SIFT features tracking for moving object with flexible aerodynamic shape based on stroboscopic imaging technique

Yu Yanan, Wang Xiangjun

(State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China)

Abstract: The research on structural dynamics, material aeroelasticity and flight control of a micro aircraft has become an important subject in recent years. In order to improve the detecting method for aircraft aerodynamic characteristics, a tracking approach for objects with flexible aerodynamic shape based on stroboscopic imaging technique was proposed. As scale invariant feature transform (SIFT) algorithm was invariant to image scaling and rotation, illumination changes and local deformations, a method that used SIFT features of any region of interest was presented to detect and track a moving object. The results show that the method is powerful in obtaining object state, location and spatial interrelation. Feature points are adequate to perform reliable matching between different locations of an object in a scene, moreover, to achieve object recognition. Experiments demonstrate that it can achieve good performance using SIFT features based on flexible aerodynamic shape as a detection method in the system.

Key words: scale invariant feature transform (SIFT); flexible aerodynamic shape; stroboscopic imaging; recognition and tracking; feature extraction

CLC number: TP391.4　　**Document code:** A　　**Article ID:** 1007-2276(2011)09-1823-06

采用频闪成像技术的柔性气动外形运动物体SIFT特征跟踪

于雅楠，王向军

(天津大学 精密测试技术及仪器国家重点实验室，天津 300072)

摘　要： 近年来，有关微型飞行器的结构动力学、材料气动弹性以及飞行控制方面的研究受到了高度的重视。为了提高飞行器的空气动力学检测能力，提出了一种基于频闪成像技术的柔性气动外形物体跟踪方法。利用尺度不变特征转换(SIFT)算法提取出的特征对于图像的尺度变换、旋转以及光照变化和局部图像变形等具有的不变特性，提出了一种利用感兴趣区域中SIFT特征对柔性气动外形运动物体进行探测与跟踪的方法。该方法在获取物体的状态、位置以及空间转换关系等方面表现出良好的性能，并且为同一场景中同一物体在不同位置之间的相互匹配提供了可靠保证。实验表明：在该实验系统中，基于频闪成像技术、利用SIFT特征作为柔性气动外形物体探测的方法具有一定的可行性。

关键词： 尺度不变特征转换(SIFT)；　柔性气动外形；　频闪成像；　识别与跟踪；　特征提取

收稿日期：2011-01-21；　**修订日期**：2011-02-19

基金项目：高等学校博士学科点专项科研基金(20070056085)

作者简介：于雅楠(1984-)，女，博士生，主要从事光电探测及视觉测量方面的研究工作。Email: jesuisyyn@126.com

导师简介：王向军(1955-)，男，教授，博士生导师，博士，主要从事精密测试技术及仪器、光电探测技术、影像与视觉测量方面的研究工作。Email: xdocuxjw@vip.163.com

0 Introduction

Due to special structures of micro aircraft vehicles, the management of environmental disturbance plays an important role in numerous technical challenges. To reduce adverse effects of gusts wind conditions and unsteady aerodynamics, exhibit desirable flight stability, and enhance structural durability, more designs employ passive flexible structure instead of conventional rigid structure [1]. Hence, further analyses on unsteady anti-air disturbance mechanism of flexible aerodynamic shape provide a number of references for practical application in various fields. This motivates us to obtain effect of external disturbance on flexible shape.

In current experiments, numerical analyses are not available to validate the predicted shape's deformation and solve problems occurred in practical flight, while another method that uses force sensors to measure deformations and acting force on model has interference effect on results. In our subject, stroboscopic imaging technique and appropriate algorithm are used to get useful database including motion parameters of object. Furthermore, flow field distribution and anti-air disturbance mechanism of flexible adaptive shape will be deduced with these original data.

Although there are a number of different ways to extract feature, most major feature extraction methods use color, texture, shape or spatial relation, all of which as well as Harris corner detector are respectively susceptible to scaling, rotation, affine transform or illumination changes [2-3]. They provide a bad basis for matching or other image processing. Compared with those feature extraction methods, scale invariant feature transform (SIFT) has been proved to perform better on the robustness and distinctiveness.

In recent years, SIFT has been successfully applied in various object detection and recognition [4]. Some applications realize distance measurement as well as depth information collection [5], while some for image registration [6-7] and 3D object recognition. General objects are mainly rigid ones with sharp transitions between different sides. On the contrary, there are also a few studies using SIFT algorithm for nonrigid object that has less structures with high contrast or high edge responses [8-9].

In order to adapt to the changes in scaling, rotation and illumination in our system, SIFT algorithm is used to extract interest keypoints. Then the same points as extracted SIFT features at different locations in a stroboscopic image are searched. Moreover, the affine model of any adjacent locations is calculated and spatial trace and corresponding motion parameters of any point at different times in varied region are found.

1 Review of SIFT

In 1999, Lowe developed a scale invariant feature transform (SIFT) descriptor based on the gradient distribution in the detected regions. The features are invariant to image scaling and rotation, the illumination changes and local image deformations [3,10]. And also, they are fast in computation and robust in the presence of noise. SIFT algorithm is composed of four stages of computation.

1.1 Construct scale space and detect interest points

It has been shown that under a variety of reasonable assumptions Gaussian function is the only possible scale space kernel [11-12]. The scale space function is produced from the convolution of a variable-scale Gaussian, $G(x,y,\sigma)$, with an input image, $I(x,y)$

$$L(x,y,\sigma)=G(x,y,\sigma)*I(x,y) \tag{1}$$

where

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \tag{2}$$

To efficiently detect stable keypoint locations in scale space, scale space extrema are used in the difference-of-Gaussian (DoG) function convolved with the image, $D(x,y,\sigma)$, which can be computed from the difference of two nearby scales separated by a constant multiplicative factor k.

$$D(x,y,\sigma)=(G(x,y,k\sigma)-G(x,y,\sigma))*I(x,y)= L(x,y,k\sigma)-L(x,y,\sigma) \quad (3)$$

In other words, the initial image is incrementally convolved with Gaussian to produce images separated by a constant factor k in scale space. Then adjacent image scales are subtracted to produce DoG images. Local extrema are detected by comparing 26 neighborhoods of a pixel within a set of three DoG images, and if it is an extremum, the pixel is detected as the keypoint.

1.2 Accurate keypoint localization

Once a feature keypoint candidate is produced by scale space extrema detection, an exact model is calculated to describe the nearby data for location, scale and ratio of principal curvatures. This step must remove some unreliable keypoints, which have low contrast or are ambiguous on the edge or boundary. Some thresholds are defined to confine the evaluation of each candidate feature point. If the value exceeds our requirements, the keypoint will be removed.

1.3 Orientation assignment for feature point

The gradient magnitude $m(x,y)$ and orientation $\theta(x,y)$ at each pixel of the image from where the keypoints are detected are respectively computed by the following functions

$$m(x,y)=\sqrt{f_x(x,y)^2+f_y(x,y)^2} \quad (4)$$

$$\theta(x,y)=\arctan\frac{f_y(x,y)}{f_x(x,y)} \quad (5)$$

The weighted histogram of 36 directions is made by gradient magnitude and orientation in the region around the keypoint, and the peak where more than 80% of the maximum value of the histogram presents is assumed to be the orientation of the keypoint.

1.4 Local image descriptor

In the process of interest point's description, an unique descriptor need to be calculated for each point to describe it and its surroundings. A keypoint descriptor is created by first computing the gradient magnitude and orientation at each sample point in a region around the keypoint location. This sample region is divided into the blocks of 4×4, and the histogram of eight directions is made at each block, with the length of each arrow corresponding to the sum of the gradient magnitudes near that direction within the region. Therefore, in this paper a 4×4×8=128 element feature vector is used for each keypoint.

Besides the major advantage of SIFT feature as a local feature described before, other main characteristics of SIFT features are as follows

(1) It is convenient to use different image features to match rapidly and accurately in the feature database.

(2) Large number of SIFT features can be extracted from even minority objects.

(3) Optimized matching algorithm based on SIFT features can meet real-time test requirement.

(4) It is easy to combine SIFT feature points with other features.

2 Proposed approaches

2.1 Extract feature points

With the advantages of SIFT algorithm, distinctive local features are extracted from an image.

Firstly, the SIFT database is created by extracting SIFT keypoints from a template image. Secondly, SIFT keypoints are extracted from an input image. The input of the algorithm is a region of interest defined by user, generally a fraction of an image.

Those SIFT features of the query image, extracted by Lowe's SIFT algorithm, are prepared for further indexing and matching to measure their similarities.

2.2 Match database

The appropriate match for each keypoint is found by identifying its nearest neighbor in the keypoints database, which is defined as the keypoint with minimum Euclidean distance for the invariant descriptor vector. Here s^m and w^n define the mth keypoint in the certain template image and the nth keypoint in the input image respectively. The keypoint with minimum Euclid distance is obtained by calculating the following expression

$$d=\arg_n \min\sqrt{\sum_i^{128}(s_i^m-w_i^n)^2} \tag{6}$$

In practice, some initial matches are usually incorrect because of ambiguous features or features that come from background noise. So a more effective method that compares the distance of the closest neighbor with that of the second-closest neighbor is used. If the distance ratio of closest to second-closest neighbors is greater than threshold, this candidate match will be rejected. In the algorithm, 0.6 is adopted as the distance ratio threshold to estimate the match.

2.3 Affine transform model estimation

The affine transform of a model point $[x\ y]^T$ to other matched point $[u\ v]^T$ can be written as

$$\begin{bmatrix}u\\v\end{bmatrix}=\begin{bmatrix}m_1 & m_2\\m_3 & m_4\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}a\\b\end{bmatrix} \tag{7}$$

where the model translation is $[a\ b]^T$ and the affine rotation, scale, stretch are represented by the m_i parameters.

There are six unknown parameters in the above transform model, so at least three matches are needed to provide a solution. The equation (7) can be rewritten to gather unknowns into a column vector as

$$\begin{bmatrix}x_1 & y_1 & 0 & 0 & 1 & 0\\0 & 0 & x_1 & y_1 & 0 & 1\\x_2 & y_2 & 0 & 0 & 1 & 0\\0 & 0 & x_2 & y_2 & 0 & 1\\x_3 & y_3 & 0 & 0 & 1 & 0\\0 & 0 & x_3 & y_3 & 0 & 1\end{bmatrix}\begin{bmatrix}m_1\\m_2\\m_3\\m_4\\a\\b\end{bmatrix}=\begin{bmatrix}u_1\\v_1\\u_2\\v_2\\u_3\\v_3\end{bmatrix} \tag{8}$$

Any number of further matches can be added to this equation. This linear system can also be written as

$$Ax=b \tag{9}$$

The least-squares solution for the parameters x can be identified by solving the following equation

$$x=[A^TA]^{-1}A^Tb \tag{10}$$

which minimizes the sum of the squares of the distances from the projected model locations to the corresponding image locations. The number of calculated matches depends on the used object and the chosen region. For a large object with heavily textured image, the more the expected number of matches, the higher the false matches.

3 Experiment

3.1 Experimental condition

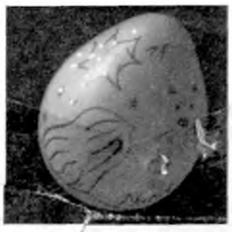
As a non-contact measurement, real-time visual detecting method performs well in recent researches with high precision and speed. Conventional visual detection methods that mainly rely on high-speed photography equipment to record dynamic parameters are limited by experimental conditions. In this paper, the stroboscopic light is installed in the enclosed experiment as the external trigger signal to obtain continuous image data of a flexible aerodynamic shape under specific flow disturbance. In the situation that there is no high-speed image equipment, continuous movement of the high-speed object is recorded and dozens of overlapping images are obtained in one picture by using stroboscopic imaging technique.

An object with rubber material surface is driven by human action to imitate autonomous flight unit such as micro aircraft. Because of the experimental conditions and especially the research object, the most original moment image data of the object affected by external disturbance airflow need to be obtained. Then changes of state, location and spatial interrelation of flexible aerodynamic shape will be calculated. Furthermore, once these changes and values of external disturbance airflow are gained, interaction relations between them could be deduced. Object's different locations in a scene detected are shown in Fig.1.

Fig.1 Object's different locations in a scene

3.2 Experimental result

SIFT feature points extracted from the template image and the candidate images are shown in Fig.2. Fig.(a) is 22 keypoints extracted from the template image. Fig.(b) is 99 keypoints extracted from the candidate image.



(a)

(b)

Fig.2 SIFT feature points extracted respectively

Then the features on the template image are matched with those on the candidate image using effective method mentioned in Sec.2. For the purpose of finding corresponding matches of one point among different locations in an image, the match approach was improved to obtain multiple matched points concurrently. Connects of one object's adjacent locations which indicate match relations between them are shown in Fig.3.

Fig.3 Connects of matching points on the object at different locations

The effect of feature extracting and matching sometimes depends on the amount and contrast of texture information as well as illumination conditions. Under the conditions of high-contrast and rich texture information, more feature points could be extracted and matched well.

Because the used textured object is a flexible aerodynamic shape one, under external disturbance, changes of state or spatial interrelation of varied region at each time point are different. It is needed to define a region of interest as the input image firstly and build up its transform model. In this paper the feature points calculated from the region of interest are shown in Fig.4.

Fig.4 Feature points in the region of interest(Using mouse, we arbitrarily select a fraction as an interest region, and obtain 8 groups of different matched points in this region on the object)

After knowing the target region and time spatial distribution of feature points extracted from this region, an affine transform is computed with the least squares approach. The transform model of detected region between the first and second location is

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0.947\,6 & 0.379\,9 \\ -0.357\,8 & -0.917\,2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 239.427\,5 \\ 66.252\,7 \end{bmatrix} \tag{11}$$

And the model between the second and third location is

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0.946\,5 & 0.232\,4 \\ -0.218\,7 & 0.973\,8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 262.820\,8 \\ 118.560\,9 \end{bmatrix} \tag{12}$$

In this way, continuous transform relations of any region of interest at different locations are acquired. In another word, changes of any point's location, state and spatial interrelation on the object's aerodynamic shape could be obtained to accurately track their motion trace. The tracking result is shown in Fig.5.

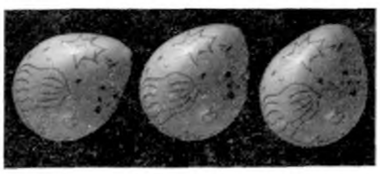


Fig.5 Track any point or line within the arbitrary region

4 Conclusion

The experimental results have proved the practical applicability of proposed features within our system for detecting and tracking textured objects with flexible aerodynamic shape. This method is used to calculate useful database including trace and motion parameters of centroid or other points of interest of the object. In the next research work, the following issues should be focused on as well:

(1) Time-consuming part of SIFT features will be shorten. Compared with other local-based features, the calculation of SIFT feature point positions is relatively slow. Therefore, it is needed to find a faster method to replace part process of it.

(2) Disadvantages of SIFT descriptor will be made up. SIFT does not involve color and global information of feature points. Therefore more useful information and powerful signals in feature description and matching process should be provided to avoid occurrence of mismatch.

(3) Three-dimensional testing using this method will be extended. At the present stage, 2D localization results are studied and detected by tracking a textured moving object with flexible aerodynamic shape. In future, 3D reconstruction technique will be used combining with the method proposed in this paper to obtain object's three-dimensional spatial information.

(4) Related theories will be derived. External disturbance values such as airflow velocity, pressure and temperature need to be combined with state, location and spatial interrelation of the object in the enclosed environment and thus to deduce interaction relations among them.

References:

[1] Ifju P G, Jenkins D A, Ettinger S, et al. Flexible-wing-based micro air vehicles [C]//AIAA Aerospace Sciences Meeting, 2002.

[2] Wang Zhouzheng, Jia Kebin, Liu Pengyu. A novel image retrieval algorithm based on ROI by using SIFT feature matching [C]// International Conference on MultiMedia and Information Technology, 2008: 338-341.

[3] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. ***International Journal of Computer Vision***, 2004, 60(2): 91-110.

[4] Fernando Lopez-Garcia. SIFT features for object recognition and tracking within the IVSEE system[C]//19th International Conference on Pattern Recognition, 2008: 1-4.

[5] Hao Meng, Kang Cheng. Object location technique for binocular stereo vision based on scale invariant feature transform feature points [J]. ***Journal of Harbin Engineering University***, 2009, 30(6): 649-652. (in Chinese)

[6] Liu Li, Peng Fuyuan, Zhao Kun, et al. Simplified SIFT algorithm for fast image matching [J]. ***Infrared and Laser Engineering***, 2008, 37(1): 181-184. (in Chinese)

[7] Liu Xiaojun, Yang Jie, Sun Jianwei, et al. Image registration approach based on SIFT [J]. ***Infrared and Laser Engineering***, 2008, 37(1): 161-165. (in Chinese)

[8] Geng Cong, Jiang Xudong. SIFT features for face recognition [C]// 2nd IEEE International Conference on Computer Science and Information Technology, 2009: 598-602.

[9] Kisku D R, Rattani A, Grosso E, et al. Face indentification by sift-based complete graph topology [C]// IEEE Workshop on Automatic Identification Advanced Technologies, 2007: 63-68.

[10] Lowe D G. Object recognition from local scale-invariant features[C]// 7th IEEE International Conference on Computer Vision, 1999, 2: 1150-1157.

[11] Koenderink J J. The structure of images [J]. ***Biological Cybernetics***, 1984, 50(5): 363-396.

[12] Lindeberg T. Scale-space theroy: a basic tool for analysing structures at different scales[J]. ***Journal of Applied Statistics***, 1994, 21(2): 225-270.

采用频闪成像技术的柔性气动外形运动物体SIFT特征跟踪

作者： 于雅楠， 王向军， Yu Yanan， Wang Xiangjun
作者单位： 天津大学精密测试技术及仪器国家重点实验室，天津，300072
刊名： 红外与激光工程 ISTIC EI PKU
英文刊名： Infrared and Laser Engineering
年，卷(期)： 2011，40(9)

参考文献(12条)

1. Wang Zhouzheng;Jia Kebin;Liu Pengyu A novel image retrieval algorithm based on ROI by using SIFT feature matching 2008

2. Ifju P G;Jenkins D A;Ettinger S Flexible-wingbased micro air vehicles 2002

3. Lindeberg T Scale-space theroy:a basic tool for analysing structures at different scales 1994(02)

4. Koenderink J J The structure of images 1984(05)

5. Lowe D G Object recognition from local scale-invariant features 1999

6. Kisku D R;Rattani A;Grosso E Face indentification by sift-based complete graph topology 2007

7. Geng Cong;Jiang Xudong SIFT features for face recog(mition) 2009

8. Liu Xiaojun;Yang Jie;Sun Jianwei Image registration approach based on SIFT (in Chinese) 2008(01)

9. Liu Li;Peng Fuyuan;Zhao Kun Slmplified SIFT algorithm for fast image matcking (in Chinese)[期刊论文]-Infrared and Laser Engineering 2008(01)

10. Hao Meng;Kang Cheng Object location technique for binocular stereo vision based on scale invariant feature transform feature points (in Chinese) 2009(06)

11. Fernando Lopez-Garcia SIFT features for object recognition and tracking within the IVSEE system 2008

12. Lowe D G Distinctive image features from scale-invariant keypoints[外文期刊] 2004(02)