http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (22): 6655-6662, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.6655.6662
© 2013 Asian Network for Scientific Information

Multi-node Cooperative Image Mosaicking Algorithm for Wireless Multimedia Sensor Networks

Xiong Zhe-Yuan, Xiong Xiao-Min, Wu Sheng-Yi, Zhang Liang
Department of Public Security Technology, Jiangxi Police College, Nanchang, Jiangxi 330100, China

Abstract: Panoramic images are important for video surveillance of Wireless Multimedia Sensor Networks (WMSN). The traditional image mosaicking approaches always accompany lots of energy and bandwidth consumption, which are not suitable for resource-constrained WMSN. For sufficient utilization of constraint resources on generating high-resolution images with wide field of view, a cooperative image mosaicking algorithm among multiple video nodes is proposed. Image block searching algorithm is applied for image registration to decrease the energy consumption, sum of absolute difference algorithm is adapted to improve the accuracy of image registration and weighted mean algorithm is applied for image stitching. Data volumes of transmission are decreased after image mosaicking, which can efficiently reduce the network loading. Simulation results demonstrate that the computation complexity of proposed algorithm is lower than other image mosaicking algorithms, under certain image registration accuracy and image quality.

Key words: Wireless multimedia sensor networks, image mosaicking, image block searching, weighted mean fusion

INTRODUCTION

Wireless Multimedia Sensor Networks (WMSN) consist of wireless sensors embedded cameras and/or microphones, improving applications like video surveillance, wildlife observation, traffic avoidance and remote monitoring. The major task of WMSN is image communication, which is really a challenge. Wireless video sensor is equipped with low end camera. The video captured by wireless node is with low image resolution and small field of view. However, high coverage, high reliability, low power consumption, high functionality is necessary for achieving higher performance in the monitoring application. It is difficult to meet the requirements of video surveillance applications in resource-constrained WMSN (Xiong *et al*., 2013a).

Recently, area coverage enhancement, image resolution improvement, viewing angle enlarging is the hotspots for WMSN. Due to the small field of view, more video nodes are required to completely cover the monitoring area, which will increase the cost of deployment (Xiong *et al*., 2013b). For video surveillance applications in WMSN, image fusion and new image coding method are used to transform low-resolution images into high-resolution images.

Tezcan and Wang (2008) study the problem of self-orientation in wireless multimedia sensor networks, which is finding the most beneficial pose of multimedia sensors to maximize multimedia coverage with occlusion-free viewpoints. Nevertheless, these methods are flawed. First, these methods are based on single topology. The multilayer topology of WMSN can have a tradeoff in power consumption, coverage, functionality and reliability. Second, the method using high-power nodes equipped rotatable camera platform higher power consumption. Third, the video nodes steering is operated by mechanical equipment, which is not adapted for real-time monitoring applications. In addition, there is a lack of guarantee that a node and its neighbor nodes covering the same area, which will reduce the reliability of WMSN. If a node fails, some areas may not be covered (Hadi *et al*., 2009).

Kansal *et al*. (2006) uses controlled motion to provide virtual high-resolution in WMSN. While the event of interest is detected, the video sensors providing a better quality image by a new motion control. This method is not adapted for real-time applications, but a small range of motion cannot achieve the requirements of a high resolution image capturing. In addition, sensor motions will consume more energy. Soro and Heinzelman (2005) proposed an application-oriented routing algorithm, reducing the overlap area of the video nodes, optimizing coverage scheme. Due to the single-layer topology, forwarding large amounts of data will reduce the network

Corresponding Author: Xiong Zhe-Yuan, Department of Public Security Technology, Jiangxi Police College, Nanchang, Jiangxi 330100, China

lifetime. In addition, the processing capabilities of video nodes are too low to filter out unwanted information, so that all raw data are transmitted through the network. Balasubramanian *et al*. (2008) proposed an image registration algorithm for two low resolution images in video sensor networks, but it did not give network topology and image communication standard. However, energy consumption and the optimal coverage of pixels is optimized depending on the communication scheme between network topology, node, node selection algorithm and video technology. Scholars did not fully consider the characteristics of resource-constrained in WMSN. Proposed algorithm is of higher computational complexity and power consumption for good image mosaic effect, which cannot meet the application requirements for WMSN.

Monitoring task in WMSN is usually done by multiple wireless video nodes. There is overlap between the images that is collected by neighboring video nodes. Image mosaicking of two or more overlapping area can generate high-resolution images with wide field of view, to meet the requirements of video surveillance. It is also possible to effectively reduce image transmission, thereby extending the network lifetime. In addition, traditional image mosaicking approaches must completely receive the desired image before image mosaicking, which may cause a noticeable delay before an entire image is received by nodes. Moreover, traditional image mosaicking approaches accompany lots of arithmetic operations. In order to obtain high-resolution panoramic images with constrained resources, a low complexity image mosaicking algorithm for WMSN is proposed in this study.

SYSTEM MODEL

High coverage, high reliability, low power consumption and high functionality are the main requirements for video surveillance in WMSN include. Most of single-layer sensor networks can meet one of requirements. Multi-layer topology can achieve a better balance between power consumption, functionality, reliability and coverage. In order to meet performance requirements of video surveillance in WMSN, two-tier architecture is used.

The network is divided into several clusters, each cluster including some video nodes and a cluster head, scattered in the monitoring area. The cluster head is not equipped with a camera, but with strong computation capability, storage capacity and communication unit, which is responsible for managing the video nodes, information gathering and processing and forwarding processing result to the base station. Video node consists of a low-power image sensor equipped an embedded sensor platform. Due to higher processing capabilities of upper node, low-resolution images can be mosaicked into a high-resolution image with wide field of view. Three assumptions are used in the model:

- Video nodes are appropriately deployed in the monitoring area to ensure full connectivity, distance from each node to neighbor nodes is within the communications radius, all nodes are immovable and the power is limited
- Each video node obtains their position through a positioning algorithm and gets the ID and residual energy of its neighbor nodes
- The area monitored by the video nodes does not contain objects that could partially or fully occlude the view of image sensors

VIDEO SENSOR SELECTION

Since all video nodes in WMSN are stationary after deployment, coverage for the monitoring area is also fixed. When monitoring area is completely covered, images with two of more overlapping areas, which are obtained from observed region, can be mosaicked to a panoramic image in an arbitrary viewpoint. In order to provide a mosaic image with desired viewpoint for users and conserve the energy of wireless nodes, a selection of video sensors whose fields of view overlap with the desired field of view should be considered. A video sensor selection method is explored to select which nodes are desired to mosaic image for avoiding redundant data and conserving nodes' energy.

If the video sensors are not constrained by limited energy, the preferable way to select sensors that jointly provide the user's desired image is by choosing those video sensors that contain parts of the scene a user is interested in and that have the smallest angle between their directions and the user's direction. However, since video sensors are battery-operated, energy conservation should be considered as well in the sensor selection method.

As the angle between the directions of a selected video sensor and the user's desired view becomes larger, it is expected that the difference in the image obtained by this sensor and the desired user's image is larger. The angle between each sensor's direction and the user's direction increases with the distance of the sensor to the user. As the angle between directions of the selected node and viewing directions requested by the user is increased, difference between the captured image and the desired image also increases, as shown in Fig. 1.

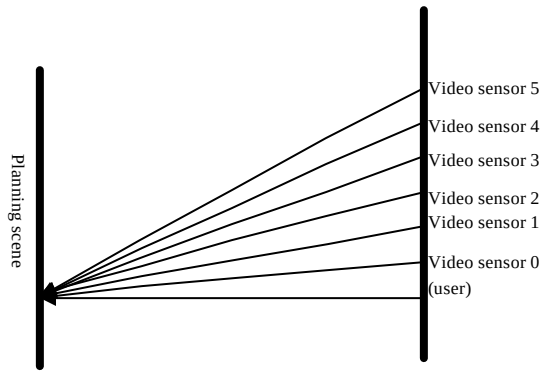

Fig. 1: Video surveillance with multiple nodes

As the sensors' visible fields are overlapped, the volume of one sensor can be partially or fully contained in the visible field of other sensors. In the absence of objects, the scene viewed by a sensor may be recovered from the images taken by neighbor sensors with overlapping views. For a given position and direction of the user's desired view, there is a group of video nodes that can provide images of the scene in response to the user's query. We label this group of cameras as a set of Candidate Cameras (CC). However, to prevent the selection of the sensors that provide very different images from the image expected by the user, we exclude from the set CC every camera for which the angle between its optical axis and the user's directional view is larger than some threshold angle.

Based on these ideas, minimum angle selection approach is introduced for the selection of sensors. The cameras are chosen by minimizing the angle between the sensor's axis and the user's view direction. This method is straightforward and will minimize the distortion between the reconstructed image and the desired image.

IMAGE MOSAICKING

Image mosaicking procedure generally involves image preprocessing, image registration and image stitching. Since the image denoising and image enhancement are computation complexity, which is a challenge for resources-constrained node, we skip image preprocessing procedure. Further, although changes in the light conditions, noise in the process of image acquisition, the image scale changing, rotation, occlusion will seriously affect the image matching result, but because of the use of homogeneous video nodes, imaging conditions such as the lighting conditions, imaging system performance is substantially the same atmospheric conditions, so gray value deviation correction is not needed.

The proposed image mosaicking algorithm steps are as follows:

- Images captured by the video nodes are encoded with JPEG standard and then be transferred to the cluster head node
- Find the two most similar image sub-block image alignment for image registration
- Obtain the result of image mosaic after image stitching

Image registration: Target images of image registration are obtained from different viewpoints in the same field, which reflects a part of the same target area, to be registered for enlarging the field of view. Image registration generates significant overhead because a large amount of pixels must be matched. Many advanced algorithms for image registration have been developed for the sake of improving either accuracy or efficiency. Image registration algorithms are generally divided into two categories: Patch-based methods and feature-based methods. For video surveillance applications studied here, video sensor nodes are homogeneous with the same resolution of the captured images and the nodes are deployed in the same horizontal plane, images are usually captured by translational or rotational movement with overlapping areas. In addition, taking into account the resource-constrained characteristics for WMSN, patch-based method, which is of less computation complexity, is adapted for image registration.

Similarity metrics are responsible for assessing intensity differences between all pixel blocks in every pair of images. A great deal of computation is generated because the calculation happens pixel by pixel. An ideal similarity metric should be able to precisely represent the similarity-degree of each block-pair using a low computational cost. Sum of Squared Difference (SSD), Sum of Absolute Difference (SAD) and Normalized Cross-Correlation (NCC) are the prevailing criteria for successful matching of each block-pair.

SSD is widely used to compute the deviation between the arbitrary data and the mean of a set of data. Each pixel in each block-pair is compared based on pixel intensity. Given an image k+1 and its reference image k and the size of the pre-defined block, m×n, I(i, j) is the intensity of a pixel at position (I, j). SSD function is as follows:

$$SSD_k(i,j) = \sum_{i=1}^{m}\sum_{j=1}^{n}(I_{k+1}(i,j) - I_k(i,j))^2 \quad (1)$$

SAD was devised to reduce the complexity of the SSD function, since the computational overhead and memory consumption of multiplication are considerably high. As shown in the follows, the squared function is replaced by the absolute function, making the equation more effective. Thus, the computational load is remarkably reduced. Like SSD, the block-pair with the minimum SAD means that the blocks have significant commonalities:

$$SAD_k(i,j) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left|I_{k+1}(i,j) - I_k(i,j)\right| \quad (2)$$

Due to its simplicity and efficiency, SAD is a proper similarity metric for image registration so as to improve the performance of a surveillance application in WMSN.

The NCC is derived from Cross-Correlation (CC), which is simplified by the expansion of SSD function. Unlike with SSD and SAD, the most similar block-pair is the pair with the greatest CC value. However, CC cannot avoid intensity variation errors caused by exposure differences when bright patches exist in either image. In general, the NCC is suitable for similarity measurement in some advanced applications. Nevertheless, the NCC function has a higher complexity than the functions of SSD and SAD. A suitable and carefully-chosen metric is the essence for correctly evaluating the similarity of two images.

Unlike patch-based methods, which seek common segments by comparing the intensity deviation of pixels on block basis, feature-based algorithms explore the prominent features of each image in the collection, create a relationship based on detected common features in each image-pair and assess the geometric transformation between the two images. Therefore, precise feature detection and adequately corresponding feature matching are vital to feature-based alignment algorithms.

Because a patch-based algorithm uses block matching to compare similarities and only considers the pixel's intensity as the similarity criterion, patch-based algorithms are desirable when the motions of source images are translation or rotation. Source images with multiple resolutions or other complex motions registered by patch-based algorithms can introduce very high computational overhead. On the contrary, features are defined by many other criteria besides intensity and are invariant to scaling or transformation. Thus, feature-based algorithms are preferable to register source images that have multi-resolutions and complex transformations. Table 1 summarizes the above discussion.

An image registration algorithm based on image block search is proposed in this section. An improved algorithm based on sum of absolute difference is employed to search the block of the highest similarity, for image registration. Firstly, the block size is determined and then searching the image blocks.

Table 1: Comparison between patch-based and feature-based image registration algorithms

Criteria	Patch-based	Feature-based
Compared basis	Block	Feature points, neighbors
Compared criteria	Intensity	Orientation, texture, curvature
Complexity	Low	High
Transform of source images	Translation, rotation	Arbitrary transformation
Salient information contained	Grayscale or color	Textures or contents

Block size determination: Patch-based image registration evaluates the degree of similarity on a block. A block must be defined before image registration. The question is how to determine the size of this block. An oversize or undersize block increases the possibility of the image misregistration. A block of a proper size should have sufficient identifying information and minimal useless data. According to Pazzi *et al.* (2010) experiments, misregistration is minimal when the block size is about one percent of the image size (Soro and Heinzelman, 2007). However, JPEG image coding algorithm is employed in this study, the source image is divided into a number of 8×8 blocks for encoding and the image information is mainly concentrated in the low frequency sub-block and therefore the block size of 8 × 8 is more suitable for image registration. The common resolutions of images captured by wireless camera sensors are 352×288 of Common Intermediate Format (CIF).

Image block searching: Once the block size is determined, traditional image mosaicking algorithm search each possible block-pair in the two images. In contrast, image block searching algorithm splits one of the images into blocks of the designated size and then the resulting blocks are matched with all blocks in the other images (Pazzi *et al.*, 2010). The block-pair searching algorithm employed here reduced the computation complexity. When the block size is proper, it can also get a better image registration results.

Although the registration accuracy of SSD algorithm is very high, the computation complexity and memory consumption is too high to be employed for WMSN. Therefore, the more simple and effective SAD method is employed to estimate the similarity between two blocks. Nevertheless, the SAD algorithm needs to be adapted to reduce the probability of misregistration.

The adaptation of SAD involves SAD_{min}, SAD_{self} and Rate. SAD_{min} represents the minimal intensity difference between blocks contained in two images, while SAD_{self} illustrates the intensity difference between each pixel as

compared to the average intensity of the block. Rate denotes the rate of SAD_{self} over SAD_{min}. SAD_{self} and SAD_{min} are computed in accordance with each block of image k generated by image block searching.

In order to obtain SAD_{min} of each block in image k, the intensity differences between each block in image k and all eligible blocks in image k+1 must be collected in advance. The smallest SAD is set to SAD_{min} and its associated block in image k+1 is the likeliest match for the block in image k. The location of the block-pair is saved as well. However, in some cases, more than one pair of blocks has the same SAD which also results in SAD_{min}. In such cases, the first block-pair deemed to be SAD_{min} is considered as the likeliest block twin of image k and k+1. The position of this block-pair will not be replaced even if other block pairs with the same SAD_{min} are detected. However, SAD_{min} cannot be zero for Rate calculation because it will be used as a denominator later on when the greatest Rate calculation is determined. In order to cater to the Rate calculation, it discards block pairs for which SAD_{min} are zero, even though they might actually be alike. The block-pair with the smallest non-zero SAD_{min} could be slightly offset, but it is still reliable, because refinements can later correct any minor misregistration at the two increasing quality levels.

If a pair of blocks has a great SAD_{self} with a small SAD_{min}, this will result in a block-pair with sufficiently distinguishing contents. When a pair of similar blocks has rich and varied information, the mismatch probability of such blocks is lower than that of two blocks that only have a minimum difference. Rate, which divides SAD_{self} by SAD_{min}, is used to identify a pair of rich content blocks with which to achieve precise image registration. Image registration algorithm based on improved SAD and image block search is as follows:

- Image k and image k+1 are split into blocks of 8×8 size
- SAD_{self} of each block in image K is calculated, the block location is obtained
- Calculating the block pair with SAD_{min} in the image k and the image k+1, to obtain the position of the blocks in image k and image k+1 with SAD_{min}
- Initialization, the Rate is set to zero
- If SAD_{min} of the block pair is zero, giving up the pair of blocks
- If SAD_{min} of the block pair is not zero and SAD_{self}/SAD_{min}>Rate

Then, Rate = SAD_{self}/SAD_{min}.

If the Rate value of a block pair is small, the pair of blocks is identified as mistaken matching, giving up the pair of blocks.

Image stitching: The purpose of image stitching is to ensure that the overlap region of images can be continuously and no visible suture. However, due to various noises and vignetting, the role of the disparity, such as radial distortion caused by movement of the optical center, some of the edges of the image after image stitching are still visible. There are a lot of image stitching algorithms, including median filtering algorithm, wavelet transform algorithm, the weighted average algorithm, which can well eliminate the stitching sutures. However, since resource constraint in WMSN, image stitching algorithm requires the computational complexity as low as possible under a certain image quality.

Three typical image stitching algorithm, median filtering algorithm is fast, but the image quality seriously; wavelet transform algorithm is of high computational complexity; weighted average algorithm can improve the signal to noise ratio of the fusion image and also have a good deal with exposure difference, convergence speed is fast enough for real-time processing, a good balance between the computational complexity and image quality can be achieved. Therefore, we employ the weighted average method. The gray level of image pixels in the overlapping area is calculated by the corresponding points in the two images with the gray-weighted averaging. I_1 (x, y) and I_2 (x, y) are the gray values of pixels in the two image to be stitched. The gray value of the overlap region images fused by weighted averaging is described as follows:

$$I(x,y) = d_1 I_1(x,y) + d_2 I_2(x,y) \tag{3}$$

where, d_1 and d_2 are the gradation weights of the overlap region in two images and $d_1+d_2 = 1$, $0 \le d_1, d_2 \le 1$。Depending on the selection method of weighted average value function, the weighted average method is divided into a gradual fade out, hat function method. This study selects gradual fade-out method, which is often used in the practical application. The Euclidean distance from pixels to the center of the image is the basis of an alternative weight. The gradual fade-out is calculated as follows:

$$I(x,y) = dI_1(x,y) + (1-d)I_2(x,y) \tag{4}$$

Where:

$$d = \frac{x_2 - x}{x_2 - x_1}$$

$x_1 \leq x \leq x_2$, d is a gradient coefficient. When d slowly change from one to zero, x_1 and x_2 are the minimum and maximum values of overlap zone axis, the images slowly transit from $I_1(x, y)$ to $I_2(x, y)$.

RESULTS AND DISCUSSION

Surveillance images of outdoor scenes are selected here for simulation experiments in the MATLAB. Two homogeneity video sensors are installed in high place to monitor vehicles and pedestrians on the same intersection to. The resolution of captured grayscale image is 560×420. After JPEG encoding, images were sent to the cluster head node for image stitching. Figure 2 and 3 shows monitor image sequences 0510 and 0800 of node k and k+1 and the mosaic image.

U.S. Crossbow's wireless sensor node platform Imote 2 is an advanced platform, with SRAM of 256 KB, radio transceiver of the transmission rate at 250 kb sec^{-1}. Assuming a cluster head node to manage two video nodes, grayscale image size of 560×420 resolution captured by two video nodes is approximately 100 k. The cluster head node can transmit two video images of each frame within a second. After mosaicking, the image data size is approximately two-thirds of the two images that are not spliced, as shown in Table 2. Two mosaic images can be sent by the cluster head per second, which can effectively save communicate energy consumption and improve the sampling frequency of the video node.

Image registration is the most important step in image mosaicking. It is of the highest computational complexity, since there are a large number of pixels to be matched. Computational complexity of image registration can be seen as the computational complexity of the entire image mosaicking algorithm. In the image registration process, the search for the most similar block pair is a step of generating a great amount of calculation. Exhaustive search algorithms search all possible block pair in the two images. If the image resolution is m×n, then the computational complexity of an exhaustive search is O(m2n2). When the block search is used, an image is divided into mn/64 blocks, then possible block pair may

Table 2: Comparison of original image size and mosaic image size

Image no.	Image k	Image k+1	Mosaic image	Ratio between original image and mosaic image (%)
0510	74.5 k	100 k	120 k	68.8
0630	94.1 k	111 k	141 k	68.7
0720	94.1 k	109 k	141 k	69.4
0800	99.6 k	104 k	121 k	59.4

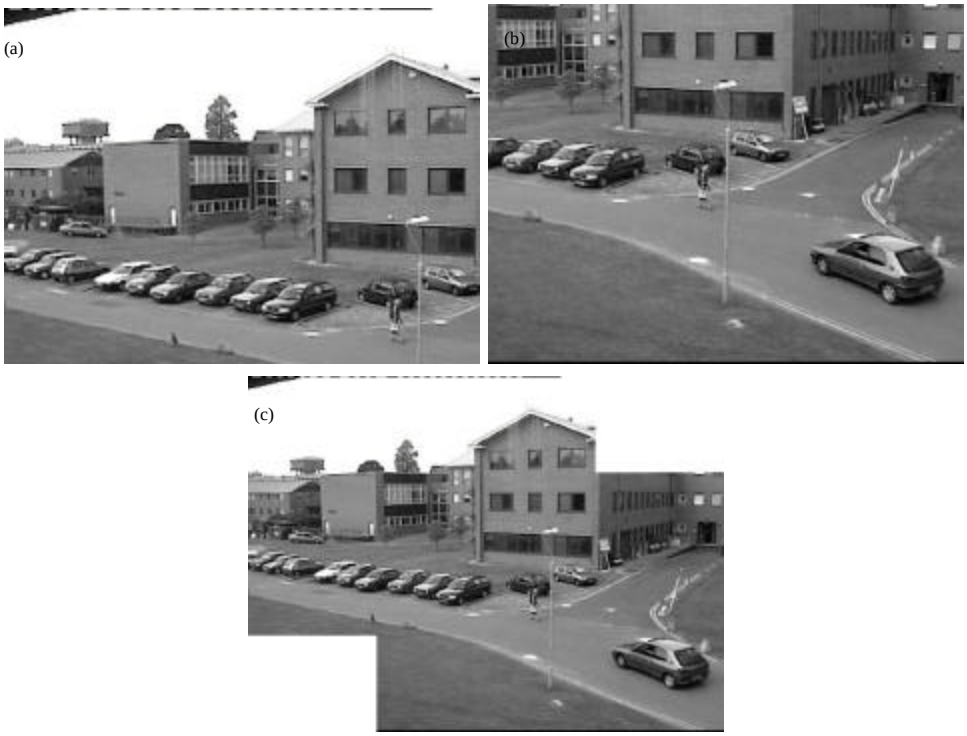

Fig. 2(a-d): Mosaic image of node k and k+1 in picture 0510, (a) Picture 0510 by node k, (b) Picture 0510 by node k+1 and (c) Mosaic image

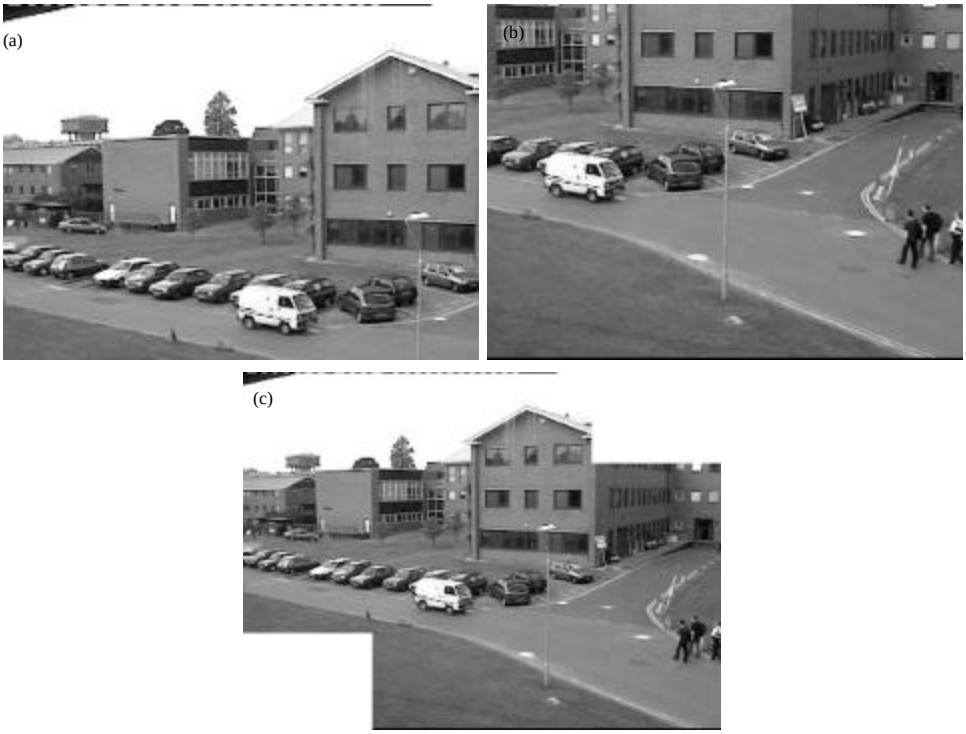

Fig. 3(a-c): Mosaic image king of node k and k+1 in picture 0800, (a) Picture 0800 by node k, (b) Picture 0800 by node k+1 and (c) Mosaic image

be reduced as m2n2/64. The calculation complexity of block search is O(m2n2/64), which is much less than exhaustive search.

CONCLUSION

An image mosaicking algorithm adapted for WMSN is proposed in this study. Image block search algorithm accelerates the speed of image registration, reducing the potential for a number of similar blocks. The SAD algorithm reduces the computational complexity of image registration and the probability of misregistration. Simulation results and analysis show that the proposed algorithm is of low computational complexity, high precision image mosaic. It can effectively improve the performance of video surveillance for WMSN, expand monitoring horizons.

ACKNOWLEDGMENTS

This study was supported by Science and technology project of Jiangxi Department of Education under Grant GJJ12762, natural science foundation of Jiangxi Province under Grant 20132BAB201057, Jiangxi university science and technology project under Grant KJLD13098, KJLD12099.

REFERENCES

Balasubramanian, H., P.K. Mitikiri and K.R. Namuduri, 2008. Image registration in low resolution visual sensor networks. Proceedings of the International Conference on Information Processing in Sensor Networks, April 22-24, 2008, St. Louis, MO, USA., pp: 551-552.

Hadi, S.A., B. Pouya, E.M. Mohsen and A. Maghsoud, 2009. High-resolution images with minimum energy dissipation and maximum field-of-view in camera-based wireless multimedia sensor networks. Sensors, 9: 6385-6410.

Kansal, A., W. Kaiser, G. Pottie, M. Srivastava and G. Sukhatme, 2006. Virtual high-resolution for sensor networks. Proceedings of the 4th International Conference on Embedded Networked Sensor Systems, November 1-3, 2006, Boulder, CO, USA., pp: 43-56.

Pazzi, R.W., A. Boukerche, J. Feng and Y. Huang, 2010. A novel image mosaicking technique for enlarging the field of view of images transmitted over wireless image sensor networks. Mobile Network Appli., 15: 589-606.

Soro, S. and W. Heinzelman, 2007. Camera selection in visual sensor networks. Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, September 5-7, 2007, London, pp: 81-86.

Soro, S. and W.B. Heinzelman, 2005. On the coverage problem in video-based wireless sensor networks. Proceedings of the 2nd International Conference on Broadband Networks, October 7, 2005, Boston, MA, USA., pp: 932-939.

Tezcan, N. and W. Wang, 2008. Self-orienting wireless multimedia sensor networks for occlusion-free viewpoints. Comput. Networks, 52: 2558-2567.

Xiong, Z., X. Fan, S. Liu, Y. Li and Z. Huan, 2013. Performance analysis for DCT-based coded image communication in wireless multimedia sensor networks. Int. J. Smart Sensing Intell. Syst., 6: 120-135.

Xiong, Z.Y., J.S. Zhang, S.Y. Wu, L. Zhang and M. Tu, 2013. A weighted ROI-based image quality assessment for WMSN. J. Applied Sci., 13: 3766-3769.