SiftKeyPre: A Vehicle Recognition Method Based on SIFT Key-Points Preference in Car-Face Image

Chang-You Zhang[1,2(✉)], Xiao-Ya Wang[1,2], Jun Feng[2], and Yu Cheng[3]

[1] Laboratory of Parallel Software and Computional Science, Institute of Software, Chinese Academy of Sciences, Beijing 100190, China
changyou@iscas.ac.cn

[2] School of Infomatics Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 050043, China

[3] Institute of Applied Mathematics, Hebei Academy of Sciences, Shijiazhuang 050081, China

Abstract. Vehicle recognition from images produced in roads bayonet provides important clues to solve vehicle crime cases. Its accuracy is not enough to meet the requirement in real conditions. We proposed a vehicle recognition method, SiftKeyPre, based on SIFT(Scale-invariant feature transform) key points preference for car-face images. Firstly, SiftKeyPre choices the SIFT key points following the DualMax algorithm to get a DualMax set. Meanwhile, Lowe set is defined as another one following Lowe algorithm. Secondly, we define a DL set under an intersection operation on DualMax set and Lowe set. For positive examples training images, we count the appearance times of each key point of DL set to compute the attention degree of each key point in base image. Finally, matching degree between the base image and a target image is evaluated with the attention degree of each matched points. SiftKeyPre method confirms a testing image based on its matching degree. Experiments results show that, under a given recall constraints, the precision of SiftKeyPre method is better than FLANN and Lowe. SiftKeyPre's computational complexity is closed to that of Lowe. Comparing with other algorithms based on training, SiftKeyPre is of lower training intensity.

Keywords: Car-face image · SIFT key points · Preference · Attention degree · Matching degree · Recognition

1 Introduction

In modern societies, the insecurity and threat events are increasing. Vehicle recognition from high-definition vehicle images produced in roads bayonet is an important source of clues on which public security departments solve cases of vehicle crime relies. For the phenomenon of faking and sheltering vehicle license plate, vehicle recognition system can not recognize car types correctly just based on the license plate. In a medium-sized city, more than 10000 images are captured at each major road bayonet by high-definition cameras per hour. At present, policeman selects out certain type of vehicle, for example black PASSAT, by "eyes of human".

H. Zha et al. (Eds.): CCCV 2015, Part II, CCIS 547, pp. 344–358, 2015.
DOI: 10.1007/978-3-662-48570-5_34

This retrieval process is non-efficiency and tedious. It is urgency for investigation on automatic recognition algorithms to identify suspect vehicle.

The basic problem in computer vision research is object classification and detection. The image object recognition is an important branch with more than fifty years history [1]. Image object recognition algorithms are divided into two basic categories, algorithms based on low-level feature and deep learning model. In algorithms based on visual feature, low-level features are extracted from images, and then these features obtained from a variety of features extraction algorithms are encoded. Finally, the appropriate classifier is designed to get classification results. Deep learning model [2] is another kind of image object recognition algorithm. Its basic idea is to learn hierarchical feature representation in supervised or unsupervised way. The objects are described from bottom to top.

Classifier design is the key step in object recognition based on low-level features. Classical classifier based on visual feature include FLANN, Lowe (classifier in the SIFT algorithm) etc. FLANN algorithm is based on key points in base image and calculates Euclidean distance one by one with key points of each testing image. It selects minimum distance and takes key-points number exceeded the given threshold as matched points. It determines whether the image object is the same according to the number of matched points. Lowe also calculates Euclidean distances between one key point of base image and all key points of testing image. If the distance is smaller than a given value, Lowe algorithm selects this key-point pair as a high quality match [3]. The object similarity in Lowe is still based on the number of matched key points. Training-based classifiers include neural networks, support vector machines, k-nearest neighbor, random forests, and so on. These algorithms need to manually mark a large number of training samples to improve the classification quality.

Big data brings huge challenges to the traditional learning algorithms. Deep learning model has powerful ability to express data naturally, so it will impact on image object detection methods. However, there are some problems like poor interpretation, high model complexity, optimization difficulty, computing intensity etc [1]. Main deep learning models include automatic coder (Auto encoder) [4], restricted Boltzmann machines (Restricted Boltzmann Machine, RBM) [5], deep belief networks (Deep Belief Nets, DBN) [6], Convolutional neural networks (Convolutional Neural Networks, CNN) [7], bio-heuristic model [8], etc. Deep learning models rely on huge amounts of training samples, and it is of high training intensity.

Classification algorithms based on training require a lot of manual marked samples, or huge amounts of training samples. It is hardly to be used in sample-limited scenarios. Missing match rate is high in classic FLANN and Lowe algorithm when the distance is close between a pair of key-points. We proposed a SiftKeyPre method based on SIFT key-points preference in vehicle image. SiftKeyPre makes a compromise by taking advantages of both the classical linear classifier and training classifiers with high computing intensity. It extracts car-face as region of interest. In SiftKeyPre method, the intersection set of Lowe's preferring key points set and DualMax's preferring key points set is used as the preferring key points set. SiftKeyPre set is used to calculate attention degree of base image. Matching degree between the base image and testing images is evaluated. SiftKeyPre is essentially a linear classifier with low-intensity training algorithms.

The structure of the rest of this paper is organized as follows. Section 2 introduces some related works on image recognition. Principle of SiftKeyPre is described in section 3. Section 4 tests the effect of SiftKeyPre through experiment. Section 5 further analyses the results and algorithm parameter to enhance the practicability of SiftKey-Pre. Section 6 summaries our works, and discusses SiftKeyPre's limitations and some further promising researches.

2 Related Works

The variation of angle and distance between the running cars and high-definition camera causes image differences in scale and orientation. Thus, we use SIFT algorithm to extracting features from images to guarantee the invariant of image scale and rotation.

Researchers investigated some auto-algorithms of car detection and vehicle recognition. Wei [9] developed a complexity-aware criterion to balance the separating capacity and retrieval efficiency based on strip feature of car in static images. This algorithm only detected the existence of vehicle, not identified its types. Some approaches classify vehicles into generic classes (vans, SUVs and bus etc.) [10-11]. These methods were not accurate enough for finding crime vehicles. Vehicles need to be classified into specific ‘make and model’ classes (MMR, Make and Model Recognition) [12].

Some scholars put forward some ideas and algorithms of vehicle recognition at early years. Sun Ze-hang [13] proposed an algorithm of vehicle recognition using Haar Wavelet decomposition for features extraction and Support Vector Machines (SVM) for classification. David Santos [14] introduced a vehicles recognition algorithm relying on the analysis of car external features. These features included shape of the car’s rear view and the car back lights. Through comparing with the features in database, system determined whether both images were matched or not. Daniel Marcus Jang [15] demonstrated a recognition application based on the SURF (Speeded Up Robust Features) algorithm. In its vehicle database, the images of each type of vehicle were photographed from 16 views. It cost approximately 16 times computing workload. Wang Quan [16] presented a MDS(multi-dimensional scaling) feature learning framework in which MDS is applied on high-level pairwise image distances to learn fixed-length vector representations of images. Images need to be uniformed caused information loss. Ferencz [17] built a classification cascade for visual recognition from one example and proposed an approach for vehicle recognition. The main contribution of this work is a classification cascade built by arranging information-rich hyper-features extracted from a single vehicle exemplar image. For running vehicles, images of exact front and lateral views are very rare.

So for from 2004, WOB (Word of Bag) are the mainstreem algorithm in image recognition. The main idea of these algorithms is clustering the features by employing K-means clustering algorithm to construct the visual vocabulary. These clustering centers are regarded as visual words. Then, they use the histogram described by appearance frequency of the visual words to represent the content of the image. By regarding the visual words histogram of each image as features vector, the classification model was abtained through SVM training [18-20]. In these algorithms, the differences of the categories are obvious. And the training of SVM was based on enough selected samples of images. In the scenario of just 100 car-face images, the effect of SVM training is limited.

3 SiftKeyPre Recognition Method

The vehicle recognition processes of SiftKeyPre consist of five steps. They are designing data structure of key points, constructing key-point pairs, preferring key-point pairs, calculating attention degree of key points, and calculating matching degrees of target images. The algorithms components are illustrated in Fig. 1.

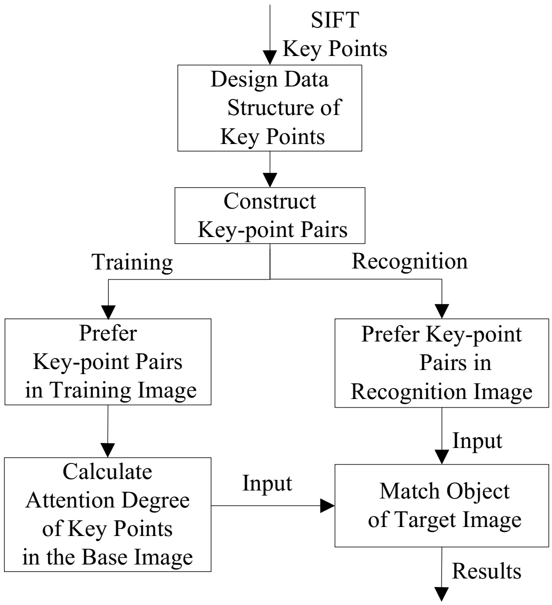


Fig. 1. Framework of SiftKeyPre method.

We call the template image of a given type of vehicle as *base image*, and the image to be matched as *target image*. In SiftKeyPre algorithms, there is just a single base image and multiple target images. For convenience, we assume that there are m SIFT key points in base image, and n SIFT key points in any one target image. Each key point is described as a vector with 2 float numbers and 128 integers. Euclidean distance is used to measure the similarity of 2 key points in the same key-point pair.

3.1 Data Structure of Key Points

Data structure of SIFT key points is consist of some parameters such as octave, scale, σ, x, y, and so on. Some of them are no contribution to SIFTKeyPre algorithm. So, we just reserve the pixel position parameter (x, y), and the key points descriptor (128 integers) to construct the data structure of key point. The key point descriptor *kp* is designed as a sequence.

kp = sequence of { x, y, d_0,d_1, …, d_i, …, d_{127} }

Where,
x, y -- is the pixel coordinate of *kp*;
d_i -- i-th component of *kp* ($0 \leq k \leq 127$).

3.2 Construct Key-Point Pairs

The similarities between key points of base image and that of target image is the foundation of image recognization. Let A stand for the key points set of the base image, and B stand for the key points set of a target image. Cartesian product of A and B builds the key-point pairs set C. As mentioned above, there is *n* key points in set A, and *m* key points in set B, then there are $n \times m$ key-point pairs in set C.

Let $A = \{A_0, A_1, \ldots, A_i, \ldots, A_{n-1}\}$,
$B = \{ B_0, B_1, \ldots, B_j, \ldots, B_{m-1}\}$.

Then,

A_0 and B produces key-point pairs $<A_0, B_0>$, $<A_0, B_1>$, ……,$<A_0, B_{m-1}>$;
A_1 and B produces key-point pairs $<A_1, B_0>$, $<A_1, B_1>$, ……,$<A_1, B_{m-1}>$
……
A_{n-1} and B produces key-point pairs $<A_{n-1}, B_0>$, $<A_{n-1}, B_1>$, ……,$<A_{n-1}, B_{m-1}>$

3.3 Key-Point Pairs Preference

Distance is the classic measurement method for evaluation. To evaluate the matching quality of a pair of points in a single key-point pair, we construct a distance matrix H with size of $n \times m$ based on their distances.

We get distances of two key points $<A_i, B_j>$ as formula (1)

$$dist(A_i,B_j) = \sqrt[2]{\sum_{t=2}^{129} (A_i[t]-B_j[t])^2} \tag{1}$$

Where,
A_i -- the *i*-th key point data in set A, ($0 \leq i < n-1$);
B_j -- the *j*-th key point data in set B, ($0 \leq j < m-1$);
$A_i[t]$ -- the *t*-th element of A_i , ($2 \leq t < 129$);

$B_j[t]$ -- the t-th element of B_j, $(2 \le t < 129)$.

These distances are all filled into a matrix H. Row of matrix H corresponds to a certain key point of the base image, and column of matrix H corresponds to that of one target images. It is said that H_{ij} denotes the distance between i-th key point in the base image and j-th key point in a target image.

According to common sense, when the distance of a key-point pair is larger than a special value, we think the both points are not similar. Their similarity is approximately 0. To reduce the computing intensity, we transform distance in H to similarity following the rules that,

(1) Smaller distance mapping to bigger similarity,
(2) If a distance is bigger than a given threshold D, similarity is 0,
(3) Similarity value range is from 0 to 1.

According to the above rules, matrix H is transformed to matching quality matrix EV according to formula (2).

$$EV_{ij} = \begin{cases} 0, & H_{ij} \ge D \\ \dfrac{D - H_{ij}}{D}, & H_{ij} < D \end{cases} \tag{2}$$

Where,

EV_{ij} – similarities in matching quality matrix;
H_{ij} – distances in H;
D – given threshold by experiments on image samples.

From formula (2), there are some zero elements in EV. These elements are no chance to be choose as matched key-point pairs. It is said that the corresponding key points of target image are out of matching.

After above pre-processing, the vital step in SiftKeyPre is to prefer real matched key-point pairs by Dual-direction evaluation. We call this matching selection as DualMax optimization. DualMax follows these steps.

(1) Let i=0;
(2)Traverse the i-th row of EV, select the maximum element,
$ema_{ij} == \max\{ e \mid e_{ij}, 0<=j<=m \}$;
mark j-th column.
(3) Traverse the j-th column, if ema_{ij} is the maximum value,
$ema_{ij} == \max\{ e \mid e_{ij}, 0<=i<=n \}$;
mark the corresponding key-point pair as a DualMax matching, and put it into set Q.
(4) Change all values of i-th row and values of j-th column to zero.
(5) Else i++; goto (2);
(6) If i >= n, finished.

The elements in set Q are preferred key-point pairs.

3.4 Attention Degree of Key Points in the Base Image

Human vision system often focuses visual attention on some special objects of the scene when processing a relative complex scene. It processes these special objects in priority so as to get main information of the scene in minimum time cost [21]. For different aims, these special objects are different. More deeply, there must be a few attentional points to represent the object.

According to this character of human vision system, we speculate that the SIFT key points extracted from car-face are of different attention degrees. To investigate into this, we give each SIFT key point an attention degree by means of a statistic on being preferred times based on a set of positive example images.

Let set L denote key-points set preferred by Lowe matching algorithm. As mentioned above, Q denotes key-points set preferred by DualMax algorithm. To further enhance the preference quality, we build set LQ as an intersection set of L and Q. All the key points in set LQ is deemed as being preferred once. We repeat this operation on a positive sample image set to get the preferred-times of each key-point.

We assume that there are S samples in a given training set. As mentioned above, A is the key point set of base image, and there are n key points in set A. Let AN denote the times of key-point preferred. The attention degrees are calculated in accordance with the following steps.

(1) $i = 0; i < S$;
(2) $LQ_i = L_i \cap Q$;
(3) for each element $kp \in LQ_i$, if kp == A[i], AN[i]++;
(4) i++, goto (1)
(5) output AN, finish.

Finally, we normalize AN according to the formula (3).

$$AD[i] = \frac{AN[i]}{\sum_0^{n-1} AN[i]} \tag{3}$$

Where, AD is a vector of the attention degrees of key points in base image.

3.5 Matching Degree of Target Image

Matching degree is a comprehensive evaluation which compounds matched key points number and their attention degrees. If the matching degree of target is bigger than the given threshold, this target image is matched with the base image.

Let vector AQ denotes the sequence of flag for key points in the base image with n components. AQ is initiated with zero. Vector AV denotes the sequence of key points in base image, and BV denotes the sequence of key points in target image. As mentioned in section 3.4, set Q denotes the preferred key points in target image. We get the matching degree in the following steps.

(1) i = 0;
(2)if key-point pairs <AV[i],BV[j]>∈Q, AQ[i] = 1;
(3) i++;
(4) if i < n, goto (2);
(5) matching degree $v = AQ \bullet AD$;
(6) if v >=V, target image is matched, finish.

Matching threshold V depends on the need of specific application.

4 Experiments

4.1 Experimental Setup

Platform: CPU-Phenom II 960T 3.0GHz* quad-core; RAM-DDR3 3.25G; Ubantu 12.04 OS; openCV library.

Data: The testing images are HD images produced at real road intersections of a city in China. There are total 1000 images, where 100 positive samples (BLACK PASSAT). Typical original images are illustrated in Fig. 2. These images are created in various angles and different distance.

* license numbers are blurred for privacy protection

Fig. 2. Examples of original images.

4.2 Experimental Results

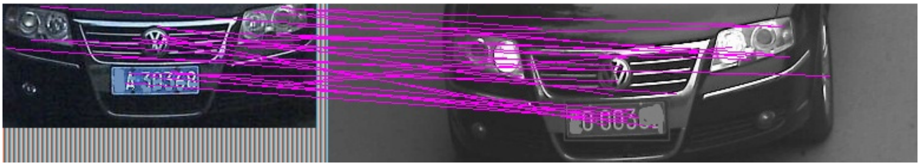
To investigate the precision and performance of our SiftKeyPre algorithm, we compare both indices among the three typical algorithms (FLANN, Lowe and SiftKeyPre) at a given recall.

In this experiment, ROI is car face extracted from original images. SiftKeyPre selects key points of high quality to determine whether the car in a target image of the same type as that in the base image or not. One of the target images' matched key-point pairs are illustrated in Fig. 3.

In Fig. 3, the base image is on the left, and the target image is on the right. Each matched key-point pair is illustrated with a line.

(1) Precision

Effectiveness of SiftKeyPre algorithm is evaluated with two indicators: precision and recall. The precision rate has negative relation with recall. It is said that the improvement



(base image) (target image)

Fig. 3. The matched key-point pairs in SiftKeyPre algorithm.

of precision followed with a drop of recall. We compared the three algorithms (FLANN, Lowe and SiftKeyPre) in their precision and recall. The results are shown in figure 4.

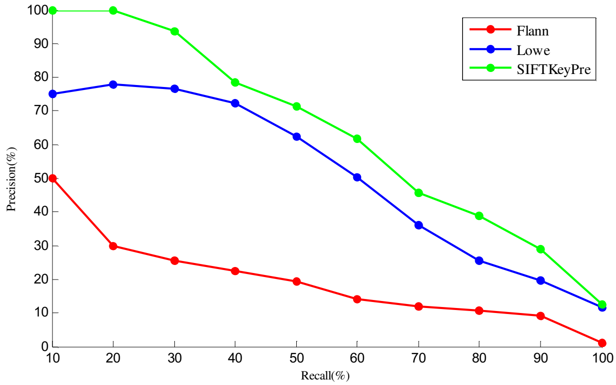


Fig. 4. Comparison between SiftKeyPre/FLANN/Low algorithms.

In Fig. 4, the abscissa denotes recall indicator, and ordinate denotes precision indicator. The experiments test a range of recall from 10% to 100% and the corresponding precision. As shown in Fig. 4, the precision of SiftKeyPre is significantly higher than that of FLANN and Lowe at a given recall. For instance, at the point of recall = 90%, the precision of SiftKeyPre is 27.95%, that of Lowe is 19.65%, and that of FLANN is 9.29%.

Precision are different between SiftKeyPre and the other two algorithms at a various recall from 10% to 100%. These differences are listed in Table 1.

Table 1. Precision differences with FLANN and Lowe algorithms.

differences with	Max	Min	Average
Lowe	+25.00%	+1.04%	+12.46%
FLANN	+35.86%	+1.29%	+16.69%

As shown in Table 1, compared with Lowe algorithm, SIFTKeyPre achieved a maximum +25% improvement in retrieval accuracy. Meanwhile, compared with FLANN algorithm, a maximum 35.86% improvement was obtained. Obviously, Sift-KeyPre performs better than the other two algorithms.

(2) Performance
SiftKeyPre algorithm is consist of two key processes (training and recognition). In training process, we get algorithm parameters such as D, attention degree et al. Training process needs only once in advance. Training with a sample image costs 0.177961s in average. For a given practical application, users seemingly unconcerned about the time cost on training.

Users more concerned about the response efficiency of recognition process. We investigated into the response time for a single target image. The results comparing with FLANN and Lowe are listed in Table 2.

Table 2. Comparison of response time with FLANN and Lowe.

Algorithm	FLANN	Lowe	SIFTKeyPre
average response time (s)	0.122930	0.081694	0.083226

From table 2, SiftKeyPre saves 32.30% than FLANN in response time. SiftKeyPre costs a little longer time than Lowe. Even so, it is well worth to exchange a performance loss of 1.88% for a precision improvement of 35.86%.

5 Analysis and Discussions

In this section, we analyze parameters of SiftKeyPre and discuss the training intensity. To be sure that the parameters of FLANN and Lowe are adjusted carefully to achieve their best precision on testing images.

5.1 DualMax Threshold

In formula (2), D is a key parameter in DualMax preferring process. In fact, D is a critical value to determine whether a distance of key-point pair maps to zero or not. A bigger D means higher quality of key-point pairs. And there are much more 0 in matrix EV. There will be less key-point pairs in DualMax set Q. Meanwhile, this will loss more key point information which contribute to the final recognition.

To balance this compromise, we develop 2 principles. (1) Gold section number is graceful to be used as the dividing line between the zero similarity and non-zero similarities; (2) For the same target image, the number of key-point pairs in Q and LQ should be roughly equal. Accordance with both principles, we determine D in the following steps. As mentioned above, matrix H has n rows and m columns. And the gold section number is 0.618.

(1) Let i =0;
(2) For i-th row, if $H_{ij} = \min\{ H_{ij} \mid H_{ij}<H_{i*}, 0<j<m \}$, $V[i] = H_{ij}$; i++;
(3) if i<n, goto (1);
(4) $R = round\ (n \times (1 - 0.618))$
(5) choose the R-th bigger number in V, let D_k=V[R], (0<=k<K)
(6) For each image in training set of K images, D is valued as average of Dks.

$$D = (\sum_{k=1}^{K} D_k) / K \tag{4}$$

5.2 Training Intensity

SiftKeyPre is a linear classifier with low-intensity training. This training is the important reason for the improvement of precision. We define the training intensity as the minimum number of training samples when the vital parameters of SiftKeyPre are convergent. To investigate into the convergence, we do three experiments from various views. (1) Overview all attention degrees of key points; (2) the transferring curves of typical key points with significant value changing; (3) the impact on recognition precision.

The aim of training is to get the attention degree of each key point in the base image. We do experiments with samples of 10, 20, 30, ..., 90 and 100 and draw the corresponding attention degree values together in the same coordinate system. The inflection point of these curves are the alternative training intensity. The attention degree changing curves are illustrated in Figure 5.

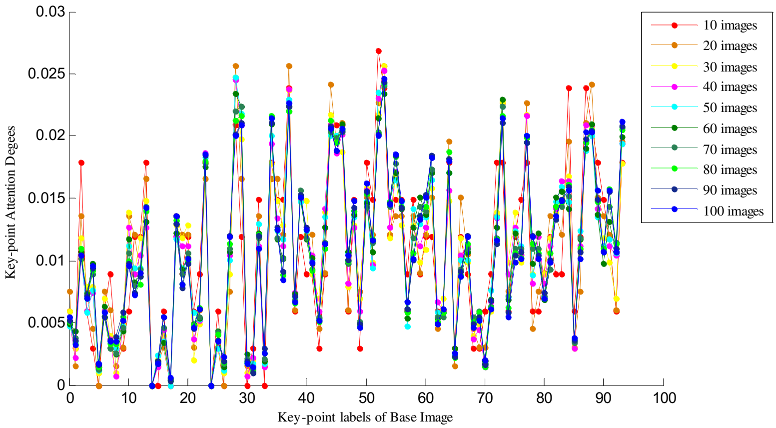


Fig. 5. Attention degree changing curves of all key points.

In figure 5, abscissa is the label of key points in base image, and the ordinate is the attention degree of these key points. As the figure legend, the training times from10 to 100 mapping to the colors from red to blue. The majority key points attention degree are converged to a stable value illustrated in a “cooler” color.

Investigating into figure 5, we find some key points (such as id=47, 52, 65 and 91) whose attention degree value fluctuates more dramatically. To show the trend more clearly, we select these 4 key points and draw the changing path in an unfolded view, as shown in figure 6.

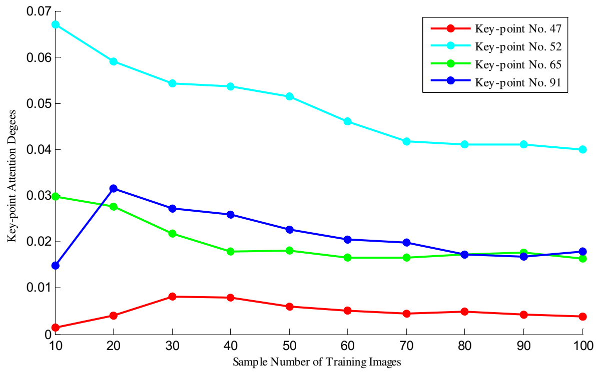


Fig. 6. Changing path of selected attention degrees in unfolded view.

In figure 6, the changing trends show that the attention degree value of key points will be stable under a certain number of training samples. This number range from 50 to 70. Then, a new question is coming. Is there any significant influence on the final precision under the training intensity of 50 and 70?

We developed another experiment on the training intensity of 0, 50, and 100. The changing trend of precision with incremental recall is illustrated in Figure 7.

From figure 7, we find that the precision under no any training is much lower than that under 50 samples’ training. It is said that training process improved the precision of SiftKeyPre. Meanwhile, when the training intensity enhanced from 50 or 100 samples, the both precisions become no obvious difference. It is said that 50 is a critical point of training intensity from the view of precision effect. The attention degrees of key points reach to convergent values.

5.3 Attention Degree

To view the attention degree of key points more clearly, we draw these points on the base image with various radius and colors according to its pixel position of (x, y) and attention degree.. The bigger radius denotes bigger attention degree. Their colors range from blue to red, mapping from the smallest to biggest attention degrees. These 2 pictures in figure 8 illustrate the changing trace of attention degrees under training intensity of 0, 50.

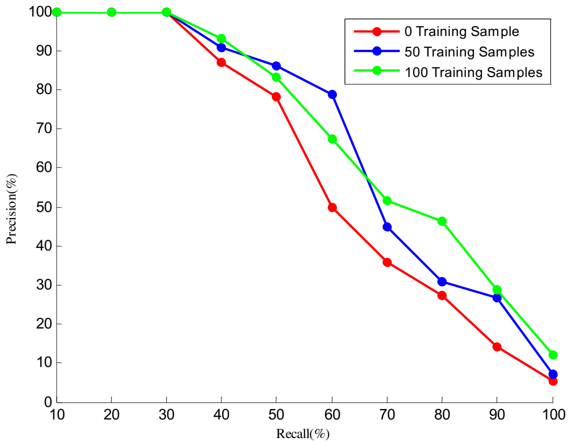
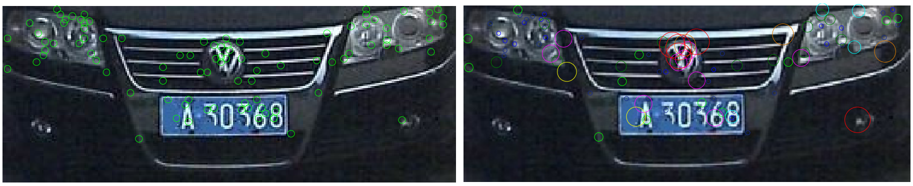


Fig. 7. Trend of precisions with incremental recall



(a)under no training (b)under training intensity of 50

Fig. 8. View of attention degrees of key points.

In the intuition of human perception system, part of key points is of significant contribution in the recognition decision. These key points are more important than others. From figure 8, we found that the most important point concentrate on the region of car logo, car light, and some distinct texture. In figure 8(a), all key points are without training, so the importance are almost in average. In figure 8(b), key points are trained under 50 positive samples. The most important points become significant in size. Training more than 50 samples do not contribute significantly to the size of key points. That proves again that the attention degrees convergence to a stable value.

Some key points with significant attention degree locate on license plate. In fact, this is a wrong matching because license plate is not a inherent part of a car. These points should be removed from the key points set.

6 Conclusion and Future Works

With the analysis on the property of vehicle images on road, we proposed a vehicle recognition algorithm SiftKeyPre based on low-level feature extraction in SIFT algorithm. SiftKeyPre consists of five steps: Design data structure of SIFT key points, construct key-point pairs, prefer key-point pairs, calculate attention degree of key points in base image, and match object of target image. Under the given recall rate, SiftKeyPre achieved obvious improvement of precision comparing with both FLANN and Lowe. As for time-consumption, SiftKeyPre algorithm cost less computing time than that of FLANN in 32.30% and almost equal to that of Lowe.

There are spaces to improve this method. For example, finer pre-processing is helpful to higher precision. Combined SiftKeyPre with support vector machines (SVM) based on WOB, neural networks, and deep learning algorithms will be a promising field in vehicle recognition system. With the huge amount of images streaming into the system, high performance algorithms are the future direction in image recognition and retrieval systems.

Acknowledgement. This paper is supported by the Natural Science Foundation of China (61379048), Hebei Province Natural Science Foundation (F2013210109). Special Project of National CAS Union – The High Performance Cloud Service Platform for Enterprise Creative Computing, Special Project of National CAS Union – Car Image Retrieval on Many core platforms, Special Project of National CAS Union (Hebei, No.14010015)–The Performance Optimization of Key Algorithms in Senile Dementia Data Processing. The authors also would like to express appreciation to the anonymous reviewers for their helpful comments on improving the paper.

References

1. Huang, K.-Q., Ren, W.-Q., Tan, T.-N.: A Review on Image Object Classification and Detection. Chinese Journal of Computers **37**(6), 1225–1240 (2014)
2. Liu, J.-W., Liu, Y., Luo, X.-L.: Research and development on deep learning. Application Research of Computers **31**(7), 1921–1942 (2014)
3. Lowe, G.D.: Distinctive Image Features from Scale-Invariant Key points. International Journal of Computer Vision **60**(91–110), 1–28 (2004)
4. Bourlard, H., Kamp, Y.: Auto-association by multilayer perceptions and singular value decomposition. Biological, Cybernetics **59**, 291–294 (1988)
5. Smolensky, P.: Information processing in dynamical systems: foundations of harmony theory. In: Processing of the Processing of the Parallel Distributed: Explorations in the Microstructure of Cognition, Foundations, vol. 1, Chapter 6. MIT Press (1986)
6. Hinton, G.E., Osindero, S., Teh, Y.: A fast learning algorithm for deep belief nets. Neural Computation **18**(7), 1527–1554 (2006)
7. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)

8. Huang, Y.-z., Huang, K.-q., Tao, D.-c., et al.: Enhanced biologically inspired model for object recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics **41**(6), 1668–1680 (2011)
9. Zheng, W., Liang, L.: Fast Car Detection Using Image Strip Features, pp. 2703–2710. IEEE (2009)
10. Chen, Z., Pears, N.E., Freeman, M., Austin. J.: Road vehicle classification using support vector machines. In: IEEE Int. Conf. Intelligent Computing and Intelligent Systems, pp. 214–218 (2009)
11. Hua, L., Xu, W., Wang, T., Ma, R., Xu, B.: Vehicle Recognition Using Improved SIFT and Multi-View Model. Journal of Xi'an Jiaotong University **47**(4), 92–99 (2013)
12. Pearce, G., Pears, N.: Automatic make and model recognition from frontal images of cars. In: 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance, pp. 373–378 (2011)
13. Sun, Z., Miller, R., Bebis, G., et-al.: A real-time precrash vehicle detection system. In: Sixth IEEE Workshop on Applications of Computer Vision, pp. 171–182, December 3–4, Orlando, Florida (2002)
14. Santos, D., Correia, P.L.: Car recognition based on back lights and rear view features. In: 2009 10th Workshop on Image Analysis for Multimedia Interactive Services, pp. 137–140, May 6–8, London, United Kingdom (2009)
15. Jang, D.M., Turk, M.: Car-rec: a real time car recognition system. In: 2011 IEEE Workshop on Applications of Computer Vision (WACV), pp. 599–605, January 5–7, Kona, HI, USA (2011)
16. Quan, W., Boyer, K.L.: Feature learning by multidimensional scaling and its applications in object recognition. In: 2013 XXVI Conference on Graphics, Patterns and Images, pp. 8–15, August 5–8, Arequipa, Peru (2013)
17. Ferencz, A., Learned-Miller, E.G., Malik, J.: Building a classification cascade for visual identification from one example. In: ICCV 2005, Wash., DC, pp. 286–293 (2005)
18. Gao, H., Dou, L., Chen, W., Sun, J.: Image classification with bag-of-words model based on improved SIFT algorithm. In: 2013 9th Asian Control Conference. ASCC 2013, pp. 1–6, June 23–26, Istanbul, Turkey (2013)
19. Bai, S.: Sparse code LBP and SIFT features together for scene categorization. In: Audio, Language and Image Processing (ICALIP), pp. 200–205 (2014)
20. Csurka, G., Dance, C.R., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: Proceedings of the 8th European Conference on Computer Vision (ECCV 2004), Prague, Czech, pp. 1–22 (2004)
21. Chen, Z., Zou, B.: Research on Region of Interest Extraction. Central South University, May 2012