieve better tracking performance in scenes with human pose and posture variations. Experiment result demonstrates that it can successfully deal with the human posture variation scenes that conventional online boosting tracking methods fails to deal with.

1. Introduction

Boosting such as AdaBoost and real AdaBoost[1] is an excellent learning method that trains a strong classifier out of many weak classifiers. Although boosting is primarily used for offline training, Oza has proposed online boosting that updates the weight of weak classifiers to adjust the target object shape [2]. This new type of online training method only updates the weight of all weak classifiers. Grabner's method not only updates the weight but also selects weak classifiers from a set of classifiers that are trained in an offline process, and they are applied to object tracking [3]. Shu and Yamashita also proposed a new online training method based on online boosting and Real AdaBoost[4]. Their method trains weak classifiers that have a hypothesis and probability density functions of positive and negative classes in the offline process. These probability functions are updated by new samples. The best weak classifiers are then selected using real AdaBoost in the online process. As a result, this online real boosting that uses real AdaBoost with online boosting can reduce the number of weak classifiers and can improve the tracking performance compared with the conventional online boosting methods.

Online real boosting selects the best of a weak classifier set dealing with tracking object shape variations such as cloth color, body shape, and poses. However, the system has to select from only a few weak classifiers that are trained offline. Thus, the new weak classifiers that are selected online are not always ideal for tracking object. In particular, if training samples selected offline do not include various pose samples, online real boosting cannot track the pose changes because it has no weak classifiers to deal with the variation in poses. Thus, conventional online boosting has to select from only a few weak classifier candidates to adjust to various object variations. We describe these features as “Hard Decision Feature”.

In this paper, we describe new adjustable feature for online boosting. This new feature, called “Soft Decision Features,” is capable of adjusting to object variation. Hard Decision feature have feature information that is a configuration and a hypothesis, but Soft Decision Feature has a lot of feature information that includes some configuration and a hypothesis. The hypothesis is shared by all feature configurations. These configurations are trained offline and have a shared hypothesis. The best feature configuration is selected to adjust to tracking the object shape online. Our new adjustable feature and online real boosting have better tracking performance for poses and posture changes.

2. Online Boosting

2.1 Online Real Boosting

The online boosting proposed by Oza updates the weight of each weak classifier to adjust to the tracking object. This method uses all weak classifiers. Grabner

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

has applied online boosting to tracking, and their method not only updates the weight but also selects weak classifiers from a weak classifier set. The online real boosting proposed by Shu and Yamashita improves the tracking performance using real AdaBoost and probability density functions. We describe online real boosting next.

Strong classifier $H_{strong}=\{h_1,...,h_M\}$ is trained offline using real AdaBoost. The weak classifiers have a hypothesis and probability density functions of positive class $W_m^+(x)$ and negative classes $W_m^-(x)$. They are divided by a selector online. Each selector has different weak classifiers. The new training samples $\{(x_1,y_1),\ldots,(x_n,y_n)\}$ in time t are used for updating the probability density as in eq.(1)

$$W_j^y(x) = W_j^y(x) + \lambda_n \quad . \qquad (1)$$

Note that $W(x)$ consists of a histogram, where y is the class label and λ_n is the weight of the training sample. Similarity variance Z is calculated using the probability density of all weak classifiers as

$$Z = \sum_{j=0}^{N} \sqrt{W_j^+ W_j^-} \quad . \qquad (2)$$

Note that N is the number of histogram bins.
The weak classifier with minimum Z is selected as the representative of the selector. The strong classifier in time t is trained by the best weak classifiers in each selector as

$$H_t(x) = \sum_{m=0}^{M} h_m(x) \quad . \qquad (3)$$

2.2. Problems of previous work

Weak classifiers are selected from a few weak classifier candidates that are trained offline. Thus, the new weak classifiers that are selected online are not always best for tracking object. If the training samples obtained offline do not include various samples of poses, online real boosting cannot track the changes in poses because it has no weak classifiers to deal with the pose variation. Thus, conventional online boosting has to select from only a few weak classifier candidates to adjust to the variation in different objects. We describe this feature as “Hard Decision Feature”.

We describe new adjustable feature for online boosting. These new feature adjusts to object variation. We call them “Soft Decision Feature”.

3. Proposed method

Soft Decision Features are trained offline and are adjusted to track object online, as shown in Fig. 1. First, we describe Soft Decision Feature. We then describe offline and online training.

3.1. Soft Decision Feature

Fig. 2 shows the concept of Soft Decision Feature. Hard Decision Feature consists of a hypothesis and a

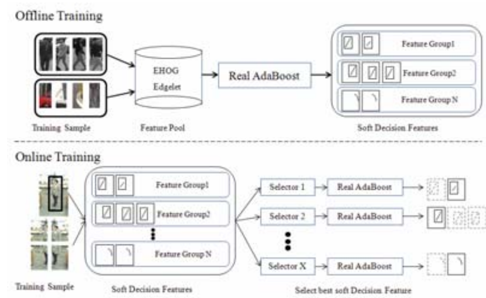

Fig. 1. Framework of our method

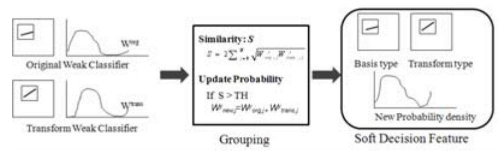

Fig. 2. Concept of Soft Decision Feature

configuration that previously defines the position, size and shape. Meanwhile, Soft Decision Feature consists of a hypothesis and some configurations. The best configurations are selected to adjust to the shape variation when tracking an object during online training.

The basis of Soft Decision feature involves the use of extended histograms of oriented gradients (EHOG) [5] and Edgelet[6]. EHOG is a gradient feature proposed by Hou. It improves HOG and speeds up feature calculation. Although a block and several cells construct the original HOG, the EHOG is constructed only by one cell. Edgelet is employed in human detection that was proposed by Wu. This feature is calculated by edge magnitude and the difference of edge orientations. Although both the EHOG and Edgelet focus on gradient information, the EHOG treats gradient variation of the local area, while Edgelet treats partial edge similarity.

The configuration of Soft Decision Feature is transformed from an original weak classifier as a transformed weak classifier in eq.(4) and eq.(5):

$$F_{trans}^{EHOG} = F_{org}^{EHOG} * T(x, y, w, h, g) \qquad (4)$$

$$F_{trans}^{Edgelet} = F_{org}^{Edgelet}(u_1 * T(x, y, g)),.., u_k * T(x, y, g))) \quad (5)$$

The transform function has the parameters of position, size, and gradient orientation. Note that x and y are the distance from the original position, w and h are the size variation, and g is the orientation variation. U_k is the elements of Edgelet.

3.2. Offline training

Fig. 3 shows the procedure of offline training using Soft Decision Feature. Weak classifier candidates are transformed into a new feature configuration according to the transformed weak

1. Initialize weight of train sample: $\lambda_n = 1/N$
2. For t=1,…,T Number of weak classifier
 2.1. For m=1,…,M Number of candidate
  2.1.1Update probability density
   if $h_m(x) \in bin_j$
   $$W^y_{org,j}=W^y_{org,j}+\lambda_n \quad y=\{0,1\} \tag{6}$$
  2.1.2 Transform feature shape
   $$h_{trans}=h_{org}*\mathrm{T}(x,y,w,h,g) \tag{7}$$
  2.1.3Update probability density
   if $h_{trans}(x) \in bin_j$
   $$W^y_{trans,j}=W^y_{trans,j}+\lambda_n \tag{8}$$
  2.1.4 Similarity
   $$S = 2\sum_{j=0}^{N}\sqrt{W^y_{org,j}W^y_{trans,j}} \tag{9}$$
  2.1.5Merge probability density
   if $S > TH$
   $$W^y_j=W^y_{org,j}+W^y_{trans,j} \tag{10}$$
  Repeat 2.1.2to 2.1.5
  2.1.6 Calculate Z
   $$Z_m = 2\sum_{j=0}^{N}\sqrt{W_j^+W_j^-} \tag{11}$$
 2.2Select best weak classifier
   $$h_t = \arg\min Z_m \tag{12}$$
 2.3Classificaion function
   $$h_t(x) = \frac{1}{2}\ln\left(\frac{W_t^+(j)+\beta}{W_t^-(j)+B}\right) \tag{13}$$
 2.4Error rate
   $$\varepsilon^y_t=\varepsilon^y_t+\lambda_n|h_t(x_n)| \tag{14}$$
   $$\varepsilon_t=\varepsilon^+_t/(\varepsilon^+_t+\varepsilon^-_t) \tag{15}$$
 2.5 Update weight
   $$\text{if } yh_t(x_m) > 0 \quad \lambda_n = \lambda_n / 2(1-\varepsilon_t) \tag{16}$$
   $$\text{else} \quad \lambda_n = \lambda_n / 2\varepsilon_t \tag{17}$$
3. Strong Classifier
   $$H(x) = \sum_{t=0}^{T} h_t(x) \tag{18}$$

Fig.3. Offline training using Soft Decision Feature

classifier and obtained probability density function in 2.1.2 and 2.1.3. Similarity degree S between the original weak classifier and the new transformed weak classifier is calculated in 2.1.4. If S is larger than threshold TH, probability density functions of the original weak classifier and the transformed one are merged. Soft Decision Feature has several configurations based on the repetition of the grouping process in 2.1 of Fig. 3.

Hard Decision Feature updates the probability density function for a weak classifier candidate as shown in 2.1 of Fig. 2 and obtains Z. Soft Decision Feature makes a group that has some configurations with similar probability density functions for an adjustable shape. Thus, online boosting using Soft Decision Feature can select the best configuration for tracking object posture.

0. Soft Decision Feature set $h_{offline\ m}(x)$ m=1,..,M
1. Divide into Selector N
2. Train sample of time t $I(x,y)$, $x=1,2,..L$, $y\{+1,-1\}$
3.Initialize weight of train sample $\lambda_l = 1/L$
4. For n=1,2,..N Number of selector
 4.1For $m=1,2,..,M/N$ Number of Soft Decision Feature
  4.1.1Update probability density
   If $h_{n,m}(x) \in bin_j$
   $$W^y_j=W^y_j+\lambda_l \quad y=\{0,1\} \tag{19}$$
  4.1.2Calculate Z
   $$Z_{n,m} = 2\sum_{j=0}^{N}\sqrt{W_j^+W_j^-} \tag{20}$$
 4.2Select best Soft Decision Feature
   $$h_n = \arg\min Z_{n,m} \tag{21}$$
 4.3Adjust feature shape
  4.3.1Transform feature shape
   $$h_{trans}=h_{org}*T(x,y,w,h,g) \tag{22}$$
  4.3.2Update probability density
   $$W^y_{n,trans,j} = W^y_{n,trans,j} + \lambda l \quad \mathrm{y}=\{0,1\} \tag{23}$$
  4.3.3 Calculate Z
   $$Z_n = \sum_{j=0}^{N}\sqrt{W^+_{trans,j}W^-_{trans,j}} \tag{24}$$
  4.3.4Select best weak classifier
   $$h_n = \arg\min Z_{n,trans} \tag{25}$$
5. Strong classifier
   $$H(x) = \sum_{n=0}^{N} h_n(x) \tag{26}$$

Fig.4. Online training using Soft Decision Feature

3.3. Online training and tracking

Fig. 4 shows the procedure of online bon feature is divided into selectors. The probability density function is updated by the new training samples of time t and is calculated to normalize factor Z in 4.1 of Fig. 4. Soft Decision Feature, which has a minimum Z, is selected as shown in 4.2 of Fig. 4. The probability density function of each transformable weak classifier is then updated in 4.3.2 and is calculated by normalizing factor Z, respectively. The transformed weak classifier, which has a minimum Z, is selected as a representative weak classifier of the selector in 4.3.2. The procedure from 4.1 to 4.3 is iterated for all selectors and obtains a strong classifier in time t.

A Particle Filter is used for tracking in our method. The weight of each particle is calculated from the confidence for a strong classifier as eq. (29). The particle that has the maximum likelihood outputs the tracking object position.

4. Experiment

We evaluated the tracking performance to compare the two types of features. The evaluation sequence was

Table 1. Tracking accuracy comparison

	Proposed method		Hard Decision Feature	
Posture	Accuracy [%]	Variance	Accuracy [%]	Variance
Standing	1.73	1.17	1.84	1.91
Sitting	2.10	4.12	7.34	4.08
Standing	2.70	4.19	Miss tracked	—
Total	4.59	13.99	7.52	39.99

taken over about 800 frames, and the person tracked changed pose and posture as he walked. First, the person walked in front of a bookshelf and shifted to the left. After sitting down, human stood up and walked to the right bookshelf. Center position accuracy was applied to the tracking results of the evaluation metric in eq. (30).

$$S = \mathrm{Dist}(\mathrm{Tracking}, \mathrm{Ground\ Truth}) / W\ [\%] \quad (30)$$

Note that *Dist* is the distance between the tracking results and the ground truth, and *W* is the width of the person in the ground truth.

4.1 Evaluation Results

The tracking accuracy of the online boosting using the Soft Decision and Hard Decision features are illustrated in Table 1. The tracking accuracy of both methods was similar in the first standing scene. In the sitting scene, the accuracy of Soft Decision Feature was higher than that of Hard Decision Feature. Soft Decision Feature was also successful in the second standing scene, where Hard Decision Feature could not track the person in this scene. Fig. 5 and Fig. 6 are the tracking results of the representative frame with pose and posture changes in the evaluation sequence. As shown in Fig. 5 (c), Soft Decision Feature correctly tracked the person, whereas Hard Decision feature miss-tracked the second standing scene as shown in Fig. 6 (c). Furthermore, the total tracking accuracy of Soft Decision Feature was higher than that of Hard Decision Feature. Therefore, Soft Decision Feature was selected as the best weak classifiers to adjust for the object pose and posture variation.

5. Discussion

Fig. 7 shows the tracking results in the sitting scene. Fig. 7 (b) and Fig. 7 (d) are human models that are obtained by superposed weak classifiers. The weak classifiers of Hard Decision Feature were selected intensively from the head area, and the tracking rectangle was off to the left, as shown in Fig. 7 (d). Thus, the probability density was updated incorrectly, and the tracking failed subsequently. Meanwhile, the weak classifiers of Soft Decision Feature were selected evenly from the entire body of the person, such as the head and back, as shown in Fig. 7 (b). The probability density was updated correctly, and it obtained a high tracking accuracy even after a change in the scene.

(a) (b) (c)

Fig. 5. Tracking results based on Soft Decision Feature

(a) (b) (c)

Fig. 6. Tracking results based on Hard Decision Feature

(a) (b) (c) (d)

Fig. 7 Human model by selected weak classifiers in online training (a),(b) Soft Decision Feature, (c)(d) Hard Decision Feature

6. Conclusion

In this paper, we described a new online boosting method using Soft Decision Feature. Soft Decision Feature transforms the configuration to adjust to pose and posture variation in tracking object during online training. The tracking accuracy of our method is better than that of the conventional method when subjected to high variation in people's poses and postures.

References

[1]R.E.Schapire, et al., Improved Boosting Algorithms Using Confidence-rated Predictions, Machine Learning, No. 37, pp. 297-336, 1999.

[2]N.Oza, et al., Online bagging and boosting, In Proc. Artificial Intelligence and Statistics, pp. 105-112, 2001.

[3]H.Grabner, et al., On-line Boosting and Vision, In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, vol. 1, pp. 260-267, 2006.

[4]L.Xu, et al., Online real boosting for object tracking under severe appearance changes and occlusion, In Proc. IEEE Conf. on Acoustics, Speech and Signal Processing, 2007.

[5]C.Hou, et al., Multiview Pedestrian Detection Based on Vector Boosting, In Proc. IEEE Conf. on Asian Conference on Computer Vision, 2007.

[6]B. Wu, et al., Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors, In Proc. IEEE Conf. on International Conference of Computer Vision, vol.1 pp. 90-97, 2005.