International Journal of Innovative Computing, Information and Control
Volume **7**, Number **9**, September **2011**
ICIC International ©2011 ISSN 1349-4198
pp. 5285–5298

REAL-TIME VIDEO STABILIZATION BASED ON VIBRATION COMPENSATION BY USING FEATURE BLOCK

Chao-Ho Chen[1], Chao-Yu Chen[2], Chin-Hsing Chen[2] and Jie-Ru Chen[2]

[1]Department of Electronic Engineering
National Kaohsiung University of Applied Sciences
No. 415, Chien-Kung Road, Kaohsiung 807, Taiwan
thouho@cc.kuas.edu.tw

[2]Institute of Computer and Communication Engineering
National Cheng Kung University
No. 1, University Road, Tainan 701, Taiwan

Received May 2010; revised January 2011

Abstract. *Based on feature-block (FB) basis, this paper presents a real-time video stabilization method by using motion compensation for repairing unstable frames captured by a vibrating camera. The basic strategy is first to check if the frame is vibrant and then select the feature block to calculate the vibration vector for compensating the vibration area. The feature block, which is selected from nine fixed observation blocks (FOBs), has certain features and hence can avoid interference of any moving-object intrusion. With the vibration vector, the vibration area can be compensated in order to adjust the vibrating frames to be stable. Experimental results show that the proposed algorithm can effectively stabilize the video frames with moderate vibration (smaller than 20 pixels) in real-time. Besides, a comparison with four reported methods demonstrates that the proposed video stabilization method can cope with the problem of moving-object interference and has less computation time. From the comparison, it also manifests the attractiveness of the proposed approach, when compared with other reported methods.*
Keywords: Video stabilization, Frame vibration, Motion compensation, Feature block

1. **Introduction.** Generally, the video frames captured by a vibrating camera always have unstable appearance which will disturb the visual perception. Besides, many applications of video processing [1-4], such as video compression, video analysis, video object segmentation, moving-object tracking and object recognition, are also seriously affected by frame vibration. Therefore, video stabilization becomes more and more important and should be developed to improve the quality of the frame's appearance and also further the following video processing. Many reported researches [5-14] in video stabilization have been proposed to overcome frame vibration for some specific situations.

Basically, video stabilization process mainly involves three parts: feature selection, vibration vector estimation and frame repair. For the vibration vector estimation, it is necessary to decide the major vibration type existed in the vibrating process and then define a model to calculate the vibration vector. In [5], an affine model is adopted for modeling the vibrant frames, but such a model is unsuitable for large variation because it will cause large error for rotational estimation. To cope with the above problem, the detection of horizontal-line in a distant place is introduced due to the fact that the horizontal-line will not be affected by the small translation of the video camera [6,7]. These reports claimed that the horizontal lines would usually appear in common images and can be detected by edge detection. However, it usually works only for the situation that the horizontal-line will exist and be far away from the camera. So, such a method may fail in the case when the horizontal line is not clear or not far away. On the other hand, an effective model based on tracking a set of features is exploited for video stabilization, but it will suffer image-blurring problem [8]. For obtaining a more accurate vibration vector of a vibrating camera, a probabilistic model is employed to estimate the camera motion and then a Kalman filter is used to remove the undesirable motion [9]. In the method of [10], it divides the video frame into four sub-regions and uses the peak amplitude value of phase correlation surfaces for calculating every local motion vector of these sub-regions. Then, the global motion vector can be derived by eliminating the undesired local motion vectors, but this usually causes erroneous estimation for the global motion owing to obstructing of the moving object. By filling the missing area in the warped frame due to vibration with mosaic processing through sampling spatio-temporal volume patches from different portions of the same video [11], the stable frames can be achieved. Anyway, this method will require heavy computations and there will be alignment error if the moving object intrudes the boundary of the frame. Without estimating the camera motion, the strategy of [12] is to utilize the parameters of the camera and refers the data of rear frame to achieve frame adjustment, but this will be unsuitable for the real-time applications. The method of [13] is also unsuitable for the real-time applications since it estimates the camera motion between consecutive frames by using the optical-flow technique, in which the regularization-based parameter smoothing will need a large number of computations. To achieve real-time video stabilization, the research of [14] is to select a feature block to calculate the motion vector [15], followed by compensating the missing area and adjusting the overlapping area. It can effectively reduce the frame vibration in real-time for the general vibration situations of moderate degree and compensate about 21 pixels. However, the requirement of the absolute background used for reference will make the application to be limited in some special situations. In other words, the above method will fail if there is a moving-object existing in the initial vibrant frames.

However, those video stabilization methods described in the above will suffer computation-intensive processing or perform only for some specific situations. To develop an effective real-time video stabilization method for general situations, this paper introduces the feature block (FB) to calculate the vibration vector which is then used to compensate the vibration area in order to adjust the overlapping area. In the proposed method, a statistical frame-differencing based judgment for whether the frame is vibrant or not is used to aid in achieving the real-time purpose of the video stabilization system. Besides, the FB is selected from nine fixed observation blocks (FOBs) that are employed to avoid the affection by the moving-object interference. The following section will describe the methodology of the proposed video stabilization method involving vibration detection, feature selection, motion estimation and frame repair. In Section 3, experimental results by simulating several representative image sequences with various vibrations are demonstrated and discussed, and a comparison with other four reported methods is also provided. Finally, conclusions are made in Section 4.

2. **Methodology.**

The proposed video stabilization method mainly includes the detection of frame vibration, selection of FB, and calculation of vibration vector, as described in Figure 1. The algorithm is firstly to check whether the video frame is vibrant or not and then to select the FB from nine FOBs. The judgment of frame vibration is based on statistical frame-difference technology. The selection of the FB is realized by checking both if there is any moving-object intruding the visited FOB and if there is the certain feature existed in that FOB. Then, the vibration vector is derived by using the selected feature block to estimate the block's motion vector between two consecutive vibrant frames. Finally, the derived vibration vector is utilized to compensate the vibrating area for adjusting the overlapping area in order to provide the stable frames.

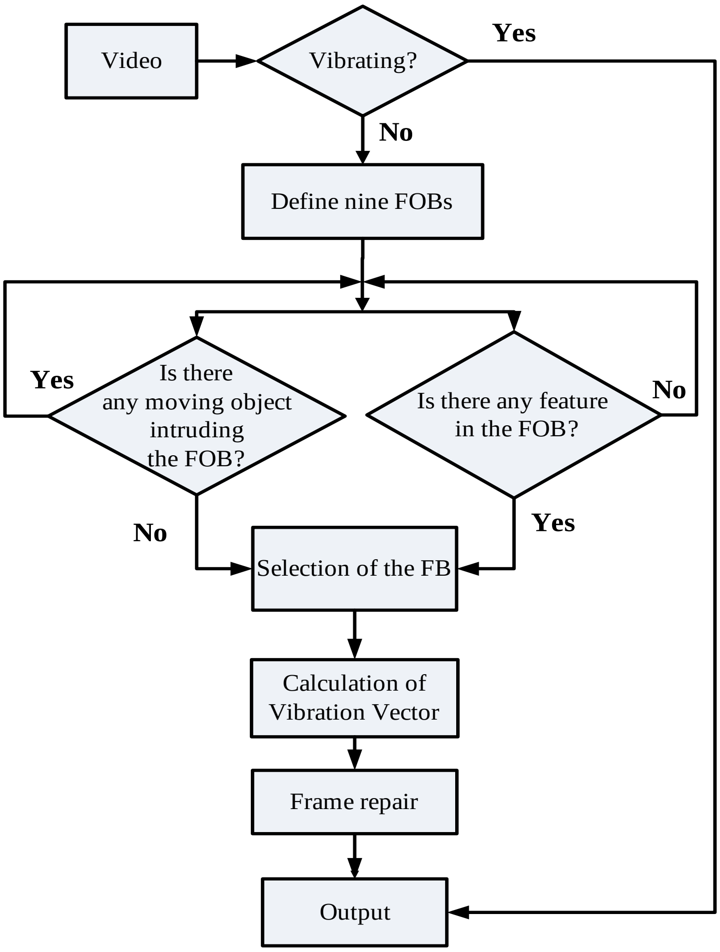


FIGURE 1. The proposed video stabilization algorithm

2.1. **Detection of frame vibration.** Basically, for the stable frames, the number of pixels with smaller values in a frame-difference image is always greater than that of pixels with larger values. Oppositely, for the vibrating frames, the number of pixel with smaller values in a frame-difference image is always smaller than that of pixels with larger values. Based on such a phenomenon, the decision rule for judging if the frames are vibrant can be deduced by the statistical data in the frame-difference images.

The difference, $d_t(x, y)$, between two consecutive frames is defined in the following equation.

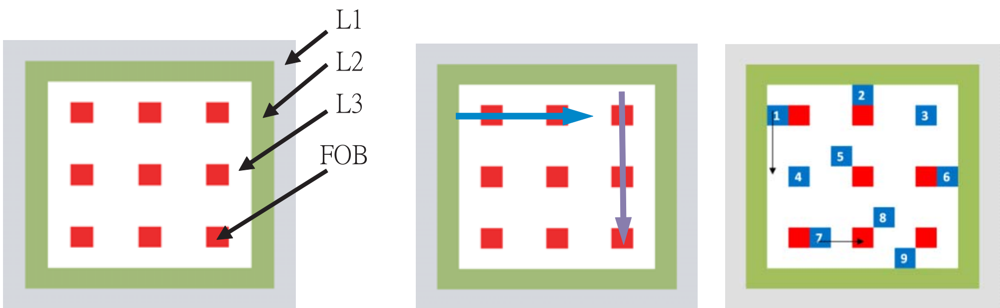
$$d_t(x, y) = |f_t(x, y) - f_{t-1}(x, y)| \tag{1}$$

where $f_t(x, y)$ and $f_{t-1}(x, y)$ denote the values of pixels (x, y) at frame-t and frame-$t-1$, respectively. In a deep analysis on a difference image, the number of pixels whose values are larger than a threshold $Th1$ is calculated and denoted as LN. Then, both frames at t and $t-1$ is considered to be vibrating if LN is larger than another threshold $Th2$. However, some special vibrating situations may make LN to be not larger than $Th2$, because of various vibrating speeds and/or intentional actions. To obtain more accurate judgment, an iterative process for checking if LN is larger than $Th2$ is necessary. The frame is judged as vibrating if $LLN/TN > r$, where LLN denotes the checking times in which LN is larger than $Th2$, TN denotes the total number of checking, and r means the vibration threshold. In the above thresholds, $Th1$ and $Th2$ can be determined by the statistical data of experiments, though they may vary with the content of a difference image. The typical value of r is greater than 0.8. The above deduction for decision rule of judging if the frame is vibrating is described in the following equation.

$$\begin{aligned} &\text{if} \quad LLN/TN > r \\ &\qquad\quad \text{VIBRATING} \\ &\text{else} \quad \text{STABLE} \end{aligned} \tag{2}$$

It is noted that "the frame is vibrating" means frame-t is relatively vibrant to frame-$t-1$.

2.2. **Selection of FB.** In the process of video stabilization, feature selection plays an important role which will lead the following vibration-vector estimation used for frame repair. In theory, the selection of unfixed feature in a frame will spend a large amount of search time and it may also be interfered by a moving-object that is intruding the feature site. To overcome such two problems, this paper proposes a novel feature selection strategy in which the FB is selected from nine FOBs.



(a) (b) (c)

Figure 2. FOBs distribution: (a) basic distribution (red blocks), (b) FOB-intruding by horizontal (blue arrow) or vertical (magenta arrow) moving object, (c) modified FOBs distribution (blue blocks)

In point of feature-block size for vibration-vector estimation, the large-size block requires more processing time. Besides, it is intractable to obtain an optimal vibration vector when using a large-size block, because the FB is frequently intruded by the moving-object in the general situations. When a moving-object is intruding the FB, the calculation for the vibration-vector is varying at any time. On the other hand, the small-size FB needs fewer computations for obtaining a vibration vector but this vector may be erroneous, because a small-size FB usually contains unreliable features. In the proposed method, the size of each FOB is defined as $k * k$ pixels and nine FOBs are fixed at some specific locations distributed over one frame, as shown in Figure 2(a), where the outermost layer L1 denotes the missing area, L2 denotes the FOB-prohibited area and L3 denotes the observation area containing these nine FOBs. For the missing area, the depth of layer L1 means the tolerable amplitude of vibration that the proposed method can stabilize the vibrating frame at most. For the FOB-prohibited area, it means FOB cannot be set on this area since such an area is left for the allowable searching scope of vibration vector derived by block motion estimation later. Nevertheless, the searching of vibration vector is seriously interfered if a horizontal or vertical moving object is directly intruding FOB, as shown in Figure 2(b), where the blue arrow denotes the horizontal moving direction and the magenta arrow denotes the vertical moving direction. For coping with this intrusion problem, the distribution of those nine FOBs is modified as in Figure 2(c), where the red block denotes the original position of FOB and the blue block denotes the modified position of FOB. For example, a moving object which is vertically crossing FOB-1 will not straightway interfer FOB-4, and another moving object which is horizontally crossing the FOB-7 will also not straightway interfere FOB-8. For the proposed distribution rule of FOBs, the basic strategy is that those FOBs are not in line of horizontal or vertical direction, i.e., each FOB can't have the adjacent neighbors each other. By such a distribution rule, it may generate several allowable distributions of FOBs. For each allowable distribution of FOBs, it can substantially reduce the influence caused by a moving object in the horizontal or vertical moving direction mentioned above.

From those nine FOBs, the optimal one is selected to work as the FB and then is used to obtain the vibration vector through block motion estimation. To estimate a correct vibration vector, the selected optimal FOB requires some significant features to correctly discriminate its original position from the vibrated position for the purpose of calculating its shift distance. If the selected FOB has no certain feature, the motion estimation will easily find the alias block, i.e., false match, which will result in generating an erroneous vibration vector. In addition, if the selected FOB contains moving-object, it is intractable to achieve the optimal block matching and thus it will also derive an incorrect vibration vector. From the above description, an FOB without certain features or with moving-object intrusion should be excluded for avoiding an erroneous vibration vector obtained. Therefore, the selected FOB needs to meet two conditions: containing certain features and no moving-object intrusion, in order to become the FB.

In the proposed algorithm of Figure 1, the histogram of gray-levels is analyzed to judge if there is the certain feature existed in an FOB. Basically, a histogram can give an estimate of the appearance probability of someone pixel in an image [1,2]. Thus, the histogram of intensity (i.e., gray-level) can be exploited to discriminate one block from another. To obtain better discriminative information, the intensity histogram is quantized into many slices. At first, the range of intensity (from 0 to 255) is divided by a slicing factor k into $256/k$ slices, where a slice means a k-graylevel interval of intensity. The FB candidate is decided by the analysis of the histogram shape of an FOB. Based on this approach, if the histogram of an FOB is like impulse-modal, the content of the FOB is regarded as monotonous. Otherwise, the content of an FOB is regarded as various. In point of fact, a monotonous block usually generates a false estimation which will lead to an erroneous vibrant vector. Hence, such a monotonous FOB is unsuitable to become an FB candidate. The above deduction reveals that only the FOB with various contents can be viewed as an FB candidate. For numerical representation, an FOB is regarded as no-feature block if the pixel number of the highest slice (i.e., the slice with the most pixels) within the histogram of the FOB is larger than $FR\%$ of total pixels of the FOB, where $FR\%$ typically ranges from 50% to 80%; otherwise, an FOB is viewed as one of FB candidates. The decision rule is described in the following equation.

$$\begin{array}{ll} \text{if} & SP > TP_{\text{FOB}} * FR\% \\ \text{then} & \text{no-feature block} \\ \text{else} & \text{feature-block candidate} \end{array} \tag{3}$$

where TP_{FOB} is the number of total pixels in an FOB, SP denotes the pixel number of the highest slice within the FOB. Figure 3 shows histograms of gray-levels of two different FOBs, where the size of FOB is set as $20 * 20$. In the subfigure (a), FOB-A is located at the sky area and thus its histogram contains only one slice of the top gray-level, as shown in the subfigure (b). For FOB-B, it is located at the outer side of a road and contains a static car, and such various contents generate a histogram including many slices of different gray-levels, as shown in in the subfigure (c). From Equation (3), if $n = 20$ and $FR = 50$, FOB-A is regarded as no-feature block and thus is given up, but FOB-B is selected as an FB candidate.

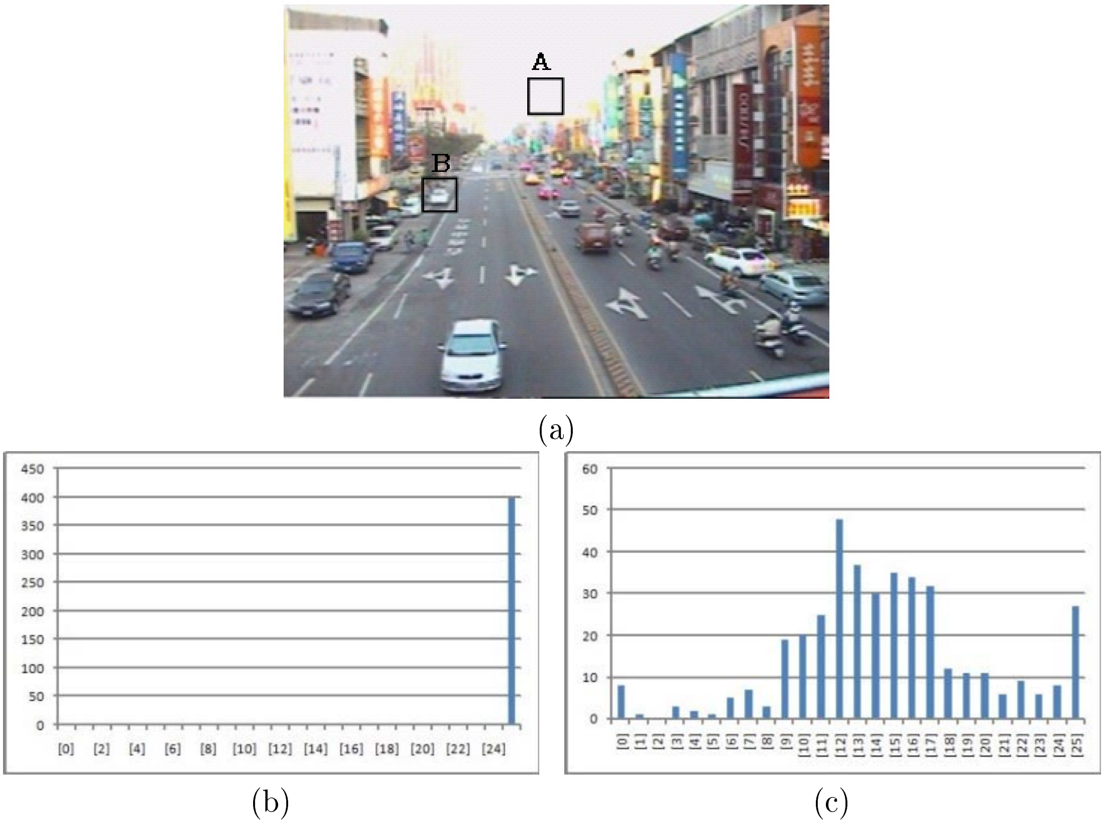
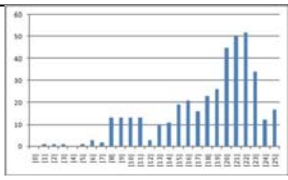
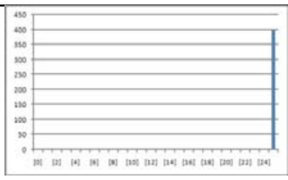
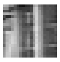
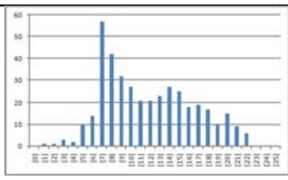
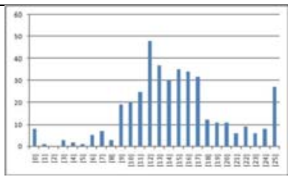
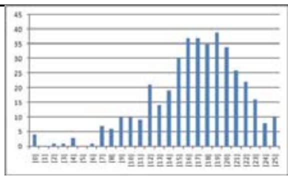
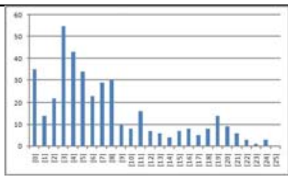
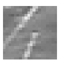
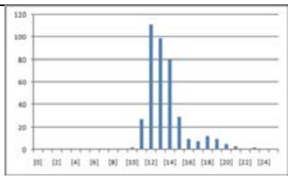
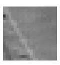
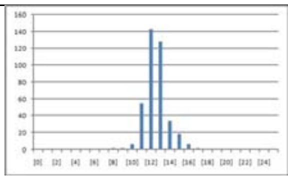
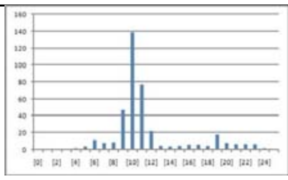


Figure 3. Histogram of gray-levels of different FOBs: (a) two different FOB positions A and B, (b) histogram of gray-levels of FOB-A, (c) histogram of gray-levels of FOB-B

Based on the decision rule of Equation (3), it may generate several FB candidates. Then, the optimal FB is selected from these FB candidates by comparing their histograms of gray-levels. Basically, for the optimal FB, the height (i.e., pixel number) of the highest slice of the histogram is lower than that of the other FB candidates. In other words, the content of an optimal FB is diverse. Therefore, the selection priority for those FB candidates will decrease with the increasing pixel number of the highest slice of the histogram. Table 1 lists the gray-level histograms of nine FOBs for the frame of Figure 3(a) and their *SP*s, where *SP* denotes the pixel number of the highest slice for an FOB. From the table, it can be observed that only FOB-2 has to be deserted because its *SP* is up to maximum value, i.e., 400. By Equation (3), if $FR\%$ is set to 50%, the other eight FOBs are selected as FB candidates. Then, the selecting priority of the optimal FB from these FB candidates is FOB-5 ($SP = 39$) > FOB-4 ($SP = 48$) > FOB-1 ($SP = 52$) > FOB-6 ($SP = 55$) > FOB-3 ($SP = 57$) > FOB-7 ($SP = 111$) > FOB-9 ($SP = 139$) > FOB-8 ($SP = 143$). Besides, to become an FB used to calculate the vibration vector, such a block can't contain moving-object, as described in the following.

Due to temporal correlation of pixels in natural videos, there is only trivial difference between two adjacent frames without involving moving-object, but the difference will become more significant when these frames contain the moving-object. Hence, this difference can be utilized to check if there is moving-object intruding an FOB. The average

TABLE 1. Gray-level histograms of nine FOBs for the frame of Figure 3(a) and their SPs, where FOB size is $20 * 20$

FOB-i	Content	Histogram	SP	Moving-object intruding
1			52	Yes
2			400	No
3			57	No
4			48	Yes
5			39	Yes
6			55	No
7			111	No
8			143	No
9			139	No

difference between two FOBs located on the same position but in the adjacent frames is defined as $Diff_{AVG}$ in the following equation.

$$Diff_{AVG} = \frac{1}{k * k} \sum_{m=1}^{k} \sum_{n=1}^{k} |B_t(m,n) - B_{t-1}(m,n)| \tag{4}$$

where $k * k$ is the size of FOB and $B_t(m,n)$ and $B_{t-1}(m,n)$ denote the values of pixels (m,n) in an FOB at frames t and $t-1$, respectively. If $Diff_{AVG}$ is larger than the intrusion threshold MT, it is considered that there is a moving-object which is intruding the FOB, as described in the following equation.

$$\begin{array}{ll} \text{if} & Diff_{AVG} > MT \\ \text{then} & \text{moving-object intruding} \\ \text{else} & \text{without moving-object intruding} \end{array} \tag{5}$$

In the above equation, MT is determined by the statistical data of experiments, though it will vary with the content of FOB. It is noted that Equation (5) may not assure if there is any moving-object intruding an FOB, but it substantially implies that the content of that FOB is changeable. Anyway, a content-changeable FOB is not suitable for becoming the FB in practice.

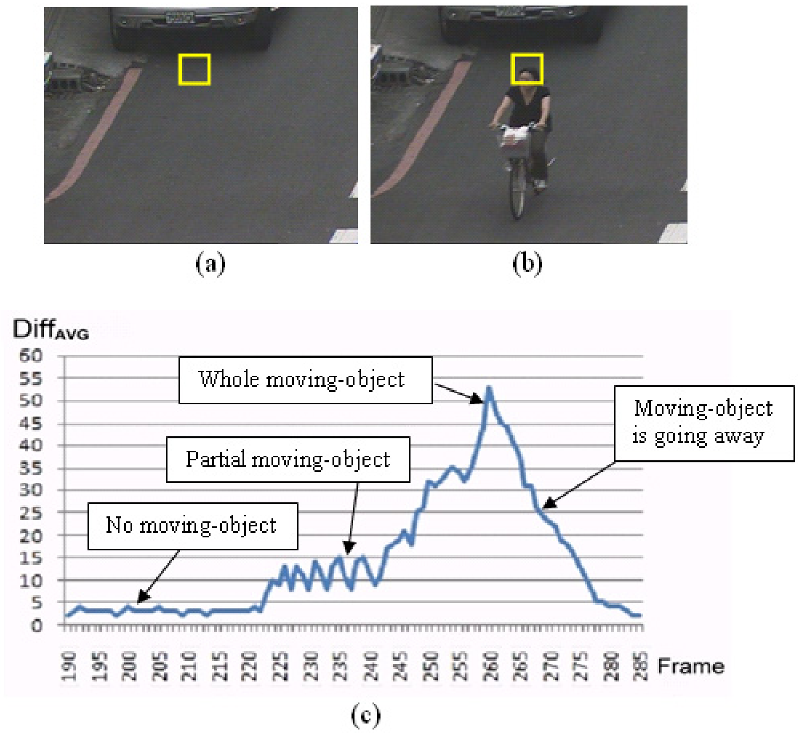


FIGURE 4. Affection in $Diff_{AVG}$ of FOB (yellow square) by the moving-object intruding. (a) Frame#200: without any moving-object, (b) Frame#265: a bicycle is intruding FOB, (c) $Diff_{AVG}$ distribution during the period that a bicycle is moving across FOB.

Figure 4 describes the affection in $Diff_{AVG}$ by the moving-object intruding. The subfigure (a) shows that Frame#200 contains no moving-object and the subfigure (b) shows that the moving-object (a bicycle) is intruding an FOB (yellow square), as shown in Frame#265. The subfigure (c) depicts $Diff_{AVG}$ distribution during the period that a bicycle is moving across FOB. It can be observed that the value of $Diff_{AVG}$ is about 5 if there is no any moving-object existed in FOB. When a moving-object is intruding FOB, i.e., the FOB contains only a partial moving-object, the value of $Diff_{AVG}$ will gradually increase to 10 ∼ 15 (at Frame#225 ∼ #240) and then increase up to 53 (at Frame#261) if the moving-object covers FOB completely. Finally, $Diff_{AVG}$ will gradually decrease when the moving-object is leaving from FOB.

From the description mentioned above, the proposed method will firstly adopt an FB candidate with the highest priority and then check if the moving-object intrusion occurs. If the selected FB candidate has no moving-object intruding problem, it is utilized to calculate the vibration vector. Otherwise, this FB candidate is deserted and then the next FB candidate with the second higher priority is adopted in turn. By such way, the optimal FB candidate with a higher priority and without moving-object intruding is derived. In Table 1, FOB-6 is the optimal FB, because that only FOB-3, FOB-6, FOB-7, FOB-8 and FOB-9 are not intruded by the moving-object, and FOB-6 has a higher priority in feature-block than the other four FOBs. On the other hand, though FOB-1, FOB-4 and FOB-5 are higher than FOB-6 in respect to feature-block priority, these three FOBs have to be deserted because they all have the moving-object intruding problem.

2.3. **Calculation of vibration vector.** In physics, vibration is commonly an oscillatory motion, i.e., a movement first in one direction and then back again in the opposite direction. A simple mechanical vibration is a continuous periodic change in a displacement with respect to a fixed reference and can be described by two factors: its amplitude and its frequency. Basically, frame vibration can be viewed as a periodic back-and-forth motion of its contents since it is a result of the displacement of a camera from an equilibrium condition, followed by the camera's response to the forces that tend to restore equilibrium. Therefore, it is necessary to obtain both direction and amplitude of motion of the frame content for restoring the content of a frame. Such motion involving direction and amplitude can be represented by a vector, called "vibration vector". As a result, the block-matching algorithms (BMA) [15] can be introduced to estimate the vibration vector, i.e., motion vector, of the frame content and then the derived vibration vector is utilized for achieving frame repair.

In the proposed method, FB is viewed as a block that is employed to estimate the motion vector by the full-search algorithm. Then, the estimated motion vector will represent the vibration vector of the whole frame for frame repair in the later processing. In the block-matching process, motion of FB of pixels, $k * k$, within both regions of L2 and L3 of Figure 2(a), is estimated, and it assumes all pixels within the block of size $k * k$ have the same uniform motion. Basically, it is necessary to find the best match between the $(k * k)$ block in the current frame and a corresponding block in the previous frame within a search window. Under the limitation of the search window, the range of the vibration vector is constrained by $[\pm\varepsilon, \pm\varepsilon]$, as shown in Figure 5. The best block match is based on the minimum block distortion measure (BDM). The distortion between the block in the current frame and the shifted block in the previous frame within the search window is defined as MAD (mean of absolute difference), as described in the following equation.

$$\text{MAD} = \frac{1}{k * k} \sum_{m=1}^{k} \sum_{n=1}^{k} \left| P_{m,n} - P'_{m,n} \right| \tag{6}$$

where $P_{m,n}$ and $P'_{m,n}$ denote pixel values of pixels within the original block and the shifted block, respectively. The search fashion of motion estimation is full search (FS) in which every pixel (or point) within the search window has to be checked. In other words, each point's MAD needs to be computed and compared so as to find the location corresponding to the minimum BDM point. Finally, if one point (a, b) has the least BDM, the best match is achieved and thus (a, b) denotes the vibrant vector.

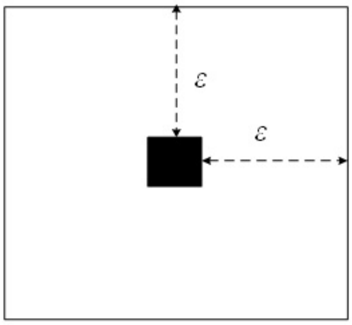


FIGURE 5. The range of the vibration vector within a search window

2.4. **Frame repair.** In the proposed frame-repair scheme, the vibration vector obtained in the above subsection is used to compensate the shift of each pixel of the current fame in order to generate stable frames. Theoretically, the partial boundary areas will miss when the frames are vibrating. This is because there are no allowable pixels which can correctly compensate such areas. So, four sides of boundary areas are regarded as the missing areas and thus the compensated area is limited in the central part, i.e., display area, as shown in Figure 6. In the figure, the original point of display area in a whole frame is set on the upper-left corner, i.e., the coordinate $(0, 0)$. Based on the above frame-repair scheme, the vibrating-area repair is achieved by using compensation with vibration vector (a, b) according to Equation (7).

$$F_{\text{repaired}}(x, y) = F_{\text{vibrant}}(x + a, y + b) \tag{7}$$

In the above equation, F_{vibrant} and F_{repaired} denote the vibrant frame and repaired frame, respectively.

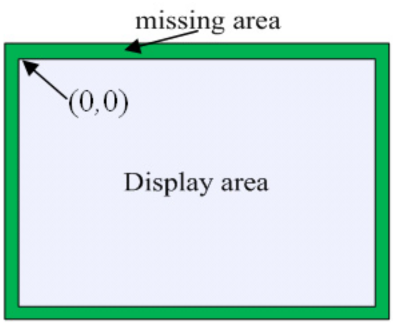


FIGURE 6. Display area (compensated area) and missing area

3. **Simulation and Discussion.** A theoretical analysis about the proposed video stabilization method has been given in the above section, but the implementation of such a stabilization algorithm can provide a realistic and interesting evaluation. For the detailed evaluation in succession, three sequences including Bicycler with vertical vibration, Donmen-1 with horizontal vibration and Donmen-2 with vertical/horizontal vibration, are used in the following simulation. The execution speed of the proposed algorithm can meet the real-time digital video recorder (DVR) system at a format of 30f/s with frame

size of 320 × 240 on Intel Core2 CPU of 2.13 GHz without any hardware acceleration. For the following experiments, FOB size is set to $20 * 20$ and the depth of layer L1 is set to 20 pixels in the proposed algorithm.

In Figure 7, using Bicycler sequence with vertical vibration, subfigures (a) ∼ (d) denote the consecutive vibrating frames and subfigures (e) ∼ (h) denote the repaired frames which are corresponding to subfigures (a) ∼ (d), respectively. From subfigures (a) ∼ (d), it can be clearly observed that the car-head of each frame appears on different positions. Also, there is the same phenomenon for the bicycler's head as indicated by red-line which is used as a reference line. This implies that display of such consecutive frames is unstable. On the other hand, the display of such repaired frames of subfigures (e) ∼ (h) is stable since every car-head and every bicycler's head appear about at the same locations, as indicated by red-line. Besides, the missing area, i.e., L1 layer, is clipped. From these subfigures, it is clear that the vibrant frames are substantially stabilized by the proposed method. Based on the above simulation, Figure 8 plots the distribution of vibration amplitudes, in which the blue curve denotes the vibration amplitude of original vibrant frames and the red curve denotes the vibration amplitude of the repaired frames. On average, the vibration amplitude with about 20 pixels is reduced to 2 pixels or so.

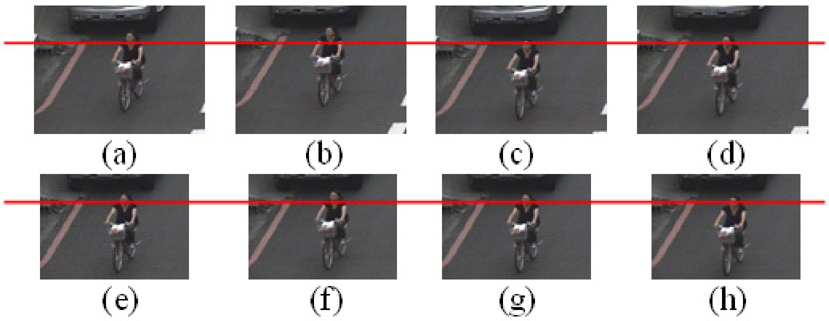


FIGURE 7. Frame-repair result of using Bicycler sequence with vertical vibration: (a) ∼ (d) the original vibrant frames, (e) ∼ (h) the repaired frames by the proposed method

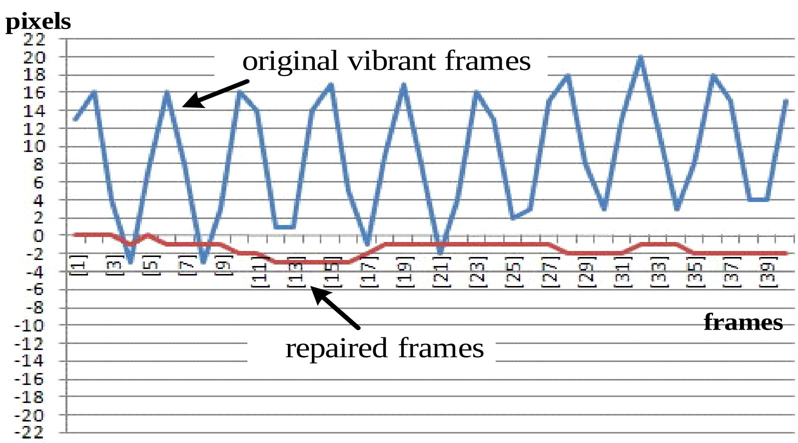


FIGURE 8. Distributions of vibration amplitudes of the original vibrant frames (blue curve) and repaired frames (red curve) using Bicycler sequence with vertical vibration

Figure 9, using Donmen-1 sequence with horizontal vibration, demonstrates the result of the proposed frame-repair method, where subfigures (a) ∼ (d) denote the consecutive vibrant frames and subfigures (e) ∼ (h) denote the repaired frames which are corresponding to subfigures (a) ∼ (d), respectively. Those subfigures (a) ∼ (d) clearly show the unstable appearances from the left billboard (yellow) of each frame. By using the proposed video stabilization method, the vibrant frames are substantially stabilized, as shown in those subfigures (e) ∼ (h). To provide a human-vision perceptual comparison, Figure 10 gives the overlapped appearance of several vibrant frames and the display with repaired frames using Donmen-2 sequence with horizontal/vertical vibration, as described in the subfigure (a) and subfigure (b), respectively. From the comparison of both subfigures, it is apparent that the subfigure (b) looks more distinct than subfigure (a) which has obvious overlapping phenomenon.

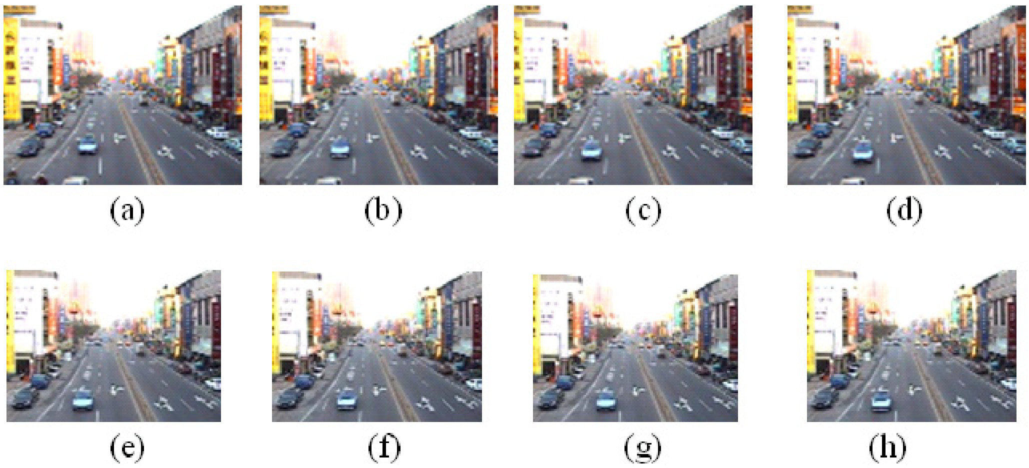


Figure 9. Frame-repair result of using Donmen-1 sequence with horizontal vibration: (a) ∼ (d) the original vibrant frames, (e) ∼ (h) the repaired frames by the proposed method

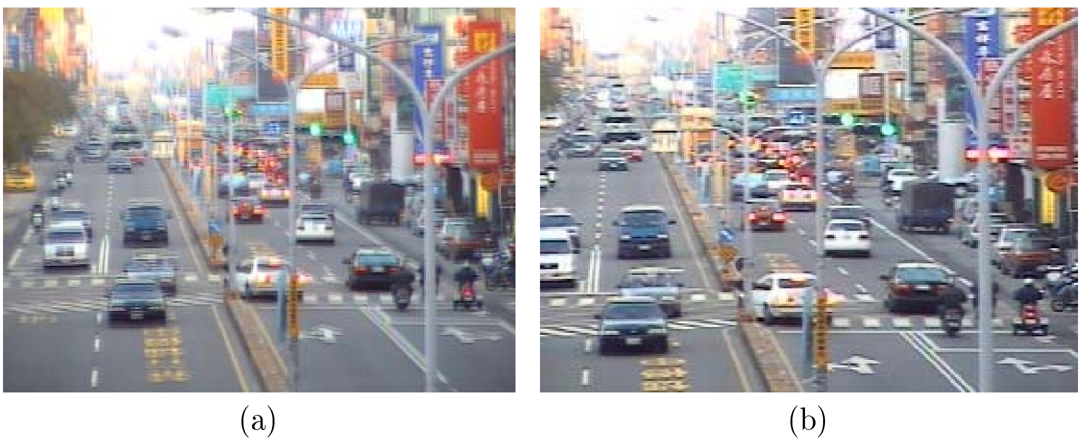


Figure 10. Frame-repair result of using Donmen-2 sequence with horizontal/vertical vibration: (a) overlapped display of several vibrant frames, (b) overlapped display of the frames repaired from (a)

To manifest the superiority of the proposed video stabilization method to four reported methods, a comparison in terms of feature basis, computation time, solution for moving-object interference, disadvantages and application, is made in Table 2. In point of processing time of video stabilization, the speed performances of [10,14] and the proposed method are better than those of both [6] and [7]. Besides, the applications of both [6] and [7] will be restricted to a camcorder mounted on a moving vehicle because they use specific feature bases, such as lane lines and sky line. The feature basis used in [10] is phase correlation and it claimed that the video stabilization will be used in the hand-held video camera. Both [14] and the proposed method are dedicated to the surveillance video camera and they obtain the feature basis by analyzing the histogram of the block. However, the feature block in the proposed method is selected from nine blocks distributed over an image frame and hence this strategy can avoid being interfered with a moving object. In general, if the feature basis is interfered with a moving object, the vibration vector is affected and thus it will result in erroneous motion compensation. The other reported methods of [6,7,10,14] had not mentioned the solution for the problem of the moving-object interference, while the proposed method can avoid the affection of the moving-object interference. With regard to disadvantages of the above five methods, video stabilization of both [6] and [7] will fail if there are not or unclear lane lines and sky line in the situation. For the other three methods, they have the same disadvantage that there may be the stabilization error, because the vibration vector of the current frame is calculated from the previous frame which may not be the original (or stable) frame, i.e., not vibrated frame.

Table 2. Comparison of five video stabilization methods

	[6]	**[7]**	**[10]**	**[14]**	**The proposed method**
Feature basis	lane lines and the road vanishing point	sky line	phase correlation	histogram analysis of one block	histogram analysis of nine blocks
Computation time	medium	more	small	small	small
Solution for moving-object interference	No	No	No	No	Yes
Disadvantages	limited usage	limited usage	maybe has small stabilization error	maybe has small stabilization error	maybe has small stabilization error
Application	only for a camcorder mounted on a moving vehicle	only for a camcorder mounted on a moving vehicle	best for hand-held video camera (but not limited)	best for surveillance video camera (but not limited)	best for surveillance video camera (but not limited)

4. **Conclusions.** This paper proposes a real-time video stabilization method by motion compensation. The basic strategy is to select the feature block from nine fixed observation blocks for calculating the vibration vector which is employed to compensate the vibrating area. In particular, the conventional problem of the moving-object interference can be solved. Experimental results manifest that the proposed algorithm can substantially stabilize the vibrating frames in real-time for the moderate vibration situation (i.e., small than 20 pixels). In addition, the proposed method has lower computation time than

other methods and can be applied to most situations, and thus it is more attractive for applications in the real-time video stabilization system.

Acknowledgment. This work is partially supported by Huper Labortories Co. and National Science Council under Grant NSC97-2622-E-151-011-CC2, Taiwan.

REFERENCES

[1] T.-Y. Chen, T.-H. (C.-H.) Chen and D.-J. Wang, A cost-effective people-counter for passing through a gate based on image processing, *International Journal of Innovative Computing, Information and Control*, vol.5, no.3, pp.785-800, 2009.
[2] T.-Y. Chen, T.-H. (C.-H.) Chen, D.-J. Wang and Y.-C. Chiou, Real-time video object segmentation algorithm based on change detection and background updating, *International Journal of Innovative Computing, Information and Control*, vol.5, no.7, pp.1797-1810, 2009.
[3] H. Wang, J. Liang and C.-C. J. Kuo, Overview of robust video streaming with network coding, *Journal of Information Hiding and Multimedia Signal Processing*, vol.1, no.1, pp.36-50, 2010.
[4] G. Cao, Y. Zhao and R. Ni, Edge-based blur metric for tamper detection, *Journal of Information Hiding and Multimedia Signal Processing*, vol.1, no.1, pp.20-27, 2010.
[5] M. Hansen, P. Anandan, K. Dana, G. van der Wal and P. Burt, Real-time scene stabilization and mosaic construction, *IEEE Conf. on Computer Vision and Pattern Recognition*, pp.54-62, 1994.
[6] Y. S. Yao and R. Chellappa, Electronic stabilization and feature tracking in long image sequences, *Univ. Maryland, Tech. Rep. CAR-TR-790*, 1995.
[7] Z. Duric and A. Rosenfeld, Stabilization of image sequences, *Univ. Maryland, Tech. Rep. CAR-TR-778*, 1995.
[8] C. Morimoto and R. Chellappa, Automatic digital image stabilization, *IEEE Conf. on Pattern Recognition*, 1996.
[9] A. Litvin, J. Konrad and W. C. Karl, Probabilistic video stabilization using Kalman filtering and mosaicking, *Proc. of IS & T/SPIE Symposium in Electronic Imaging, Image and Video Communications and Processing*, vol.5022, pp.663-674, 2003.
[10] S. Etrurk, Digital image stabilization with sub-image phase correlation based global motion estimation, *IEEE Trans. Consumer Electronics*, vol.49, no.4, pp.1320-1325, 2003.
[11] Y. Wexler, E. Shechtman and M. Irani, Space-time video completion, *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, vol.1, pp.120-127, 2004.
[12] Y. Matsushita, E. Ofek, X. Tang and H. Y. Shum, Full-frame video stabilization, *IEEE Conf. on Computer Vision and Pattern Recognition*, vol.1, pp.50-57, 2005.
[13] H. C. Chang, S. H. Lai and K. R. Lu, A robust real-time video stabilization algorithm, *Journal of Visual Communication and Image Representation*, vol.17, no.3, pp.659-673, 2006.
[14] C. H. Chen, C. W. Hsu and C. H. Chen, Video stabilization by feature-block motion estimation, *Proc. of the 8th IEEE International Conference on Intelligent Systems Design and Applications*, Kaohsiung, Taiwan, pp.557-560, 2008.
[15] C. H. Chen, T. Y. Chen, D. J. Wang and Y. F. Li, Multipath flatted-hexagon search for block motion estimation, *Journal of Information Hiding and Multimedia Signal Processing*, vol.1, no.2, pp.110-131, 2010.