A Biologically-inspired Vision Based Approach for Detecting Multiple Moving Objects in Complex Outdoor Scenes

Zhengzheng Tu [1,3],Aihua Zheng [1], Erfu Yang [2], Bin Luo [1,3],Amir Hussain [2]

[1]*School of Computer Science and Technology, Anhui University, Hefei 230601, China*

[2]*Division of Computing Science and Mathematics, University of Stirling, Stirling FK9 4LA, UK*

[3]*Key Lab of Industrial Image Processing & Analysis of Anhui Province, Hefei 230039, China*

Zhengzheng Tu,

e-mail: tuzhengzheng@gmail.com

Abstract:

In the human brain, independent components of optical flows from the medial superior temporal (MST) area are speculated for motion cognition. Inspired by this hypothesis, a novel approach combining independent component analysis (ICA) with principal component analysis (PCA) is proposed in this paper for multiple moving objects detection in complex scenes—a major real-time challenge as bad weather or dynamic background can seriously influence the results of motion detection. In the proposed approach, by taking advantage of ICA's capability of separating the statistically-independent features from signals, the ICA algorithm is initially employed to analyze the optical flows of consecutive visual image frames. As a result, the optical flows of background and foreground can be approximately separated. Since there are still many disturbances in the foreground optical flows in the complex scene, PCA is then applied to the optical flows of foreground components so that major optical flows corresponding to multiple-moving objects can be enhanced effectively and the motions resulted from the changing background and small disturbances are relatively suppressed at the same time. Comparative experimental results with existing popular motion detection methods for challenging imaging sequences demonstrate that our proposed biologically inspired vision-based approach can extract multiple-moving objects effectively in a complex scene.

Keywords: Motion Cognition, Optical Flow, Independent Component Analysis, Principal Component Analysis, Moving Objects Detection

1 Introduction

Detecting or segmenting moving objects in complex scenes is a challenging but essential task in many industrial applications such as cognitive vision-based robot tasks. The main difficulties come from two aspects, namely appearance complexity and motion complexity. The reality that objects may have very similar features like textures and colors to the background will lead to appearance complexity, thus making it challenging to distinguish the object from its background using appearance features. Motion complexity mostly originates from dynamic background or bad weather, since bad weather or illumination makes moving objects unclear whereas dynamic background brings redundant motion to the foreground moving objects, which makes it difficult for the robot's vision system to identify moving objects in outdoor environments.

1.1 State-of-the-Art in Motion Detection

In the past years, three kinds of major methods were proposed to detect motion or moving objects. They are temporal difference, background modeling and subtraction, and optical flow methods. The main idea of temporal difference methods is to simply threshold the difference of consecutive frames[1]. The selection of the threshold highly depends on the noise or object's movement intensity in the frame sequences. Temporal difference methods work much faster than other motion detection methods, but they have a big disadvantage that some holes always exist inside the moving objects, and motion detection is very sensitive to noise and dynamic background.

The background modeling and subtraction approach is the most popular method for detecting motions [2,3].Background modeling and subtraction algorithms aim to distinguish a moving foreground from a static, or slow changing background. To be specific, background subtraction aims to compare a static background image with the current frame, pixel by pixel, therefore building the background model of a video clip is the most important step in these types of methods. After building a probability density function of the intensity at each individual pixel for a scene, the next step is to compare this background model with the current frame for the sake of finding a significant difference between them, and detecting these as motion areas. Along with the recent emergence of many models and segmentation strategies, background subtraction approaches can be classified as predictive or non-predictive methods. In brief, non-predictive methods build an intensity's probability density function for the pixels in a frame, for example, the classic single Gaussian [4] algorithm is adopted to indicate the statistical distribution of a pixel in a static environment. If the intensity value of a pixel is quite different from the corresponding value in background model, it will be regarded as belonging to the foreground. But predictive methods[5] develop a dynamical model for the scene and predict the intensity of the current pixel from past observations.

The third type of motion detection methods, the well-known optical flow can describe the apparent motion of objects in a visual scene by movement of brightness patterns from two consecutive frames. Consequently, optical flow can provide important information about the spatial location of the moving objects and their speeds. The motions are acquired by segmenting the optical flow field into regions corresponding to different objects. Meanwhile, as optical flow is independent of some image features like colors or textures, it will not be influenced even if the moving objects are similar to the background. The two fundamental works of optical flow were conducted by Horn & Schunck (H-S)[6] and Lucas & Kanade (L-K)[7]. The optical flow based approaches and evaluation methodology have been widely used for motion detection[8,9]. L-K's method does not perform well enough in dense flow field[8]. On the contrary, the H-S method can detect any minor motion and provide a 100% flow field which we need. Therefore, the H-S method is used for optical flow computation in our research.

In addition, many new methods for motion detection have been proposed in recent years. The survey in [10] suggests that subspace learning models are well suited for background subtraction. Some simple and useful motion detection methods just as [11] combine optical flow with Principal Component

Analysis (PCA) and significantly reduce the dimension of data. Furthermore, some improved subspace learning models like Robust Principal Components Analysis (RPCA)[12,13] and Robust PCA via Principal Component Pursuit[14] have been used to model background or detect moving foreground. As a new trend, using saliency detection methods and models could be very helpful in locating and detecting objects [15,16].

1.2 Challenges in Complex Scenes

Detecting foreground moving objects in the complex outdoor scenes has become a new hot issue since bad weather such as fog or rain, the dynamic background caused by camera's movement, and shaking trees or water flows all bring great difficulties to robots' outdoor task in avoiding the interference of complex scenes and locating the major moving objects correctly. We focus our research on challenging sequences with complex scenes such as the examples shown in Fig.1 and Fig.2. The scene in Fig.1 was shot in foggy weather, it can be seen that the contrast between the background and foreground moving objects is very low. The images in Fig.2 were captured by an on-board camera, the background and foreground are moving simultaneously at different speeds, and many small disturbances like traffic lights or distant pedestrians also exist.

t=1 t=3 t=5

Fig.1 Sequences from an outdoor scene with bad weather

t=475 t=500 t=525

Fig.2 Sequences from an on-board camera

Detecting the motion of objects in complex environments is a critical step in the task of vision-based autonomous moving robots. A variety of techniques have been developed in the last decade for suppressing of dynamic background or distinguishing of major moving objects from complex scenes. Considering only small differences exist between consecutive frames in a video clip compared to the large number of video frames, background modeling and subtraction methods and their extended methods are popular. However, the single Gaussian model has less-than-ideal performance for complex or dynamic backgrounds, thus a mixture of several Gaussians is proposed to model the background. By far the most popular is the Gaussian Mixture Model (GMM) [17,18] which models the distribution of the values at every pixel observed over time by a weighted mixture of Gaussians. A GMM background model can handle the multimodal nature of many real scenes and get good results for repetitive background motion, for example, tree leaves or branches. The GMM model has become a very popular algorithm amongst the computer vision field from the beginning, and it is still under development by researchers, with enhanced algorithms such as adaptive GMM[19,20]. However, GMM and its enhanced algorithm are able to tolerate the variations of the background scene, but still have some defects, such as sometimes losing all the foreground objects whose motions in consecutive frames are not strong enough.

As a powerful pixel-based background subtraction technique, the ViBe algorithm[21] chooses background samples to build background models. It first stores a set of pixel intensity values at the same location or in the neighborhood for a number of previous frames, and then compares the current pixel intensity value to this set. If the current pixel does not belong to the background, the algorithm will update the background model through randomly selecting background samples. Many applications have shown that it performs much better than other state-of-the-art methods in both aspects of computational efficiency and detection accuracy, but we have found that it is so sensitive to the dynamic background that it detects many redundant motions in the background, which is adverse for detecting and tracking the major moving objects in the robot's outdoor tasks.

1.3 Main Contributions

It is important for a vision-based robot to discover and detect moving targets of interest. The researchers expect a robot to perceive motions from the visual sensor intelligently like a human. In neuroscience, K.Y. Park et al. [22, 23] found that visual motions are perceived in the medial superior temporal (MST) area of the human brain. They also observed that motion field from visual perception is separated into independent components at the MST area. N. Ohnishi et al. [24, 25] applied the idea into the visual navigation of a robot. They used ICA of optical flows successfully for the robot to avoid collision with obstacles.

In most situations, foreground objects have distinguishable motion patterns, therefore foreground and background can be separated approximately as two different signals by applying ICA of optical flows into moving objects detection. Experiments show that moving objects are detected successfully in complex scenes. We are the first to adopt this strategy to solve the problem, which is quite different from the existing methods.

Afterwards we use PCA to project the foreground optical flows onto an orthogonal subspace with the purpose of detecting moving objects accurately. As a result, PCA can make major motions more prominent and suppress redundant noises in the background. Multiple moving objects can be subsequently extracted effectively from complex outdoor scenes.

In brief, illuminated by the above concept of the biology-inspired vision and subspace-based algorithms, in this paper we propose a novel method to detect multiple moving objects in complex outdoor scenes. This new method combines the advantages of the aforementioned existing algorithms with our proposed solution to provide better performance.

2 Proposed Biologically Inspired Approach

As explained previously, all the existing methods have their own drawbacks when they are applied to complex dynamic scenes. To detect major motions, a hybrid intelligent method is proposed in the paper. Firstly, as optical flows are considered to be fundamental information used by robots for neurobiological data processing [26], the complete motions in the consecutive frames are detected by the H-S optical flow method for subsequent analysis. Then, as mentioned before, the observation that the complex visual motions can be recognized separately by the independent components of motion field at the MST area has inspired us to apply ICA to the optical flow field. ICA can extract statistically-independent sources from signals because the independent components of optical flow field of the natural scene are separate background and foreground optical flows. As dynamic background or bad weather can cause some redundant motions in the foreground component, we then apply PCA to the foreground optical flows to suppress these redundant motions so that the major optical flows corresponding to multiple moving objects can be extracted effectively. Finally, through the post-processing of morphological operation and the combination of temporal motions and spatial frames, multiple objects are detected accurately. The flowchart of the proposed approach is given in Fig. 3.

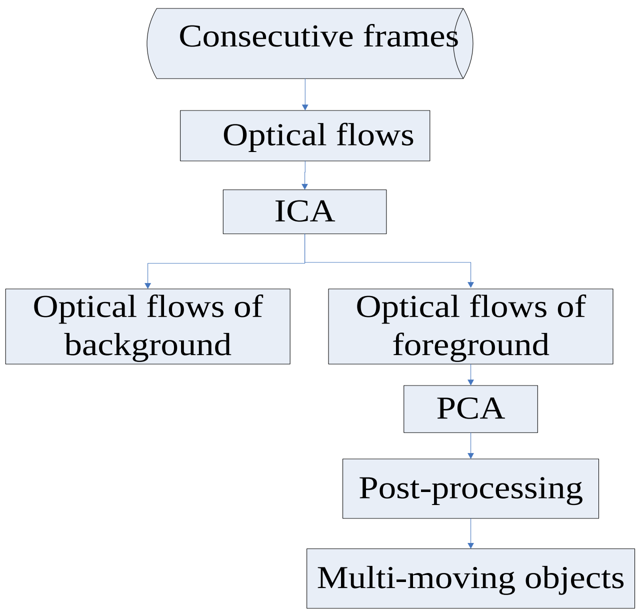


Fig.3 Flowchart of the proposed approach

2.1 Optical Flow Field of Consecutive Frames

As mentioned above, optical flows have been widely used for motion cognition at the MST area of the brain. In [25] the Lucas-Kanade optical flow method with pyramids was used for the purpose of finding the dominant plane. Unlike [25], what we need initially is just a complete optical flow field which includes minor motion of objects. Since the Lucas-Kanade algorithm is a local window based method that can not solve for optical flow in all cases, the Horn and Schunck method [6] is employed in our research.

Let $I(x, y, t)$ represent the brightness of a pixel at (x, y) coordinates and the t^{th} frame. According to [6], the image constraint at $I(x, y, t)$ with Taylor series can be expressed by the following equation:

$$\frac{\partial I}{\partial x}\partial x+\frac{\partial I}{\partial y}\partial y+\frac{\partial I}{\partial t}\partial t=0 \qquad (1)$$

And its simple form is

$$I_x u + I_y v + I_t = 0 \qquad (2)$$

where $I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t}$ are the derivatives of the image at (x,y,t) in the corresponding directions, and $u = \partial x/\partial t$ and $v = \partial y/\partial t$ are the x and y components of the velocity or optical flow of $I(x, y, t)$, respectively.

The H-S method uses a global smoothness term to constrain the estimated velocity field $V=(u,v)$, minimizing

$$\int_D (\nabla I \cdot V + I_t)^2 + \lambda^2(\|\nabla u\|_2^2 + \|\nabla v\|_2^2)dx \qquad (3)$$

which is defined over a domain D. The magnitude of λ reflects the influence of the smoothness term. Iterative equations are used to minimize (3) and obtain image velocity:

$$u^{k+1} = \bar{u}^k - \frac{I_x[I_x\bar{u}^k + I_y\bar{v}^k + I_t]}{\alpha^2 + I_x^2 + I_y^2} \qquad (4)$$

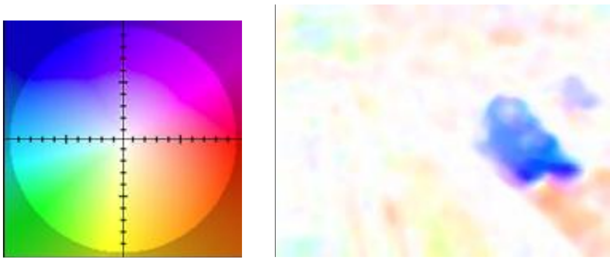
$$v^{k+1} = v^k - \frac{I_y[I_x\bar{u}^k + I_y\bar{v}^k + I_t]}{\alpha^2 + I_x^2 + I_y^2} \qquad (5)$$

where k denotes the iteration number, α is a weighting factor, u^0 and v^0 denote initial velocity estimates which are set to zero, and $\bar{u}^k$ and $\bar{v}^k$ denote neighborhood averages of u^k and v^k. Through the above constrained minimization, the optical flow vector (u^{k+1}, v^{k+1}) for the (k+1)-th frame is calculated. The benefit of the optical flow method is that the optical flow vector can reflect the motion intensity by the length of vector and the motion orientation by the direction of vector. Figure 4 shows the consecutive frames from a bad weather scene and the corresponding optical flow field.

(a)k^{th} frame (b) $(k+1)^{th}$ frame (c) the corresponding optical flows

Fig.4 Optical flow field by H-S method

For the visualization of optical flow, we use the Middlebury color coding of optical flow vectors. The Middlebury optical flow benchmark is a new benchmark and evaluation method for optical flow algorithms[9]. Fig.5(a) shows the color strategy of displaying the optical flow vector field, the direction of the optical flow vector is coded by hue and the length is coded by saturation. The color of a pixel represents a vector starting in the middle of the coding circle and pointing to the corresponding color. The Middlebury color coding of the example in Fig.4 can be seen in Fig.5(b).



(a) Middlebury color coding scheme (b) Middlebury color coding of optical flow in Fig.4

Fig.5 Middlebury color coding result

2.2 Independent Component Analysis of Optical Flows

ICA is a statistical method with the purpose of finding a linear representation of nongaussian data so that the components are statistically independent, or as independent as possible[27]. Just as ICA separates signals into independent components, the MST area in the brain separates the motion field into independent visual components[22,28] as an important visual motion perception mechanism. According to previous research, we assume that the optical flow field of a natural complex scene is a linear combination of the foreground and the background optical flow field, so we can apply ICA to consecutive optical flow fields to detect the foreground by separating the moving foreground and the relatively still background.

In order to give a simple representation, ICA is often sought as a linear transformation of the original data. Let us assume that $x=(x_1,x_2, ...,x_m)^T$ is an observed zero-mean m-dimensional random variable, and $s = (s_1, s_2, ..., s_n)^T$ is its n-dimensional transform. Then ICA will determine a constant (weight) matrix W so that the linear transformation of the observed variables

$$s = Wx \tag{6}$$

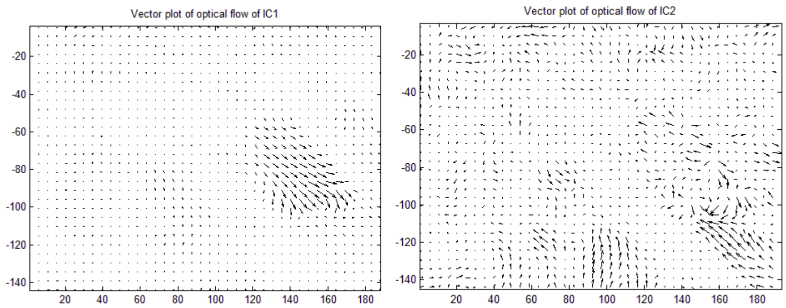
has some suitable properties[27] that the transformed components s_i are statistically as independent from each other as possible. W is a constant (weight) matrix.

When applied to the optical flow field (u^k, v^k) (u^k or v^k is a $m \times n$ matrix), the observed variable x_k can be taken from the matrix u^k and v^k. Let u_0 and $v_0 \in R^{m\times n}$ be the corresponding vector representation of u^k and v^k with pixels arranged in lexicographic order, then the observed vector $x_k = \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \in R^{m\times n\times 2}$ is generated. Thus, through this transformation, a three-dimensional $(m\times n\times 2)$ matrix of optical flow field (u^k, v^k) changes into a one-dimensional vector x_k.

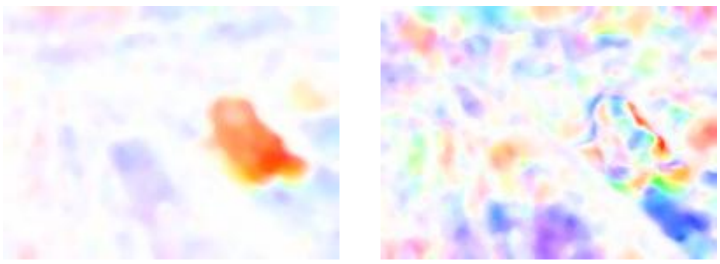
As mentioned above, we assume that optical flow fields are linear combinations of optical flow fields of the foreground and the background which are seen as different signal sources, so we apply ICA to two consecutive optical flow field vectors (x_k, x_{k+1}) observed by the outdoor robot in a complex environment, and can get the optical flow fields of the foreground and the background separately as follow:

$$(U,V)=\alpha_1 (U,V)_{foreground}+\alpha_2 (U,V)_{background} \quad (7)$$

where (U,V) is defined as the two consecutive optical flow field which is equal to (x_k, x_{k+1}), $(U,V)_{foreground}$ and $(U,V)_{background}$ are the optical flow fields of the moving foreground and the relative still background. The observed optical flow field *(U,V)* is approximately expressed by a linear combination of $(U,V)_{foreground}$ and $(U,V)_{background}$, where α_1 and α_2 are the mixture coefficients. Fig.6 shows the results of the foreground and the background optical flow fields from an outdoor scene with fog and their corresponding Middlebury color coding.



(a)Foreground optical flow field (b) Background optical flow field



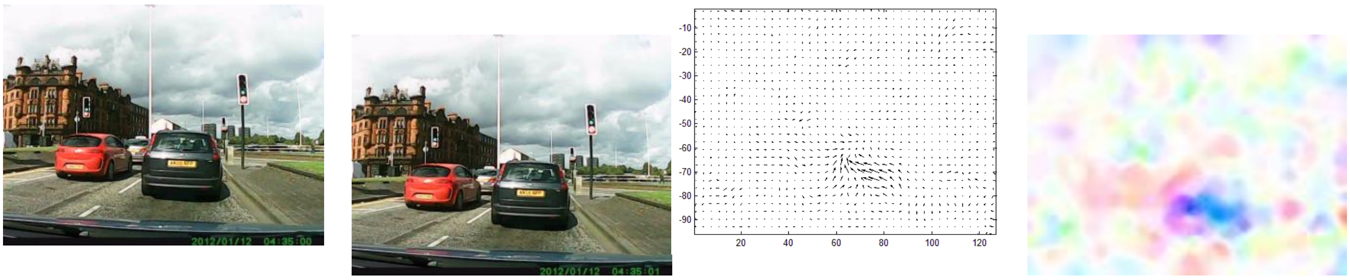
(c) Middlebury color coding of (a) (d) Middlebury color coding of (b)

Fig. 6 Foreground and background optical flow field after ICA from bad weather scene

In comparison with Fig.4 and Fig.5, Fig.6 verifies that the optical flow field from bad weather scene can be separated into the optical flows of the foreground and the background by ICA. From Fig.4(c) and Fig.6(a), it is not difficult to see these optical flows behind the bigger car resulted from fast motion in Fig.4(c) has been suppressed in foreground optical flows in Fig.6(a). Therefore, the major foreground objects are more accurate. In a word, the motion or noise from the background is reduced or suppressed in the foreground component.

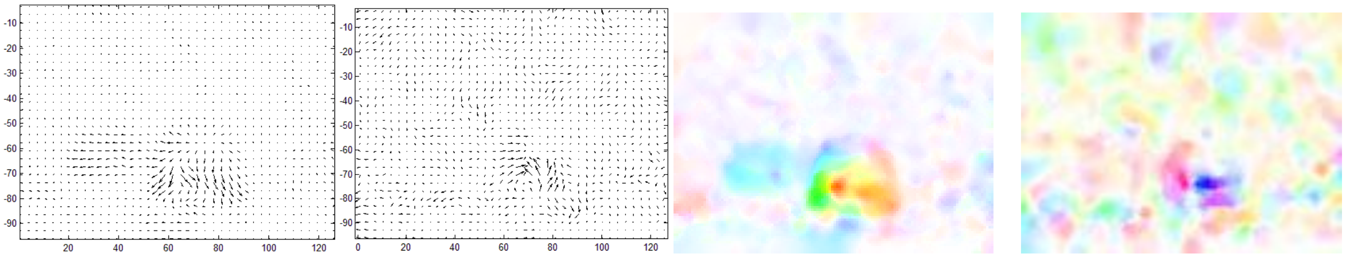
In order to test whether the idea is robust, we carried out another experiment on the vehicle-mounted video(Fig.7). This video has the problem of both the objects and the background moving simultaneously. In addition, in the scene there are dynamic lights and clouds which bring more difficulties to object detection.

Figure 7 shows the first and second important steps in our method. Fig.7(c) is corresponding optical flow field of consecutive frames(Fig.7(a),(b)) in the vehicle-mounted video. Fig.7(e) and Fig.7(f) are the foreground optical flows and the background optical flows after performing ICA. In this very complex scene, we can see optical flow based ICA can effectively extract the foreground objects (Fig. 7(e)) compared to the optical flow result (Fig. 7(c)). The major moving objects are more obvious and the majority of noise from the background is suppressed in Fig. 7(e).



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

(a)(b):Consecutive frames in vehicle-mounted video;(c)(d):Corresponding optical flows and Middlebury color coding



(e) (f) (g) (h)

(e)(f):The foreground and the background optical flows after the ICA; (c)(d): Middlebury color coding of (e)(f)

Fig. 7 Foreground and background optical flow field after ICA from vehicle-mounted camera

From the comparative experiments, it is easy to see that as one of the independent components, the foreground optical flows can detect major motions of objects effectively. Therefore, ICA is suitable for separating the optical flows of the real complex scene into the independent optical flow sources, and this capability is demonstrated to be effective even in bad weathers or in dynamic backgrounds from the complex outdoor environment.

2.3 Principal Component Analysis of Foreground Optical Flows

Although after performing optical flow and ICA, the results improve, some problems still exist. For example, the foreground optical flows include some redundant obvious motions or noises, which will affect the accuracy of the motion detection, since the dynamic background arises from the moving camera or disturbances in outdoor scenes. In order to get satisfactory results, we also use PCA [29] to project the original data onto orthogonal subspaces whose directions are not mutually correlated, with the major data information present in the first several principal components. As a result, the foreground optical flows can be enhanced by detecting major motions. Figure 8 shows the PCA procedure for the foreground optical flows.

$(U,V)_{foreground}$ is the optical flow field of the foreground. A mask of size 3×3 slides respectively through $U_{foreground}$ and $V_{foreground}$ is shown in the first step of Fig.8. A matrix X is constructed using all the data covered by the mask whose center pixel is (i,j). The covariance matrix is calculated as follow[11]:

$$\Sigma = \overline{X}^{T}_{2\times 9}\overline{X}_{\times 9} \tag{8}$$

where $\overline{X}$ is the optical flow matrix after the mean is removed.

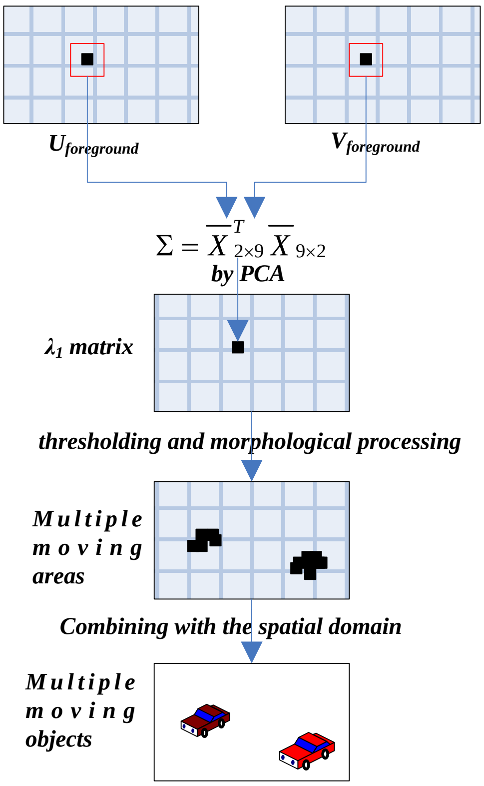
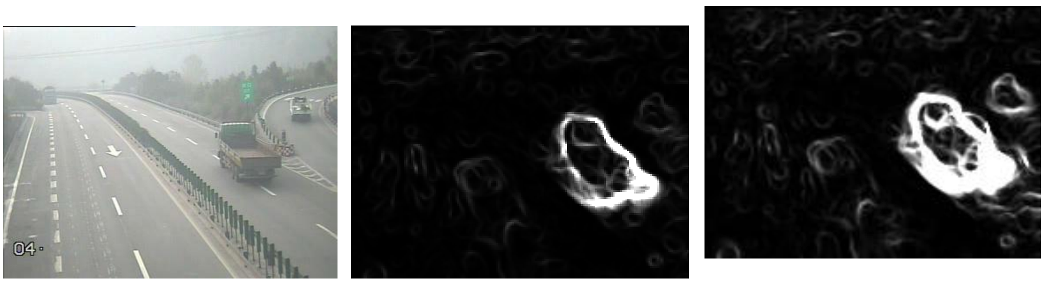


Fig.8 Multiple moving objects detection by PCA based on the foreground optical flow field

After the eigen-decomposition, the major eigenvalue λ_1 is assigned to the central pixel of the mask. Motion detection is accomplished by thresholding the eignenvalue(λ_1) matrix. In order to verify the effectiveness and necessity of this step, a comparison of PCA for optical flows[11] is shown in Fig.9.

Fig.9(a) is the original video frame, there are two moving trucks, a bigger one and a smaller one. Fig.9(b) is visualization of major eigenvalue λ_1 matrix from PCA for all optical flows. Fig.9(c) is also visualization of major eigenvalue λ_1 matrix from PCA for foreground optical flows. We can see that these moving trucks in Fig.9(c) have stronger brightness, which makes these objects more prominent and the noises in the background not so obvious relatively. This advantage is quite beneficial to the next step of thresholding the eignenvalue matrix to detect moving objects.

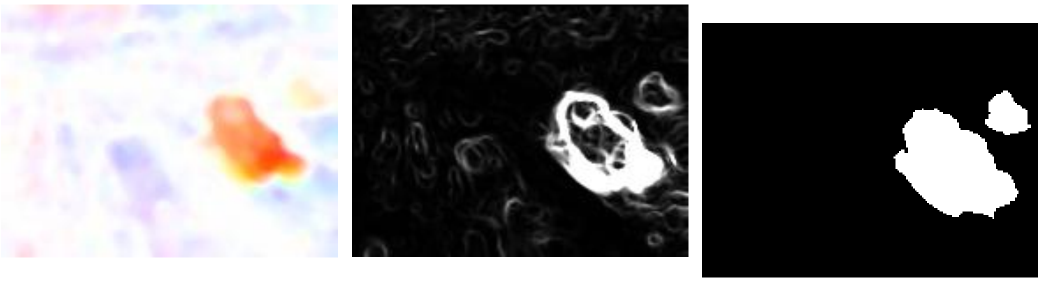


(a)The original video frame (b) PCA based on optical flows (c) PCA based on foreground optical flows

Fig.9 Comparison between major eigenvalue λ_1 matrixs of PCA for optical flows and PCA for foreground optical flows

Then, by thresholding the matrix of eigenvalue $\lambda_{1,}$ and performing a morphological filling operation, we can acquire the complete major motion areas. Finally, by combining motion areas with the Canny edge feature of the spatial frame, multiple moving objects can be detected accurately.

Fig.10 shows experimental results of every step in PCA based on the foreground optical flows. Fig.10(a) is the Middlebury color coding of the foreground optical flows (Fig.6(a)) in which there are two foreground moving objects of red coding, but some motions or noises in the background still exist. From the image of $\lambda 1$ matrix after PCA(Fig.10(b)), we can see that the motions of two trucks are more obvious and they can be extracted well (as shown in Fig.10(c)) by subsequent thresholding and morphological processing. Figs.10(d)~(f) show the following steps that combining the temporal motion with the spatial Canny edge feature of the corresponding frame with the purpose of producing the final accurate multiple moving objects.



(a) (b) (c)

(a) Middlebury color coding of foreground optical flows (b)the image of $\lambda 1$ matrix after PCA for foreground optical flows (c) thresholding and morphological processing result

(d) (e) (f)

(d) Canny edge detection result of the corresponding frame (e) the mask of combining the motion areas and spatial edge feature (f) multiple-moving objects detection result

Fig.10 Result of every stage of multiple-moving objects detection by PCA based on foreground optical flows

When any objects that are moving at exactly the same speed as the robot, although this special situation is less likely to be encountered in practice, it is still important to be well considered in the paper. Actually this special case makes it difficult for the robot's vision system to identify moving objects, our method can overcome the potential drawback in one of the following ways. The first method is by appropriately adjusting the speed of a vehicle-mounted platform or robot periodically, we can avoid omission of moving objects with the same speed. Another strategy is detecting new moving objects at first, then tracking them to avoid losing them in case the speed of the object(s) becomes the same as the robot/vehicle-mounted platform.

3 Simulation Results

The first dataset of our own including the two different long videos with complex outdoor scenes are preliminarily used for evaluating the performance of the proposed method. One is in foggy weather, another scene is from a vehicle-mounted camera with variable background due to the camera moving forward while the targeted cars are also moving. The first video has RGB image sequences of size 384×288 pixels and the vehicle-mounted video has RGB images of size 320×240 pixels. Our method was implemented in MATLAB R2010.

We first halved the width and height of the input images so that the program could run efficiently. To verify the effectiveness of the proposed technique, as shown in Fig.11 and Fig.12, some meaningful comparison experiments were conducted separately with H-S Optical Flow[6], Optical Flow and ICA [25], Optical Flow and PCA-based motion detection[11] which is one of methods using subspace learning to detect motion, adaptive Gaussian Mixture Models(GMM) background modeling[20], and the popular ViBe algorithm[21].

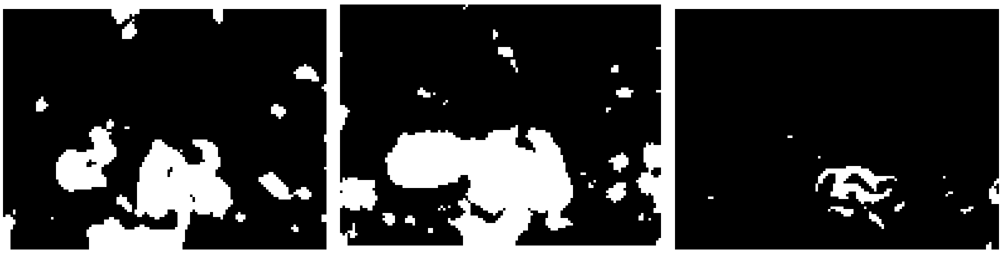
The difficulty in the first outdoor scene lies in the fact that the strong fog in the video makes the motion of the smaller truck inconspicuous and causes changes in illumination, which causes noise. From the first comparison experimental results(Fig.11), we find that the classic H-S method can detect all the obvious motion but contains a lot of noise due to bad weather or changes in background (Fig.11(a)),and there is a significant optical flow ghosting behind the vehicle due to the past motion of the vehicle. From the detected foreground based on optical flows and ICA(Fig.11(b)), we can see the ghosting behind the vehicle is much less than Fig.11(a), but there are still some redundant noise which is mistakenly recognized as motion. PCA based on optical flow misses the smaller moving object despite detecting the bigger one (Fig.11(c)). The classic adaptive GMM background modeling (Fig.11(d)) and the popular ViBe algorithm(Fig.11(e)) all detect the foreground motions but also detect ghosting. In comparing to all the above methods, our method(Fig.11(f)) can detect moving objects accurately and also eliminate ghosting.

(a) Optical flow (b)Optical flow and ICA (c)Optical flow and PCA

(d) Adaptive GMM (e) ViBe (f) Our method

Fig.11 Object detection compared with Optical flow[6], Optical flow and ICA[25], Optical flow and PCA[11], Adaptive GMM[20], ViBe [21] for bad weather scene

The second preliminary experiment was conducted on the vehicle-mounted video(Fig.2) with the dynamic background and moving objects, which also contained dynamic lights and clouds. This kind of complex scene makes foreground motion detection more challenging. From the comparative experiments shown in Fig.12, it is obvious that simply combining the optical flows with ICA is not good enough for the final objects detection result given in Fig.12(b). But there is a good indication from Fig.12(b) that the detected moving objects are more complete compared to the optical flow result (Fig.12(a)). Optical flow based PCA still does not detect the smaller moving object just as in Fig.11(c). From Fig.12(d) and (e), it is clear that both the adaptive GMM background modeling method and the ViBe method have not performed well. Adaptive GMM loses the objects since the objects are not moving quickly, and the ViBe method detects lots of background motion from the obvious moving background. However, our method based on both ICA and PCA (Fig.12(f)) outperforms those classic methods as it is able to detect the moving objects effectively in the complex scene.



(a) Optical Flow (b)Optical flow and ICA (c)Optical flow and PCA

(d) Adaptive GMM (e) ViBe (f) Our method

Fig.12 Object detection comparison with Optical flow[6], Optical flow and ICA[25], Optical flow and PCA[11], Adaptive GMM[20], ViBe [21] from vehicle-mounted camera

In order to explore effect of our method, we also carried out further extensive comparative evaluation using an additional dataset[30] with complex scenes, including manually segmented ground truths.

This change detection benchmark dataset has 11 video categories including all kinds of situations, such as low frame rate, PTZ, thermal, turbulence, etc. Since this paper is proposed for complex outdoor scenes, we select the data of ‘bad weather’ and ‘dynamic background’ to conduct statistical evaluation. The ‘bad weather’ category has four videos describing different moving foreground in heavy snow weather. And there are six videos in ‘dynamic background’ category depicting outdoor scenes with strong redundant motion in background. Two videos represent boats on shimmering water, two videos show cars passing next to a fountain, and the last two depict pedestrians, cars and trucks passing in front of a tree shaken by the wind.

In order to show every methods’ respective characteristic, we also list partial object detection results in Fig.13.

Fig.13 Partial object detection results comparison using dataset[30]

Considering that the 'bad weather' videos in dataset[30] include only various snow weather, we added our own videos of 'fog' with manually segmented ground truths into the statistic quantitative comparison experiments in order to make the results more comprehensive and objective. The total frames of these datasets are more than 40000.

Two popular metrics Recall (= TP / (TP + FN)) and Precision(=TP / (TP + FP)) [30] are used to assess the output of a foreground detection algorithm given a series of ground-truth segmentation maps. These metrics involve the following quantities: the number of true positives (TP), which counts the number of correctly detected foreground pixels; the number of false positives (FP), which counts the number of background pixels incorrectly classified as foreground; and the number of false negatives (FN), which accounts for the number of foreground pixels incorrectly classified as background. Values of Recall and Precision are all the bigger, the better.

For each method, the above metrics are first computed for each video in each category. For example, the Recall metric for a particular video *v* in a category c is computed as follow: $\mathrm{Re}_{v,c} = TP_{v,c} / (TP_{v,c} + FN_{v,c})$.

Then, a category-average metric for each category is computed from the values of the metric for all videos in a single category. For example, the average Recall metric of category *c* is given by $\mathrm{Re}_c = \frac{1}{|N_c|}\sum_v \mathrm{Re}_{v,c}$, where $|N_c|$ is the number of videos in category *c*.

At last, overall-average Recall is given by $\mathrm{Re} = \frac{1}{|N|}\sum_c \mathrm{Re}_c$, where |N| is the number of categories we selected.

Table 1 and 2 are Recall and Precision quantitative results comparison on 'Bad Weather' and 'Dynamic Background' videos respectively, table 3 is overall quantitative results comparison.

Table1: Quantitative results comparison on 'Bad Weather' videos

'Bad Weather' Videos	Optical flow[6]	Optical flow and ICA[25]	Optical flow and PCA[11]	Adaptive GMM[20]	ViBe[21]	Our method
Recall (Re)	0.4600	0.4455	0.1296	0.5693	0.4019	0.5838
Precision (Pr)	0.3472	0.4324	0.2511	0.6821	0.7833	0.4952

Table2: Quantitative results comparison on 'Dynamic Background' videos

'Dynamic Background' Videos	Optical flow[6]	Optical flow and ICA[25]	Optical flow and PCA[11]	Adaptive GMM[20]	ViBe[21]	Our method
Recall (Re)	0.6267	0.7838	0.0494	0.8019	0.5330	0.8079
Precision (Pr)	0.6071	0.7092	0.2545	0.6328	0.3205	0.7805

Table3: Overall quantitative results comparison

overall Videos	Optical flow[6]	Optical flow and ICA[25]	Optical flow and PCA[11]	Adaptive GMM[20]	ViBe[21]	Our method
Recall (Re)	0.5434	0.6147	0.0895	0.6856	0.4675	0.6959
Precision (Pr)	0.4771	0.5708	0.2528	0.6575	0.5519	0.6379

From these statistic quantitative results in table 1 and table 2, we see that the proposed method gets highest Recall and Precision values for dynamic background sequences among all above methods, which means it is quite suitable for this type of complex background. Visually, the results of our method for dynamic background are close to ground-truth references shown in right five columns in Fig.13, especially for it wiping off the noises resulted from dynamic background such as shaking leaves and water flows. However, other methods, such as adaptive GMM which has good overall performances(shown in table 3), still generates lots of background noises(Fig.13). The popular ViBe also has the same disadvantage as adaptive GMM in dynamic background(Fig.13). From Fig.13, it is clear that Optical flow and ICA method suppresses the noises in background to some extent in comparison with Optical flow method, which is demonstrated by statistical comparison results in tables, but is still not good enough. We also see Optical flow and PCA method just detects contours of motion, this leads to its lowest Recall and Precision among all methods.

From detection results for ‘Bad Weather’ videos in left two columns in Fig.13, we can see our method is not good enough for minor objects detection in heavy snow (shown as the ‘blizzard’ column in Fig.13), it mistakes some snow areas as foreground objects, but it still performs much better than most methods such like Optical flow, Optical flow and ICA, Optical flow and PCA, as same as the statistical comparison results in table1.

From the overall statistic results comparison in table 3, we can see our method gets highest Recall among all comparison methods, and adaptive GMM gains highest Precision value. The proposed method still needs to be improved for minor objects detection in bad weather by restraining background noises just like heavy snow. And enhancing accuracy of foreground optical flows will also be helpful to increase Precision of our method in some degree. The thinking about these disadvantages will guide us to improve this biologically inspired vision method in the following work.

4 Conclusion and Future Work

Inspired by motion cognition at the MST area of the human brain, we propose a novel approach which combines ICA of optical flow with PCA to detect moving objects in the context of complex outdoor scenes. This new approach is totally different from existing widely used foreground object detection methods. Firstly, the classic H-S optical flow method is adopted to get a complete optical flow field including as much minor motion as possible. The ICA algorithm is then applied to the complete optical flow field of the complex outdoor scene, which can approximately separate the optical flows of both the foreground and background of consecutive frames. The results have shown that optical flow combined with ICA was much better than the pure optical flow method for motion detection of objects. In order to solve the problem of obvious motion coming from the dynamic background or bad weather, PCA was applied to optical flows of the foreground components to enhance the major motion and relatively suppress the redundant motion from the background. Finally, by using morphological processing and combining motion areas with spatial edges, the major optical flows corresponding to multiple moving objects can be well detected. The experimental results have demonstrated that the proposed ICA and PCA-based approach can extract multiple moving objects effectively in complex outdoor scenes.

As shown in the statistic quantitative results, the proposed method needs to be improved for minor objects detection in bad weather in the following work, through restraining background noises. Thinking about how to get more accurate preliminary foreground optical flows will also be good for final detection accuracy. In addition, since optical flow information is often used in combination with electronic data from other sensors to get richer inputs, we will consider using the data from more sensors such like positioning of the robot, which can be combined with optical flows, to have a better sense of relative motion due to the robot moving versus other objects moving, then the robot can carry on its navigation, detecting or tracking interested object well. Furthermore, making our method more computationally efficient for real-time application will also be investigated in future work. We will develop a predictive moving objects detection method that moving objects in current frame can be predicted by the results of previous frame, which could enhance the operation speed greatly.

Considering identification of salient non-moving objects in the background would be important for the case of a moving robot, although we have not attempted to identify salient non-moving objects at this stage, but we will try to fulfill this important outstanding task for future work.

Acknowledgements

This research is supported by The Royal Society of Edinburgh (RSE) and The National Natural Science Foundation of China (NNSFC) under the RSE-NNSFC joint project (2012-2014) [grant number 61211130309] with Anhui University, China, and the “Sino-UK Higher Education Research Partnership for PhD Studies” joint-project (2013-2015) funded by the British Council China and The China Scholarship Council (CSC). Amir Hussain and Erfu Yang are also funded, in part, by the UK Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/I009310/1], and the RSE-NNSFC joint project (2012-2014) [grant number 61211130210] with Beihang University, China. We also thank Dr Andrew Abel for providing some good suggestions for this paper.

REFERENCES

[1] Paragios N, Deriche R. Geodesic active contours and level sets for the detection and tracking of moving objects. IEEE Transactions on Pattern Analysis and Machine Interface, 2000;22(3): 266-280.

[2] Edelstein A, Rabbat M. Background Subtraction for Online Calibration of Baseline RSS in RF Sensing Networks. IEEE Transactions on Mobile Computing. 2013; 12(12):2386-2398.

[3] Heikkila M, Pietikainen M. A texture-based method for modeling the background and detecting moving objects. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006; 28(4):657-662.

[4] Cavallaro A, Ebrahimi T. Video object extraction based on adaptive background and statistical change detection. In: Proc. of SPIE VCIP, 2000.

[5]Shimada A, Yoshinaga S, Taniguchi R. Adaptive Background Modeling for Paused Object Regions. ACCV2010 Workshops, 2010; Vol. 6468, p.12-22.

[6] Horn BKP, Schunck BG. Determining optical flow. Artificial Intelligence. 1981; Vol.17, Issues 1-3, p. 185-203.

[7] Lucas B, Kanade T. An iterative image registration technique with an application to stereo vision. In: Proceedings of the 7th IJCAI, 1981; p. 674-679.

[8] Sun D, Roth S, Black MJ. Secrets of optical flow estimation and their principles. 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2010; p.2432-2439.

[9] Baker S, Scharstein D, Lewis J, Roth S, Black MJ, Szeliski R. A database and evaluation methodology for optical

flow. International Journal of Computer Vision. 2011; 92(11):1-31.

[10] Bouwmans T. Subspace learning for background modeling: A survey. Recent Patents on Computer Science. 2009; 2(3):223-234.

[11] Liu K, Yang H, Ma B, Du Q. A joint optical flow and principal component analysis approach for motion detection, 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP).March 2010; p.1178-1181.

[12] Candes E, Li X, Ma Y, Wright J. Robust Principal Component Analysis?. ACM, May 2011; Volume 58, No. 3.

[13] Guyon C, Bouwmans T, Zahzah E. Foreground Detection by Robust PCA solved via a Linearized Alternating Direction Method. International Conference on Image Analysis and Recognition(ICIAR) 2012; p.115-122.

[14] Bouwmans T, Zahzah E. Robust PCA via Principal Component Pursuit: A Review for a Comparative Evaluation in Video Surveillance.Special Isssue on Background Models Challenge, Computer Vision and Image Understanding(CVIU). 2014.

[15]Tong Yubing, Faouzi Alaya Cheikh, Fahad Fazal Elahi Guraya, Hubert Konik, Alain Trémeau. A Spatio-temporal Saliency Model for Video Surveillance. Cognitive Computation , 2011; 3(1): 241-263.

[16]G. C. H. E. de Croon, E. O. Postma, H. J. van den Herik. Adaptive Gaze Control for Object Detection. Cognitive Computation, 2011; 3(1): 264–278.

[17] Stauffer C, Grimson E. Learning patterns of activity using real time tracking, IEEE Trans. Pattern Anal. Mach. Intell., 2000; 22(8): 747-757.

[18] Mukherjee D, Wu QMJ, Nguyen TM. Multiresolution Based Gaussian Mixture Model for Background Suppression, IEEE Transactions on Image Processing. 2013; 22(12): 5022-5035.

[19] Lee D. Effective Gaussian mixture learning for video background subtraction, IEEE Trans. Pattern Anal. Mach. Intell. 2005; 27(5): 827-832.

[20] Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction. Proceedings of the 17th International Conference on Pattern Recognition. 2004; Vol.2, p.28-31.

[21] Barnich O, Van Droogenbroeck M. ViBe: A Universal Background Subtraction Algorithm for Video Sequences, IEEE Transactions on Image Processing. 2011; 20(6):1709-1724.

[22] Park KY, Jabri M, Lee SY, Sejnowski TJ. Independent components of optical flows have MSTd-like receptive fields, in:Proceedings of the 2nd International Workshop on ICA and Blind Signal Separation. 2000; p. 597-601.

[23] Jabri MA, Park KY, Lee SY, Sejnowski TJ. Properties of independent components of self-motion optical flow. In: Proceedings of IEEE International Symposium on Multiple-Valued Logic, 2000; p. 355-362.

[24] Ohnishi N, Imiya A. Dominant plane detection from optical flow for robot navigation. Pattern Recognition Lett. 2006; Vol.27, p.1009-1021.

[25] Ohnishi N, Imiya A. Independent component analysis of optical flow for robot navigation, Neurocomputing. 2008; Vol. 71, Issues 10-12, p. 2140-2163.

[26] Vaina LM, Beardsley SA, Rushton SK. Optic Flow and Beyond. Kluwer Academic Publishers, Dordrecht, 2004.

[27] Hyvarinen A, Oja E. Independent component analysis: algorithms and application. Neural Networks. 2000; Vol.13, p.411-430.

[28] Zemel RS, Sejnowski TJ. A model for encoding multiple object motions and self-motion in area MST of primate visual cortex. Neuroscience. 1998; Vol.18, p.531-547.

[29] Bishop CM, Nasrabadi NM. Pattern recognition and machine learning. New York: springer, 2006.

[30] www.changedetection.net