

## Research article

# SAFER under vehicle inspection through video mosaic building

*Andreas Koschan, David Page, Jin-Choon Ng, Mongi Abidi, David Gorsich and Grant Gerhart*

### The authors

Andreas Koschan, David Page, Jin-Choon Ng and Mongi Abidi are based at Imaging, Robotics and Intelligent Systems Lab, Department of Electrical and Computer Engineering, The University of Tennessee, Knoxville, Tennessee, USA. David Gorsich and Grant Gerhart are based at Tank-Automotive Research, Development and Engineering Center, US Army RDECOM, Warren, Michigan, USA.

### Keywords

Inspection, Road vehicles, State security, Robotics

### Abstract

The current threats to US security, both military and civilian, have led to an increased interest in the development of technologies to safeguard national facilities such as military bases, federal buildings, nuclear power plants, and national laboratories. As a result, the imaging, robotics, and intelligent systems (IRIS) laboratory at the University of Tennessee has established a research consortium, known as security automation and future electromotive robotics (SAFER), to develop, test, and deploy sensing and imaging systems. In this paper, we describe efforts made to build multi-perspective mosaics of infrared and color video data for the purpose of under vehicle inspection. It is desired to create a large, high-resolution mosaic that may be used to quickly visualize the entire scene shot by a camera making a single pass underneath the vehicle. Several constraints are placed on the video data in order to facilitate the assumption that the entire scene in the sequence exists on a single plane. Therefore, a single mosaic is used to represent a single video sequence.

### Electronic access

The Emerald Research Register for this journal is available at

[www.emeraldinsight.com/researchregister](http://www.emeraldinsight.com/researchregister)

The current issue and full text archive of this journal is available at

[www.emeraldinsight.com/0143-991X.htm](http://www.emeraldinsight.com/0143-991X.htm)

Industrial Robot: An International Journal  
Volume 31 · Number 5 · 2004 · pp. 435–442  
© Emerald Group Publishing Limited · ISSN 0143-991X  
DOI 10.1108/01439910410551854

## 1. Introduction

In the present situation, mobile sensor platforms are becoming key elements of safety and surveillance. The imaging, robotics, and intelligent systems (IRIS) laboratory at the University of Tennessee has established a program for security automation and future electromotive robotics (SAFER) to develop, test, and deploy sensing and imaging systems that augment the missions of current and future mobile sensor platforms. In essence, SAFER seeks to deploy “sixth-sense” technologies such as thermal imaging cameras, laser range scanners, and other advanced sensors and to incorporate autonomous intelligence into these sensors through the development of fusion and processing algorithms.

### 1.1 SAFER overview

As shown in Figure 1, the three fundamental elements of a robotics platform are sensing, processing, and mobility. The main focus of the SAFER program is the development of processing algorithms. Currently, a variety of sensors and mobile platforms are available. The link that requires additional research is the processing to bring these elements together. Specific technologies that SAFER has targeted include processing and fusion of 2D video from visual and thermal cameras. For the development of these technologies SAFER promotes the notion of “SFC bricks” to achieve an interchangeable sensor suite. A sensing, fusion, and communications (SFC) brick is a three-module concept. The sensor module contains one – or integrates multiple sensors – to collect data around the robot environment. The fusion module processes this data and incorporates reasoning and analysis. Finally, the communications module transmits this information to appropriate end users. The SFC brick concept allows the user to easily deploy and upgrade the system as new sensor bricks become available.

### 1.2 Targeted mission

To center research efforts, SAFER firstly targets a specific mission for the inspection of vehicle undercarriages. The key design of this system is for the robot and sensors to have a low profile for navigation underneath a vehicle such as a car or truck. Figure 2 shows this mission. This figure shows a prototype platform that has the flexibility

This work is supported by the University Research Program in Robotics under grant DOE-DE-FG02-86NE37968, by the DOD/RDECOM/NAC/ARC Program, R01-1344-18, and by the US Army under grant Army-W56HC2V-04-C-0044.

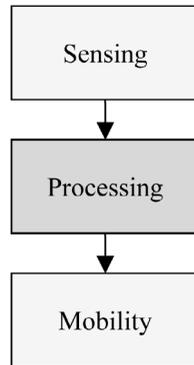
# Report Documentation Page

*Form Approved  
OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>05 OCT 2004</b>	2. REPORT TYPE <b>N/A</b>	3. DATES COVERED <b>-</b>	
4. TITLE AND SUBTITLE <b>SAFER under vehicle inspection through video mosaic building</b>		5a. CONTRACT NUMBER	
		5b. GRANT NUMBER	
		5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) <b>David Gorsich; Grant Gerhart; Andreas Koschan; David Page; Jin-Choon Ng; Mongi Abidi</b>		5d. PROJECT NUMBER	
		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>USA TACOM 6501 E 11 Mile Road Warren, MI 48397-5000</b>		8. PERFORMING ORGANIZATION REPORT NUMBER <b>14328</b>	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) <b>TACOM TARDEC</b>	
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release, distribution unlimited</b>			
13. SUPPLEMENTARY NOTES <b>Pub. in Industrial Robot: An Internationa Journal Volume 31, Number 5, 2004, pp. 435-442</b>			
14. ABSTRACT			
15. SUBJECT TERMS			
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>	<b>SAR</b>
			18. NUMBER OF PAGES <b>8</b>
			19a. NAME OF RESPONSIBLE PERSON

**Figure 1** The fundamental elements of a robotics platform. The SAFER program focuses on the processing component through fusion of multiple sensors



to accommodate a variety of different sensors. The platform is able to navigate under a vehicle steered by a remote operator. The cavity just above the IRIS logo contains a mirror system to allow different sensors to view up and under a vehicle. This prototype is able to maneuver completely under standard cars and trucks.

The SAFER program is using this platform to experiment with different sensor configurations to study the under vehicle inspection problem. Vehicle inspection is traditionally accomplished through security personnel walking around a vehicle with a mirror on the end of a stick. That person is able to view underneath a vehicle with the mirror to identify contraband such as weapons, bombs, or other security threats. However, the challenge is to overcome the problem that the mirror-on-a-stick approach only allows partial coverage under a vehicle, and the mirror is often restricted by ambient lighting conditions such as poor lighting or sunlight glare. The prototype above seeks to overcome these issues by allowing complete coverage with the low-profile robot

and extending beyond visible inspection by using thermal cameras. Additionally, the stand-off capabilities that the prototype offers is an attractive alternative where potential harm to security personnel is possible. A mirror-on-a-stick solution puts personnel in harms way, but the remote wireless links of the prototype in the figures allows the user to remain at a safe distance.

However, there are times when video tends to be a cumbersome format for referencing visual information. As the camera moves through the scene, the inspection personnel involved must be watching the video sequence the entire time it is running or risk missing important details. If the personnel see that something is amiss or are momentarily distracted, they need to rewind the video or remotely move the mobile platform back to center on the area in question.

Suppose that all the visual information in a single video sequence captured by the surveillance camera were somehow represented by a single, large, high-resolution image that encompasses the entire scene. This image would be a mosaic composed of all the individual video frames taken by that single camera. It has been argued that mosaics provide efficient and complete representations of video sequences (Irani *et al.*, 1996; Zheng, 2003). A mosaic representation eases the inspection process by removing the inter-frame redundancies seen in video sequences, since a mosaic represents each spatial point in the sequence only once. This representation of a video sequence shortens the inspection time by allowing inspection personnel to refer disparate spatial points quickly during inspection. This concept is shown in Plate 1.

The motivation for this work stems from the need to create high-resolution images by building mosaics from a series of infrared and color video data acquired for under vehicle inspection. Video

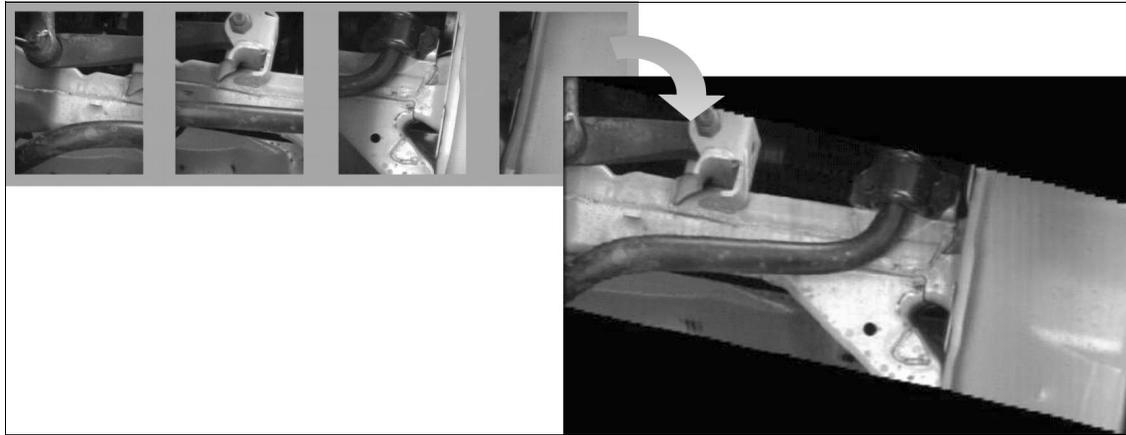
**Figure 2** These images depict a prototype for the SAFER mobile sensor platform: (a) the configurable sensor bay has a view portal just above the IRIS logo in this image and (b) the low-profile platform can navigate remotely under vehicles



(a)



(b)

**Plate 1** Mosaics as concise representations of video sequences

was obtained from a mobile platform moving along the underside of vehicles for the purpose of threat detection, using both standard video as well as infrared modalities. Our aim was to devise a method that is capable of creating very high-resolution images from video sequences from both modalities. We employ multi-perspective mosaic construction paradigms to devise our solution since techniques for multi-perspective mosaic building are well-suited for creating mosaics from image sequences where the camera's optical center moves during data acquisition. Constructing multi-perspective mosaics of infrared and video images is one of the features included in the SAFER vehicle inspection system which is described in more detail in Page *et al.* (2004).

The rest of this paper is organized as follows. Section 2 discusses the method for multi-perspective mosaic construction and the underlying constraints. Experimental results are shown in Section 3, and the paper concludes in Section 4.

## 2. Multi-perspective mosaic construction

To address the under vehicle inspection problem, we propose to align video images by building multi-perspective mosaics. The term “multi-perspective mosaic” originates from the aim to create mosaics from sequences where the optical center of the camera moves; hence, the mosaic is created from camera views taken from multiple perspectives. This is opposed to panoramic mosaic building techniques, which aim to create mosaics traditionally taken from a panning, stationary camera. In other words, panoramic mosaic construction techniques create 360° surround views for stationary locations while the objective of multi-perspective mosaic building is to create very large high-resolution, billboard-like images from

translating camera imagery. The paradigms associated with building multi-perspective mosaics, as described by Peleg and Herman (1997), are straightforward. For a video sequence, the motion exhibited in the sequence must first be determined. Then, strips are sampled from each video frame in the sequence with the shape, width, and orientation of the strip chosen according to the motion in sequence. These strips are then arranged together to form the multi-perspective mosaic.

For instance, for a camera translating sideways past a planar scene that is orthogonal to the principal axis of the camera, the dominant motion visible in the scene would be translational motion in the opposite direction of the camera's movement. A strip sampled from each frame in the sequence must be oriented perpendicular to the motion; therefore, in this case, the strip is vertically oriented. The width of a strip would be determined by the magnitude of the motion detected for the frame associated with that strip.

### 2.1 Constraints

In this work, we have placed certain constraints on the movement of the mobile platform to match the scenario just described. These restrictions greatly simplify the mosaic construction process, and a systematic method of acquiring data of the scene would most likely obey these restrictions. First, it is assumed that the camera is translated solely on a single plane that is parallel to the plane of the scene. It is also assumed that the viewing plane of the camera is parallel to this plane of the scene. Finally, it is assumed that the camera does not rotate about its principal axis.

The collective effect of these constraints is that motion between frames is restricted to pure translational motion. An ideal video sequence would come from a camera moving in a constant direction while the camera's principal axis is kept

orthogonal to the scene of interest. A camera placed on a mobile platform may be used for this purpose. The platform may then be moved in a straight line past the scene. If the scene is larger than the camera's vertical field of view, several straight line passes may be made to ensure that the entire scene is captured. A single pass will produce one mosaic. Figure 3 shows a characteristic acquisition set-up.

To accelerate mosaic construction, we suppose that the scene is roughly planar. This simplifies the processing to finding only one dominant motion vector between two adjacent frames, and using that motion as the basis for registration of the images. The assumption of a planar scene, of course, does not hold for most under vehicle scenes, as there will always be some parts under the vehicle closer to the camera than others. This situation results in a phenomenon called motion parallax: objects closer to the camera will move past the camera's field of view faster than objects in the background. We assume, however, that these effects are negligible and will not adversely affect the goal of creating a summary of the under vehicle scene.

Other systems, for example, an *X-Y* plotter beneath a plate of glass with the car being driven over the inspection site, are not considered here due to their lack of flexibility in the inspection process. Our remotely controlled mobile inspection platform is able to revisit an area of interest beneath the car and to zoom into that area for a more detailed description if desired. Another approach could be to measure the movement of the wheels utilizing information from wheel encoders to determine the translation between one frame and the next. This would significantly simplify the computation of the translation between the frames and it would avoid the computationally-intensive Fourier processing. However, these data are not always reliable due to slippage, spinning, and other tire-soil interactions. In addition, the amount of translation within two consecutive video frames relates to the distance

between the optical center of the camera and the object surfaces beneath the car. Thus, the movement of the sensing platform by a specific distance beneath the undercarriage of a sports car results in a much larger inter-frame translation than the same movement of the sensing platform carried out beneath the undercarriage of a truck.

## 2.2 Distortion correction

The general framework of our algorithm is shown in Figure 4. Following the description of Chen (1998) for the general mosaic construction process, we divide the process into three processing steps. We correct each image in the sequence for barrel distortion and perspective distortion in the distortion correction step. In the registration step, we compute the motion associated with each frame in the sequence. Finally, in the merging step, we select strips from each image and paste them together based on the motion computed in the registration stage.

The barrel-distortion correction problem addressed in the correction step is fairly common. The parameters for correction used in this work were chosen manually, since we are only interested in reducing the more extreme distortions at the edge of the images. It is not required in our work that the correction be completely precise. Perspective distortion is performed to make the sequence appear as though the camera's principal axis was orthogonal to the plane of the scene. This step would not be necessary if the camera were pointed straight up at the vehicle underside during acquisition. In practice, due to vehicle ground clearance issues, the camera is usually pointed at an angle. Hence, perspective distortion correction

Figure 3 Video acquisition set-up using a camera mounted on a mobile platform

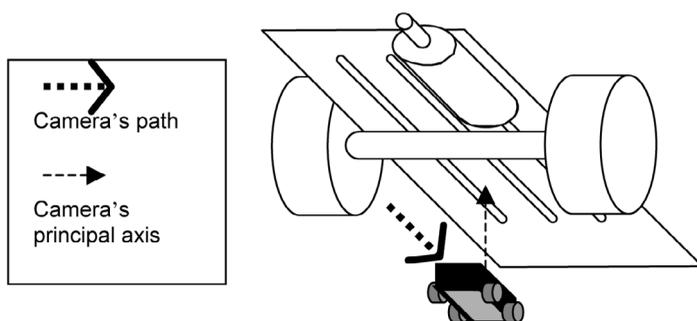
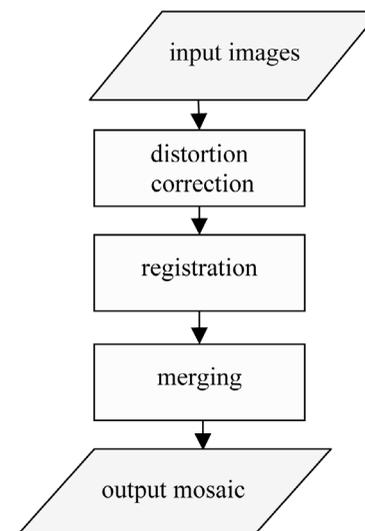


Figure 4 Flow chart of the mosaic construction process



is used to compensate for this. The reason we do this is that every element in the sequence displays pure translational motion and not more general affine motions. For this effort, the parameters for barrel and perspective distortion correction were chosen manually.

### 2.3 Registration using phase correlation

The registration step consists of computing the translational motion for each frame in the sequence. For any frame in the sequence, its motion vector is computed relative to the next frame of the sequence. The motion vector  $(u, v)$  may consist of shifts in the horizontal ( $u$ ) and vertical ( $v$ ) directions. Owing to motion parallax, there may be more than one motion vector present between two adjacent frames. Our aim is to compute, for a pair of adjacent frames, one dominant motion that may be used as the representative motion. Dominant motion is computed adopting the phase correlation method described by Kuglin and Hines (1975), since this technique is capable of extracting dominant inter-frame translation even in the presence of many smaller translations. Phase correlation has also proven to be applied successfully for tile inspection by Costa and Petrou (2000).

Phase correlation relies on the time shifting property of the Fourier transform. The Fourier transform of an image produces a spectrum of frequencies measuring the rate of change of intensity across the image. High frequencies correspond to sharp edges, low frequencies to gradual changes in intensity, such as lighting changes on large, angled planar surfaces. The spectrum  $F(\xi, \eta)$  is a frequency-signature of the contents of the image. By correlating the spectra of two images, the lines along which they match can be established, and the translation between the two can be found.

According to the property of the Fourier transform, a translation within the image plane corresponds to an exponential factor in Fourier domain. Suppose, we have two images, one being a translated version of the other, with a displacement vector  $(x_0, y_0)$ . Given the Fourier transforms of the two images,  $F_1$  and  $F_2$ , then the cross-power spectrum of these two images is defined as:

$$\frac{F_1(\xi, \eta)F_2^*(\xi, \eta)}{F_1(\xi, \eta)F_2(\xi, \eta)} = e^{j2\pi(\xi x_0 + \eta y_0)}, \quad (1)$$

where  $F_2^*$  is the conjugate of  $F_2$  and  $\xi$  and  $\eta$  are variables in the frequency domain corresponding to the displacement variables  $x, y$  in the spatial domain. The inverse Fourier transform of the cross-power spectrum, ideally, is zero everywhere except at the location of the impulse indicating the displacement  $(x_0, y_0)$  that corresponds to the

translation motion between the two images. The inverse Fourier transform of the cross-power spectrum is also referred to as the phase correlation surface. If there are several elements moving at different velocities in the picture, the phase correlation surface will produce more than one peak, with each peak corresponding to a motion vector. By isolating the peaks, a group of dominant motion vectors can be identified. This information does not specify individual pixel-vector relationships, but does provide information concerning motions in the frame as a whole. In our case, the strongest peak is selected as being representative of the dominant motion. Note that in our implementation, all images were resized to  $256 \times 256$  images prior to computing the Fourier transform, in order to simplify our DFT computation algorithm.

The results of the phase correlation algorithm may be affected by a phenomenon called discrete Fourier transform leakage (DFT leakage). DFT leakage occurs in most Fourier transforms of real images, and is caused by the discontinuities between the opposing edges of the original image. In order to deal with DFT leakage, a mask based on the Hamming window is applied to each image prior to calculating its Fourier transform. The equation for the one-dimensional Hamming window, which would provide the 1D weights of the tapering window, is

$$H(x) = 0.54 + 0.46 \cos\left(\frac{\pi x}{a}\right). \quad (2)$$

The resulting tapering window removes the discontinuities at the sides of the image while preserving a majority of the information towards the center of the images. All images are therefore tapered prior to computing their Fourier transforms applying equation (1). In addition, we apply restrictions to the search region within the phase correlation surface, based on the motion we would expect to see in the video sequence. The search region parameters are determined by minimum and maximum values for the horizontal and vertical motion vectors,  $u_{\min}, u_{\max}, v_{\min},$  and  $v_{\max}$ . These search region boundaries aid in reducing incorrect inter-frame motion estimates.

### 2.4 Merging and blending

Once the horizontal and vertical motions between two images have been computed by means of phase correlation, strips are acquired from one of the images based on those motions. One of the motions will correspond to the direction in which the camera moved during acquisition; this is called the primary motion. The other motion, which may be due to the camera deviating from a straight

path, or the camera's tilt, will be orthogonal to the primary motion and is called the secondary motion. The width of the strips is directly related to the primary motion. Adjacent strips on the mosaic are aligned using the secondary motion.

Although the strips may be properly aligned, seams may still be noticeable due to small motion parallax, rotation, or inconsistent lighting. A simple blending scheme is used in order to reduce the visual discontinuity caused by seams. Suppose in the mosaic  $D_m$  we have two strips sampled from two consecutive images,  $D_1$  (the image on the left) and  $D_2$  (the image on the right). The blending function is a one-dimensional function that is applied along a line orthogonal to the seam of the strips. For a coordinate  $i$  along this line, the intensity of its pixel in  $D_m$  is determined by:

$$D_m\left(b - \frac{w}{2} + i\right) = \underbrace{\left(1 - \frac{i}{w}\right)}_{A_1} \underbrace{D_1\left(c_1 + \frac{w_1}{2} - \frac{w}{2} + i\right)}_{B_1} + \underbrace{\left(\frac{i}{w}\right)}_{A_2} \underbrace{D_2\left(c_2 - \frac{w_1}{2} - \frac{w}{2} + i\right)}_{B_2}, \quad (3)$$

$$i = 1 \dots w,$$

where  $c_1$  and  $c_2$  are the coordinates corresponding to the centers of  $D_1$  and  $D_2$ , respectively,  $w_1$  and  $w_2$  are the widths of the strips sampled from  $D_1$  and  $D_2$ ,  $w = \min(w_1, w_2)$ , and  $b$  is the mosaic coordinate corresponding to the boundary between the two strips. The terms  $A_1$  and  $A_2$  are weights for the pixel intensities for  $D_1$  and  $D_2$ , while  $B_1$  and  $B_2$  are the pixel intensities themselves. For color images, this function is applied to the red, green, and blue components of the image. At a seam, this function adds weighted pixel values from the images that intersect at the seam. The weights of each pixel in a strip are a function of the distance of the pixel from the intersecting seam; the weights increase as pixels get closer to the midpoint of the strip from which they are sampled, and decrease as they get further apart. At the seam, the weights for pixels from both strips in an adjacent pair are equal, so that both adjacent images contribute equally to the pixel values at the seam. This simple blending technique has been chosen to accelerate the mosaic building process.

Note that results of higher image fidelity may be obtained for the color image mosaic when applying the (more computationally costly) technique of Hasler and Süssstrunk (2004). Nevertheless, this technique cannot be applied to the IR video sequence. Furthermore, note that blending merely creates more visual appealing imagery by

minimizing artifact seams. However, both blended and non-blended mosaics can be made available to aid inspection tasks.

After the blending is complete, the two strips have been successfully amalgamated. The process is then repeated for each subsequent frame in the video sequence. After each cycle of the merging process, the vertical and horizontal displacement of the last strip in the mosaic is recorded, and this information is used as the anchor for the next strip in the mosaic. Once every frame in the video sequence has been processed, the mosaic is complete.

### 3. Experimental results

Two image modalities were used to acquire the data used in this work: color video (visible-spectrum) and infrared video. The color video sequences for the under vehicle inspection efforts used in this work were taken using a Polaris Wp-300c Lipstick video camera mounted on a mobile platform. Infrared video was taken using a Raytheon PalmIR PRO thermal camera mounted on the same platform. The Lipstick camera has a focal length of 3.6 mm, a 1/3 in. interline transfer CCD with 525-line interlace and 400-line horizontal resolution while the Raytheon thermal camera has a minimum 25 mm focal length (36° horizontal and 27° vertical field-of-view). The tapering window parameter,  $a$ , was set to  $256/1.75 = 146.286$  for both sequences. We choose this value because it gives us a compromise between completely darkening the edges of each frame while retaining detail at the center of each frame (256 being the pixel-wise dimension of the resized frames). The search region parameters were set to:

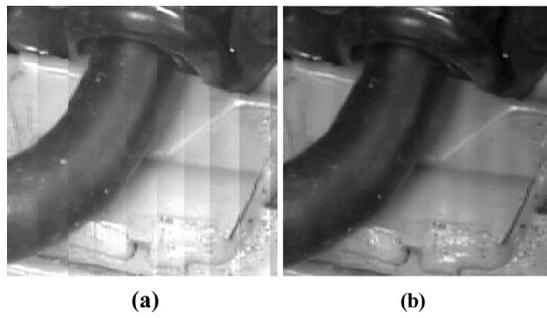
$$u_{\min} = -30, \quad u_{\max} = 30, \quad v_{\min} = -170,$$

$$v_{\max} = 0.$$

Here, we present the results of our mosaic building algorithm on two video sequences. The first, referred to here as UnderV3, is a visible-spectrum color video sequence. The second, IR1, is an infrared color video sequence. The necessity of applying a blending technique to the stitched mosaic for creating visual appealing mosaics is shown as example in Figure 5. The figure shows the results of creating mosaics (a) without blending and (b) with blending. Note the reduced discontinuities at the seams separating each strip in the mosaic after blending.

Finally, Figures 7 and 8 show the results of constructing mosaics of the UnderV3 and IR1 video sequences. Figure 6(a) shows four sample

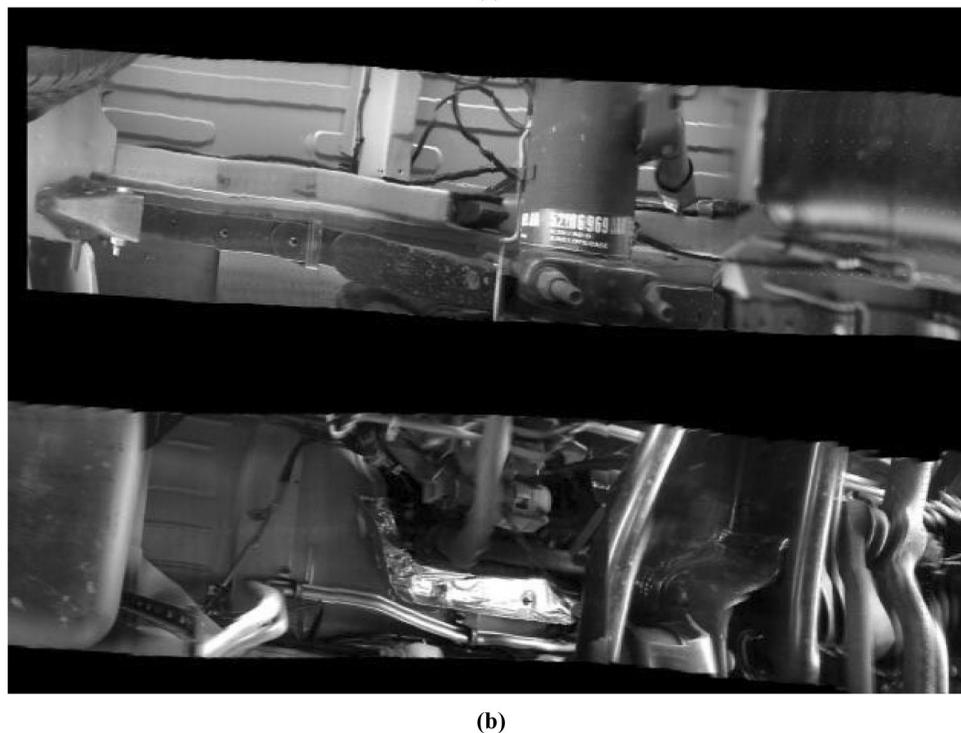
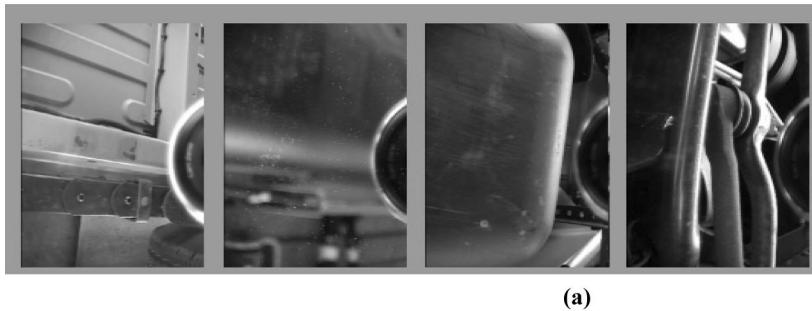
**Figure 5** Results of mosaic building: (a) without blending and (b) with blending



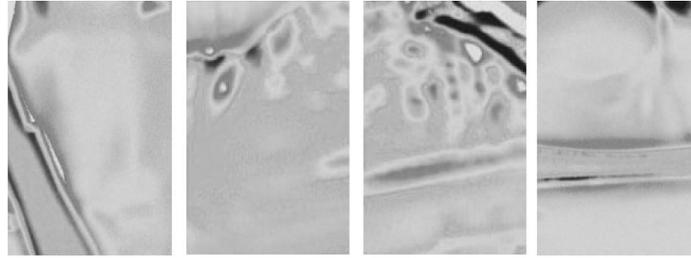
frames from color video sequence UnderV3 which has been acquired with a camera pointing to the undercarriage of a Dodge Ram. Two parts of the constructed mosaic of sequence UnderV3 are shown in Figure 6(b). The mosaic was

created from 196 frames. Figure 7(a) shows four sample frames from infrared video sequence IR1 which has been acquired in the same manner as the color video sequence UnderV3 but with an infrared camera. The mosaic is composed of 679 frames. From these results, it can be seen that our algorithm is capable of providing a good summary of these video sequences. There are still discontinuities visible in the mosaic due to motion parallax or absence of visual details that can be used to compute inter-frame motion (most noticeable in a large portion of the IR1 mosaic). Still, this algorithm performs well considering there are many parts of the IR1 sequence that display large homogenous areas. Well-known local-motion analysis techniques such as the Lucas and Kanade (1981) motion analysis algorithm may have problems identifying good global motion vectors for these sequences.

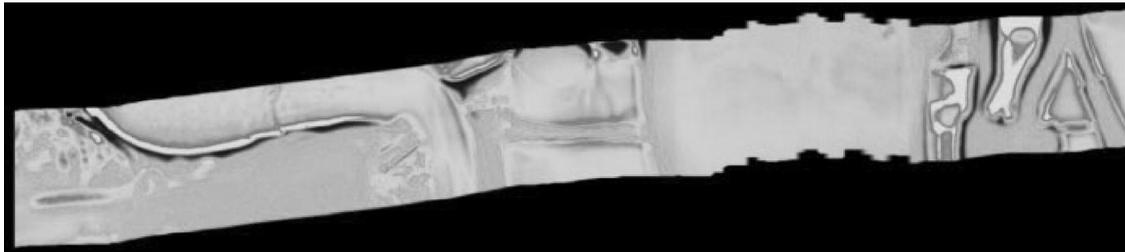
**Figure 6** (a) Sample frames from color video sequence UnderV3 which has been acquired with a camera pointing to the undercarriage of a Dodge Ram. (b) Two parts of the constructed mosaic of sequence UnderV3



**Figure 7** (a) Sample frames from infrared video sequence IR1 which has been acquired with a camera pointing to the undercarriage of a Dodge Ram, and (b) mosaic of IR1 sequence



(a)



(b)

#### 4. Conclusions

We have presented a method for building mosaics from video sequences for under vehicle inspection. The method uses phase correlation to perform registration and is capable of building mosaics from video sequences captured using infrared and visible-spectrum modalities. Given that many of the image sequences used here often display large homogenous areas with little visual detail, the phase correlation method is demonstrated to be a fairly robust registration method.

#### References

- Chen, Y. (1998), "Image stitching-comparisons and new techniques", Technical Report CITR-TR-30, Department of Computer Science, Center for Image Technology and Robotics, The University of Auckland, New Zealand.
- Costa, C.E. and Petrou, M. (2000), "Automatic registration of ceramic tiles for the purpose of fault detection", *Machine Vision and Applications*, Vol. 11, pp. 225-30.
- Hasler, D. and Süsstrunk, S. (2004), "Mapping colour in image stitching applications", *J. Visual Communication and Image Representation*, Vol. 15 No. 1, pp. 65-90.
- Irani, M., Anandan, P., Bergen, J., Kumar, R. and Hsu, S. (1996), "Efficient representations of video sequences and their applications", *Signal Processing: Image Communication*, Vol. 8, pp. 327-51.
- Kuglin, C.D. and Hines, D.C. (1975), "The phase correlation image alignment method", *Proc. Int. Conf. on Cybernetics and Society*, pp. 163-5.
- Lucas, B. and Kanade, T. (1981), "An iterative image registration technique with an application to stereo vision", *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 674-9.
- Page, D., Fougerolle, Y., Koschan, A., Gribok, A., Abidi, M., Gorsich, D. and Gerhart, G. (2004), "SAFER vehicle inspection: a multi-modal robotic sensing platform", *Proc. SPIE Unmanned Ground Vehicle Technology*, Vol. 5422.
- Peleg, S. and Herman, J. (1997), "Panoramic mosaics by manifold projection", *Proc. Conf. on Computer Vision and Pattern Recognition*, pp. 338-43.
- Zheng, J.Y. (2003), "Digital route panoramas", *IEEE Multimedia*, Vol. 10, pp. 57-67.