IMPROVED DESIGN FOR INCIDENT DETECTION IN TRANSPORTATION SYSTEMS USING REAL-TIME ACQUISITION OF PERSPECTIVELY CORRECT IMAGES

by

**Eleftheria Tzamali[1], George Akoumianakis[2], Antonis Argyros[3] and
Yorgos J. Stephanedes[4] [5]**

**Institute of Computer Science,
Foundation of Research and Technology – Hellas (FORTH)**

**Vassilika Vouton, P.O. Box 1385
GR-711-10, Heraklion, Crete, Greece**

---

[1] R&D Engineer, Institute of Computer Science, Foundation of Research and Technology – Hellas (FORTH), Vassilika Vouton, P.O. Box 1385, GR-711-10, Heraklion, Crete, Greece

[2] R&D Engineer, Institute of Computer Science, Foundation of Research and Technology – Hellas (FORTH), Vassilika Vouton, P.O. Box 1385, GR-711-10, Heraklion, Crete, Greece

[3] Assoc. Researcher, Institute of Computer Science, Foundation of Research and Technology – Hellas (FORTH), Vassilika Vouton, P.O. Box 1385, GR-711-10, Heraklion, Crete, Greece

[4] Visit. Researcher, Institute of Computer Science, Foundation of Research and Technology – Hellas (FORTH), 65 Ag. Ioannou, Athens 15341, Greece, email: yste@ath.forthnet.gr

[5] Corresponding author.

ABSTRACT

Advances in machine vision techniques have led to algorithms and integrated systems that can be applied in transportation engineering to improve surveillance and control. Despite these advances, certain problems in the effective integration of machine-vision based systems at complex intersections and complex freeway sections still remain. These are related to increasing system performance in the identification, analysis and detection of the traffic state in real time. This work examines the feasibility of providing transformed visual input to existing machine-vision based systems, in order to gain increased efficiency and cost-effectiveness of integrated transportation systems. Two transformations are developed, *homography-based transformation* and *panoramic image reprojection*.

Homography-based transformation operates on video of the road scene, provided by classical cameras, and seeks to transform any view to a top-down view. This transforms the three-dimensional problem of image analysis for, e.g., road event detection to a two-dimensional one. *Panoramic image reprojection* employs panoramic cameras to reduce required hardware, and the complexity and cost incurred in obtaining the desired road view. The image re-projection technique allows the reconstruction of undistorted, perspectively correct views from panoramic images in real time.

Tests at sites in Spain, the U.K. and Greece are performed on-line and off-line in combination with operating machine-vision based incident detection systems. Test results indicate that the two methods simplify the input provided to machine vision, and reduce the workload and amount of hardware in implementing complex machine-vision based systems for incident detection. Both modules can be integrated into incident detection systems to improve their overall efficiency and ease of application.

Keywords:

Incident, detection, machine vision, panoramic vision, homography, image reprojection, testing.

INTRODUCTION

Advances in machine vision techniques have led to development of algorithms and integrated systems that can be applied in transportation engineering to improve traffic surveillance and control (Stephanedes and Filippi 1996, Blosseville et al 1993, Michalopoulos and Jacobson 1993). Despite these advances, certain problems in the effective integration of machine-vision based systems at complex intersections and complex freeway sections still remain. These are related to increasing system performance in the identification, analysis and detection of the traffic state in real time (Stephanedes 2004).

In particular, once a road scene is defined, multiple cameras have to be installed such that the total input from their images fulfils the detection requirements, and appropriate windows have to be placed on the road scene such that image artefacts and resulting errors are minimized. Effects from artefacts include, e.g., effects from shadows, sun glare, angle of vehicle movement, distance between vehicles, and unusual vehicle size. Any of the effects can substantially increase the identification error and, therefore, reduce the effectiveness of machine vision.

The goal of this work is to examine the feasibility of providing transformed visual input to existing machine-vision based systems, so that higher-level image analysis is facilitated, leading to increased efficiency and cost-effectiveness of integrated transportation systems. To this end, two such transformations and methods have been developed, *homography-based transformation* and *panoramic image reprojection*. This paper presents the theory behind the employed view transformations as well as important aspects of their implementation. Results from testing at different sites, across countries, with the proposed techniques are also included.

BACKGROUND

Homography-based transformation operates on video of the road scene, provided by classical cameras, and seeks to transform any view to a top-down view by employing a homography-based view transformation module. The top-down view reduces design error since the user can treat it essentially in two dimensions, thus simplifying the placement and the ensuing analysis of the scene. Appropriately placing the window on the desired scene is a long and tedious process that often has to be repeated when results are not satisfactory or when design changes. Unfortunately, in practice, cameras can rarely be placed at top-down locations in the field. Most often, cameras are placed in the median or along the roadside, providing a view of the oncoming or departing vehicles at an angle. However, placing windows along side vehicle movement and at an angle is a difficult three-dimensional scenario for the analyst and the user. A major reason is that, at such placement, sensitivity of machine vision detection to small changes in window placement is high, increasing with the distance of the approaching or departing vehicle.

Through this transformation the three-dimensional problem of higher-level image analysis for, e.g., road event detection becomes a two-dimensional one, and the subsequent analysis is simplified. The transformation is valid for planar surfaces or surfaces that can be well approximated by planes. This assumption holds very often in traffic management applications, in which analysis is focused on planar or close-to-planar parts of a road. Additionally this assumption holds because, in most cases, traffic-monitoring cameras are placed at a distance from the imaged scene that is substantially larger than the depth variations within the scene.

Planes have been heavily studied in computational vision owing to their abundance in man-made environments, as well as to their attractive geometric properties. As reported in the literature, planes have been successfully employed in diverse applications such as feature matching (Lourakis et al 2000, Lourakis et al 2000a) and grouping (Van Gool et al 1998), camera self-calibration (Triggs 1998), obstacle detection (Lourakis and Orphanoudakis 1998), 3D reconstruction and scene analysis (Baillard and Zisserman 1999, Bartoli et al 2000, Criminisi and Zisserman 1998, Irani et al 1998, Kaucic et al 2001), camera relative positioning (Mohr et al 1992), object recognition (Rothwell et al 1995), visual measurement (Criminisi et al 1999), image mosaicing (Zoghlami et al 1997) and augmented reality (Simon et al 2000). Elegant theoretical results related to planes have also been derived (Zelnik-Manor and Irani 2002). The homography-based view transformation employed in this work is a direct consequence of such a result.

Panoramic image reprojection employs panoramic vision to reduce the number of required cameras and the complexity and cost incurred in obtaining the desired road view at complex intersections and complex freeway sections. Technologically, this capability can be implemented using perspective cameras, multi-camera systems or panoramic cameras. In particular, it would be possible to use perspective cameras and alter their gaze direction via pan-tilt platforms. Alternatively, multi-camera systems could be used, in which cameras jointly provide a wide field of view. Both alternatives, however, may present significant mechanical, perceptual, and control challenges (Argyros et al 2004).

Panoramic cameras can provide a wide field of view (up to 360°) (Benosman and Kang 2001), and allow effortless and instantaneous switching of viewing directions, thus emerging as an effective sensor in several application domains. Nevertheless, because of their optical properties, panoramic cameras provide a distorted view of the imaged scene. For this reason, an image re-projection technique is employed that allows the reconstruction of undistorted, perspectively correct views from panoramic images, in real time. The two methods are described in the next section.

TRANSFORMATION DESCRIPTION

Homography-based transformation

An important concept in projective geometry[1] (Hartley and Zisserman 2000) is the plane homography H, a non-singular 3x3 matrix which relates two uncalibrated retinal images of a 3D plane. More specifically, if x is the projection in one view of a 3D point X on the plane and x' is the projection of the same 3D point in a second view, then the two projections are related by the linear projective transformation:

$$x' \simeq Hx \tag{1}$$

where, vectors and arrays are represented using projective (homogeneous) coordinates, 3D points are written in upper case, their image projections in lowercase (e.g. X and x), and the symbol " $\simeq$ " denotes equality of vectors up to an arbitrary scale factor.

Since eq. (1) holds up to an unknown scale factor, the 3x3 homography matrix H can be fully determined by eight parameters. This, in turn, implies that, if a correspondence of at least four points between two different views of a certain planar surface is established, the homography connecting these two views of the plane can easily be estimated through the solution of a linear system of equations. This is because, from each point correspondence, we can derive

---

[1] For a more detailed treatment of the application of projective geometry to computer vision, the interested reader is referred to (Hartley and Zisserman 2000).

two constraints (equations) on elements of *H*. Thus, four points is the theoretic minimum number of points that suffice to fully determine *H*. In practice, more elaborate minimization schemes are employed to come up with a more accurate estimation of *H* (Lourakis et al 2002).

The above procedure is typically employed in vision to segment a scene into 3D planes (Lourakis et al 2002). In the particular application domain in this paper, the same theoretical result is used in a slightly different manner. The user defines a part of the road that corresponds to a planar surface that has the shape of an oblong (quadrilateral with parallel sides forming right angles). An example of such region is defined by points p_1, p_2, p_3 *and* p_4 in the image at the left part of Figure 1.

-- Figure 1 --

Owing to perspective projection and the obliqueness of the view, the oblong maps to a general quadrilateral on the image plane. Since the selected region is by definition an oblong, we know that, if it is observed in a top-down view, it will appear as an oblong on the image plane, too. Thus, we may define four new points, i.e., p'_1, p'_2, p'_3 and p'_4, such that

- p_i corresponds to p'_i,
- all p'_is form an oblong, and
- the sides of the oblong are proportional to the true dimensions of the road part so that the final, transformed image does not appear distorted.

Provided that the above constraints are fulfilled, the exact coordinates of p'_is are not important. The point selection process results in four point correspondences, which can be used to define a homography *H* connecting the original view of the road to the new, “virtual”, top-down view. We can, therefore, solve the linear system of eight equations based on eq. (1), which will lead to the estimation of homography matrix *H* connecting the view of the two planes. Once this is achieved, the view transformation becomes a simple process. Since eq. (1) holds for all points belonging to the plane, eq. (1) also guarantees that each point *p* in the quadrilateral $p_1 p_2 p_3 p_4$ is mapped on the point with coordinates $p' \simeq Hp$.

In practice, the algorithm operates inversely. More specifically, the rectangular region in the target (top-down) view is scanned in a row-major manner and, for each point *p',* we determine the pixel

$$p \simeq H^{-1} p' \qquad (2)$$

in the original image, from which p' should “borrow” the intensity or the color. This guarantees that all pixels of interest (i.e., all points in oblong $p_1' p_2' p_3' p_4'$) are assigned some color.

Equation (2) typically results in real-numbered image coordinates *p*. For producing images of better quality the four image points with integer coordinates that are closer to p are determined and their colors are used in a bilinear interpolation scheme to determine the color of point p'.

Interestingly, once the homography matrix *H* is computed in a preprocessing step, the mapping of point coordinates between the original and the transformed view can be computed once and saved as a look-up table that can subsequently be used to implement the transformation. This relieves the whole operation from the expensive matrix-vector multiplication involved in eq. (2) and makes the transformation amenable to real-time implementation by even moderate hardware.

Acquiring perspectively correct images from panoramic views

Panoramic cameras are cameras that are able to capture simultaneously a full 360^o visual field. This ability, in conjunction with other beneficial geometric properties, becomes increasingly important in a number of applications, such as robot navigation based on visual information, tele-presence and teleconferencing systems, and web-applications.

Alternatives to covering the full 360^o field of view include either a standard camera mounted on a pan and tilt device, or the use of a number of cameras, arranged on a ring so that they have overlapping visual fields. Both alternatives have a number of disadvantages compared to the option of a panoramic camera. In the case of a standard camera mounted on a pan and tilt device, the complexity of the system is increased because of the need for controlling that device. Moreover, with such a configuration it is impossible to have an instant view of the whole visual field, since camera motion introduces certain delays. This disadvantage is particularly important in cases where the camera is about to observe a dynamic scene, such as one involving fast-moving objects.

This problem can be alleviated when a number of cameras with overlapping visual fields are used. The cameras can be synchronized so that all images are acquired simultaneously. In this case, however, the complexity and cost-effectiveness of the system increase because of the need for multiple cameras, multiple digitization devices (frame-grabbers) as well as synchronization hardware. Moreover, the construction of a full panorama of the environment is a time consuming operation, and its accuracy heavily depends on the quality of the calibration of the overall system.

Panoramic vision is achieved by a combination of a standard camera with a hyperbolic or parabolic mirror that is placed on top of a standard CCD camera. Several types of panoramic camera are commercially available, differing in characteristics such as mirror type, mirror placement with respect to the imaging device, and optical characteristics of the imaging device. Figure 2 illustrates the geometry of a typical panoramic camera. An imaging device (typically an ordinary perspective camera) is located at a point O. O is also the center of the Cartesian, 3D coordinate system of the camera. The optical axis of the camera coincides with the *Z*-axis of the coordinate system. The mirror of the panoramic camera is, in turn, placed so that its axis of symmetry coincides with the optical axis of the camera. Assume now a 3D point $P(X,Y,Z)$ on an object of the environment. A 3D ray of light emanating from this point reflects on the mirror surface at the 3D point $P_m(X_m,Y_m,Z_m)$. The illumination of this point is then projected onto the photosensitive surface of the camera (image plane π) at coordinates $p_{pan}(x_{pan},y_{pan})$. This process results in a ring-like image, see Figure 6.

- - Figure 2 - -

Surveillance of freeways and complex road junctions for the purpose of incident detection is an application area that can take advantage of such a wide, panoramic field of view. It appears natural that the ability to combine information from a panorama can lead to the coverage of a broad area and to the integration of information from multiple visual directions. As several products for vision-based traffic incident detection are commercially available, the increased sensing capabilities (in terms of visual field) of a panoramic camera can be combined with the power of the image analysis algorithms already available.

However, because of the specific type of image projection, several geometric distortions are introduced. For example, straight 3D lines do not necessarily map to straight lines on the image, as it happens in typical, perspective images. These geometric distortions hinder the application of existing image analysis algorithms to machine-vision based incident detection; such algorithms have been developed under the assumption that they will be employed in images acquired by typical cameras.

It is, therefore, necessary to retrieve perspectively correct images out of the visual information provided by panoramic cameras that have the geometry depicted in Figure 2. This is achieved by inverting the image projection process. More specifically, assume that a panoramic image like the one of Fig. 6 is available. We also assume that a virtual, photosensitive imaging plane π' is located at a known 3D position and orientation with respect to the camera coordinate system (e.g., π' in Figure 2). Then, we will reconstruct the image that this “virtual camera” would acquire.

To this end, we assume that, through a camera calibration process, all camera parameters are known, either as a result of a self-calibration algorithm or of an off-line, grid-based calibration method (Lourakis and Deriche 2000). The matrix of the intrinsic calibration parameters has the well-known form (Hartley and Zisserman 2000) presented in eq. (3):

$$K = \begin{bmatrix} a_u & -a_u \cot\theta & u_0 \\ 0 & a_v / \sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

The parameters a_u and a_v of eq. (3) correspond to the focal distances in pixels along the axes of the image, θ is the angle between the two image axes, $-a_u \cot\theta$ is the camera skew, and (u_0, v_0) are the coordinates of the image principal point. In addition to the intrinsic camera parameters, it is important to know the shape and positioning of the mirror with respect to the camera coordinate system. Then, for a specific imaged point $p_{per}(x_{per}, y_{per})$ we can define the 3D line ℓ connecting O and $p_{per}(x_{per}, y_{per})$.

Moreover, we can analytically compute the 3D point $P_m(X_m, Y_m, Z_m)$ as the intersection of line ℓ with the mirror. Further, we can analytically compute the 3D ray ℓ' corresponding to the reflection of the 3D line ℓ on the mirror surface. Last, we intersect the 3D line ℓ' with the plane π' of the "virtual" camera. Similarly to the case of the homography-based transformation, once the parameters of the camera are known and plane π' is positioned with respect to it, the mapping of intensities between the original, panoramic image and the new, perspectively corrected image remains constant and can be efficiently implemented through a look-up table operation. Figure 7 shows an example re-projected image that has been derived using input from the panorama of the image in Figure 6.

We have developed a system that performs, in real-time, the role of multiple virtual perspective cameras using the video input from a 360-degree panoramic camera, based on the above construction. The input to the system application can be either a sequence of images (saved on disk in bitmap or Windows metafile format), or live input from an actual panoramic camera. The virtual perspective camera can be at any position and can look at any direction, within the visual field of the panoramic camera. Multiple different views can be opened simultaneously, and the number of views is constrained only by the speed of the host machine. Five predefined views (front view, back view, right view, left view and bird's-eye view) and two special views (half size and double size) enable quick resizing of the already defined view.

Besides the predefined views, the user is also able to define custom views, by defining a number of parameters related to the position, looking direction, resolution, field of view of the virtual perspective camera as well as the parameters of the panoramic camera, i.e., CCD, lens and parameters of the mirror. In that sense, the system can deliver any possible perspective view and receive input from any panoramic camera. The parameters of any user-defined view can be saved on disk for later retrieval and use.

TESTING

Each of the developed transformations is implemented as a separate module that enables real time video transformation through a highly configurable user interface. In order to test the developed transformations a series of tests have been performed. For all tests, on-line data are used from different sites across countries, and image analysis is performed through off-line assessment and on-line testing. This section describes the testing methodology, tests performed, results and performance obtained in these tests, for each transformation module. On-line input data are from sites in the cities of Barcelona, Spain; Southampton, UK; and Athens, Greece.

A. Tests of the homography-based module

The goal of testing the homography-based module is to quantify the potential of the transformation in producing top-down views of an arbitrary view obtained from a classical machine vision camera, leading to incident detection at performance comparable to incident detection from raw images. The tests in this section include, (a) pre-test sensitivity analysis, (b) off-line feasibility assessment, and (c) on-line testing.

(a) Pre-test sensitivity analysis

In preparation for testing, a number of classical camera views were processed by the homography-based transformation, and the sensitivity of results to changes in exogenous factors was assessed in a qualitative manner. Particular emphasis was placed in testing several practical camera setups that could be employed by the transportation and traffic engineer in the field. It was also expected that findings, especially with respect to camera placement, would assist with the design of the actual testing in this work.

Findings from this analysis indicated that, in most cases, system performance is acceptable. In particular, from a total of twenty-seven sensitivity tests, only one produced unacceptable results (percentage of 3.7%).

More specifically, findings indicated what was theoretically expected, i.e., performance is influenced by the relative location and exact position of the camera in the field. Performance improves,

(a) as the distance from the observed scene increases, and
(b) as the optical axis of the camera becomes perpendicular to the plane of the road.

In addition, the quality of the homography-based transformation does not depend on the traffic conditions or on the lighting conditions. This is as expected since the transformation is of a geometric nature, not of a photometric one. Owing to its nature, it determines the re-mapping of pixel intensities in space, not the pixel intensities themselves.

These findings are directly related to the assumptions inherent to the proposed transformation, and lead to better understanding on camera placement that should result in improved system performance.

(b) Off-line feasibility assessment

Prior to on-line testing, we sought to assess whether, for an existing incident detection algorithm, for a typical detection scenario (here, stopped vehicles), and for road scenes for which such detection can be made successfully using raw images, the detection can also be performed using images that have been transformed by the homography-based module. In all cases, it was sought to not introduce *new* false alarms to the detection process, i.e., alarms that would be caused only by the transformation.

For the purposes of this assessment, classical machine vision sequences were transformed and fed into an operational incident-detection machine-vision algorithm previously developed by the research team. Further, the homography-based module was integrated with the detection algorithm. Camera location was selected to increase system performance, adopting the recommendations from the findings during the pre-test sensitivity analysis.

Background

An incident detection algorithm based on background subtraction, previously developed by the research team, has been used to assess the quality of the images produced by the homography-based module. The algorithm specializes in detecting stopped vehicles, and has been reported elsewhere (McDonald and Stephanedes 2002). To be sure, the logic from other incident detection tools, including single-station (Michalopoulos and Jacobson 1993) and two-station algorithms (Stephanedes and Chassiakos 1993), could alternatively be employed. The method operates as follows. In each transformed image two subtractions are performed;

subtracted are, (a) a selected (also transformed) background image giving information about vehicle's existence and (b) its previous (transformed) frame of the sequence giving evidence of vehicle's movement. The results from these subtractions are combined to provide evidence of stopped vehicles. A number of parameters control the performance of the underlying algorithm. These parameters are related to the dimensions of the region of interest, threshold values and time duration of the incident to be detected. In these tests, all threshold values are kept constant.

Assessment results

Assessment analysis was performed over a period of eighteen months, to gain an indication of possible performance changes over time; no such changes were observed. Typical assessment scenarios sought to detect the queues of stopped vehicles that are formed as a result of either traffic lights or traffic congestion. When no such queues were detected, it was sought to confirm the normal situation on the road. The scenarios took place on busy 2-3 lane urban arterials in Barcelona, Southampton and Athens. Typical time duration of each scenario was 2-20 min. and included both stopped-vehicles and normal-traffic cases.

Figure 3 shows snapshots of the integrated application while running. On the upper-left part of the dialog the original video appears. A red contour indicates the selected region of interest. The transformed region of interest is also depicted. The results of the analysis are demonstrated on the dialogs at the right of the interface. When stopped vehicles are detected, a warning appears on the dialog.

- - Figure 3 - -

The results indicate that the integrated homography-based transformation and detection algorithm are successful in delivering top-down views of the road and in detecting stopped vehicles from the transformed images. In particular, in all cases in which detection had been possible with the raw images, it was also performed with the transformed images, and no new false alarms were introduced by the image transformation. Further, the method performed correctly under different light conditions and at the same basic parameter settings for all tests. Calibration information was neither available nor needed; the transformation was easily derived, by exploiting information present in the images.

While the above results are encouraging, they relate to detection of only one type of incident, i.e., stopped vehicles. As pointed out in the literature (Stephanedes and Chassiakos 1993), other types of incident may be harder to detect, and this could influence the performance of the homography-based transformation. Therefore, additional tests with other types of incident will be forthcoming.

(c) On-line tests

On-line testing was performed following successful feasibility assessment. In particular, the transformation module acquired classical machine vision on-line sequences, and the transformed images were, in turn, fed into a commercially available incident detection machine-vision system that has been successfully operating at one of the sites. The machine vision system counted the number of vehicles that passed in each lane of the road, in both the original (unprocessed) and resulting (transformed) sequences. The correct figures were also known as vehicle counts by human visual inspection of the taped sequences. The comparison of these figures quantifies the quality of the transformed images. In these on-line tests, camera location was selected to increase system performance, adopting the recommendations from the findings during the pre-test sensitivity analysis. For these tests, a typical test sequence had duration of thirty minutes.

A presentation of the tests performed and the results obtained follows. A sample image from the test video stream is presented in Figures 4 and 5.

- - Figure 4 - -

- - Figure 5 - -

Results and test performance

The results obtained from the test are summarized in Table 1. The first and second rows show the number of vehicles (volume) counted by machine vision in the original and in the transformed sequence, respectively. The third row is “ground truth”, i.e., the actual count of vehicles, provided by human, visual inspection. Each of the first three data columns provides the number of vehicles counted for each of the three lanes of the road, and the last column provides the total number of vehicles for all lanes.

TYPE OF SEQUENCE	**Left lane**	**Middle lane**	**Right lane**	**All lanes**
	Vehicle Count	**Vehicle Count**	**Vehicle Count**	**Vehicle Count**
Original sequence	58	114	56	**228**
Transformed sequence	56	107	58	**221**
Ground truth	57	112	61	**230**

Table 1. Performance results of homography-based transformation

From the vehicle counts in Table 1 it can be seen that for the left lane, the original and the transformed sequence perform equally well. For the middle lane, the original sequence gives better results and, for the right lane, the transformed sequence gives better results.

B. Tests of the panoramic reprojection module

A beneficial property of panoramic cameras is their increased field of view, which enlarges the observation area. Nevertheless, a problem that arises in the use of such cameras is the geometric distortion of panoramic images, which is due to characteristics in their acquisition process. Owing to this distortion, panoramic images need perspective correction to become usable by existing machine vision techniques for incident detection. This work has developed such a geometric correction algorithm, which is realised by the panoramic reprojection module. The goal of testing this module is to determine whether the corrected images can be used in machine-vision-based, traffic monitoring applications, and to test the performance of the module.

Testing methodology

The panoramic reprojection module acquired panoramic image sequences, and the resulting output sequence was, in turn, fed into the commercial machine-vision based incident detection system that had been operating at the site. The machine vision system then performed vehicle counting in the resulting (transformed) sequence.

A key aspect of the testing methodology is that the panoramic video was synchronised with the video of the classical machine vision camera that has been used for testing the homography-based module (see above). Thus, the results of processing (a) the classical camera output, (b) the homography-based transformation, (c) the transformed panoramic video, and (d) the real vehicle count were comparable, at least for the common area view of all these views.

Tests performed

A panoramic camera was mounted at a height of approximately nine meters from ground level. The mounting location was selected such that it allowed the best panoramic view of traffic movements for the camera. A commercial machine vision camera and the panoramic camera provided synchronised video streams of the activity in the road that was being observed by the cameras.

To allow easy comparison of the corrected panoramic sequence with the results from the homography-based transformation, results from the same Athens site are presented in Figures 6 and 7. In particular, Figure 6 gives example of the original panoramic image; and Figure 7 presents the corrected panoramic image from Figure 6. These images are to be compared to the original machine vision image of Fig. 4, and the homography-based transformation of the machine vision image of Figure 5.

- - Figure 6 - -

- - Figure 7 - -

Results and test performance

The results obtained from the processing of traffic sequences are summarized in Table 2, for three different lanes of the observed road. For the purpose of immediate comparison, the results obtained by the testing of the homography-based transformation are repeated in this table.

TYPE OF SEQUENCE	**Left lane**	**Middle lane**	**Right lane**	**All lanes**
	Vehicle count	**Vehicle count**	**Vehicle count**	**Vehicle count**
Original machine vision sequence	58	114	56	**228**
Transformed machine vision sequence	56	107	58	**221**
Corrected panoramic sequence	41	100	86	**227**
Ground truth	57	112	61	**230**

Table 2. Comparative performance results of homography-based transformation and panoramic-image reprojection

The results in Table 2 indicate that, although the principle of transformation is sound, the lane-by-lane performance of the vehicle counting process in the corrected panoramic images is worse when compared to the performance of the vehicle counting process that is applied to the original machine-vision camera images or to the homography-based transformed images. The reason for this loss of performance is attributed primarily to the height of the camera placement and the specific type of panoramic camera used. More specifically, the camera should be placed closer to the road, to exploit a larger portion of its visual field and increase the resolution of the output images. Further improvement in system performance should result from using a panoramic image with a better quality CCD array and higher resolution.

SUMMARY

A homography-based transformation has been developed for transportation systems, which can be used to simplify monitoring and detection applications that require the use of a camera input from the road scene in real time. The transformation allows the user to place the actual camera along the side of the vehicle movement and at an angle. Furthermore, the technique allows the user to view the real scene from a top-down view. This reduces the three-dimensional image-analysis problem to a two-dimensional one, thus facilitating higher-level image analysis.

Results indicate that the homography-based transformation, when integrated with an existing incident detection algorithm is successful in delivering top-down views of the road, and no new false alarms are introduced by the transformation. For all road lanes, the original and the transformed sequence perform similarly, indicating potential applications of the new algorithm in traffic monitoring and detection applications that are based on machine vision. Performance will further improve if camera placement respects the constraints that result from the assumptions inherent to the transformation.

The results obtained from testing the panoramic image reprojection technique indicate that the basic principle of the transformation algorithm is sound. However, the panoramic camera

should be placed closer to the road, to exploit a larger portion of its visual field. In this manner, the resolution of the output images will be increased with direct favourable impact on the quality of the results. A panoramic image with a better quality CCD array and higher resolution should be used for further improving the results.

The results reflect the potential of the new methods to simplify the input that is provided to machine vision techniques, and to reduce the workload, potential for human error, and total cost in installing, calibrating and updating complex machine-vision based hardware and software used for incident detection. Although research to date has focused on a limited number of incident types, application of the methods to locations at different sites and across a number of countries provides important feedback on the transferability potential of the new methods.

ACKNOWLEDGEMENT

This work was performed as part of project PRIME, which was partially funded by the European Union. Participating organisations were from the U.K., Greece, Spain, Germany, and the Netherlands. Data provided by cooperating sites in these countries are gratefully acknowledged.

REFERENCES

Argyros, A. A., Tsakiris, D. P., and Groyer, C. (2004). “Bio-mimetic centering behavior: Mobile robots with panoramic sensors.” *IEEE Robotics and Automation Magazine, Special Issue on Panoramic Robotics,* IEEE, New York, NY, 21-30.

Baillard, C., and Zisserman, A. (1999). “Automatic reconstruction of piecewise planar models from multiple views.” *Proc., Computer Vision and Pattern Recognition,* IEEE, New York, NY, 559–565.

Bartoli, A., Sturm, P., and Horaud, R. (2000). “A projective framework for structure and motion recovery from two views of a piecewise planar scene.” *Technical Report RR-4070,* MOVI / Institut National de Recherche en Informatique et en Automatique, France.

Benosman, R., and Kang, S. B., eds. (2001). *Panoramic vision.* Springer-Verlag, New York, NY.

Blosseville, J., Morin, J., and Lochegnies, P. (1993). “Video image processing application: Automatic incident detection in freeways.” *Proc., Pacific Rim Transportation and Technology Conference,* ASCE, Reston, Va., 69-76.

Criminisi, A., Reid, I., and Zisserman, A. (1998). “Duality, rigidity and planar parallax.” *Proc., European Conference in Computer Vision,* Springer-Verlag, Berlin, Germany, 846–861.

Criminisi, A., Reid, I., and Zisserman, A. (1999). “A plane measuring device.” *Image and Vision Computing*, 17, 625–634.

Hartley, R., and Zisserman, A. (2000). *Multiple view geometry in computer vision.* Cambridge University Press, Cambridge, MA.

Irani, M., Anandan, P., and Weinshall, D. (1998). “From reference frames to reference planes: Multi-view parallax geometry and applications.” *Proc., European Conference in Computer Vision*, Springer-Verlag, Berlin, Germany, 829–845.

Kaucic, R., Hartley, R., and Dano, N. (2001). “Plane-based projective reconstruction.” *Proc., International Conference in Computer Vision*, IEEE, New York, NY, I, 420–427.

Lourakis, M., and Deriche, R. (2000). “Camera self-calibration using the singular value decomposition of the fundamental matrix.” *Proc., Asian Conference in Computer Vision*, Springer-Verlag, Berlin, Germany, 403–408.

Lourakis, M. I. A., and Orphanoudakis, S. C. (1998). “Visual detection of obstacles assuming a locally planar ground.” *Proc., Asian Conference in Computer Vision*, Springer-Verlag, Berlin, Germany, 2, 527–534.

Lourakis, M. I. A., Halkidis, S. T., and Orphanoudakis, S. C. (2000). “Matching disparate views of planar surfaces using projective invariants.” *Image and Vision Computing*, 18, 673–683.

Lourakis, M. I. A., Argyros, A. A., and Orphanoudakis, S. C. (2002). “Plane detection in an uncalibrated image pair.” *Proc., British Machine Vision Conference*, British Machine Vision Association, Worcs, UK, 2, 587-596.

Lourakis, M. I. A, Tzurbakis, S. V., Argyros, A. A., and Orphanoudakis, S. C. (2000a). “Using geometric constraints for matching disparate stereo views of 3D scenes containing planes.” *Proc., International Conference On Pattern Recognition*, International Association for Pattern Recognition, Buffalo, NY, 419–422.

McDonald, M., and Stephanedes, Y. J., eds. (2002). *Prediction of congestion and incidents in real time, for intelligent incident management and emergency traffic management*, Fin. Report, PRIME IST-13036, Brussels, Belgium.

Michalopoulos, P., and Jacobson, R. (1993). “Field implementation and testing of machine vision based incident detection system.” *Transportation Research Record*, 1394, 1-7.

Mohr, R., Morin, L., and Grosso, E. (1992). “Relative positioning with uncalibrated cameras.” *Geometric invariance in computer vision*, J. L. Mundy and A. Zisserman, eds., MIT Press, Cambridge, MA, 440-460.

Rothwell, C. A., Zisserman, A., Forsyth, D. A., and Mundy, J. L. (1995). "Planar object recognition using projective shape representation." *International Journal in Computer Vision*, 16(1), 57–99.

Simon, G., Fitzgibbon, A., and Zisserman, A. (2000). "Markerless tracking using planar structures in the scene." *Proc., Int'l Symposium on Augmented Reality,* IEEE, New York, NY, 120-128.

Stephanedes, Y. J. (2004). "Intelligent Transportation Systems." *The Engineering Handbook, 2nd Ed.*, R. C. Dorf, ed., CRC Press, Boca Raton, Florida, Ch. 86, ISBN 0-8493-1586-7.

Stephanedes, Y. J., and Chassiakos, A. (1993). "Freeway incident detection through filtering." *Transportation Research Journal*, 1C:2, 219-234.

Stephanedes, Y. J., and Filippi, F., eds. (1996). *Applications of advanced technologies in transportation engineering,* ASCE, Reston, Va., 742 pages, ISBN 0-7844-0146-2.

Triggs, B. (1998). "Autocalibation from planar scenes." *Proc., European Conference in Computer Vision,* Springer-Verlag, Berlin, Germany, 89–105.

Van Gool, L., Proesmans, M., and Zisserman, A. (1998). "Planar homologies as a basis for grouping and recognition." *Image and Vision Computing,* 16, 21–26.

Zelnik-Manor, L., and Irani, M. (2002). "Multiview constraints on homographies." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, New York, NY, 24(2), 214–223.

Zoghlami, I., Faugeras, O., and Deriche, R. (1997). "Using geometric corners to build a 2D mosaic from a set of images." *Proc., Computer Vision and Pattern Recognition*, IEEE, New York, NY, 420–425.

LIST OF FIGURES

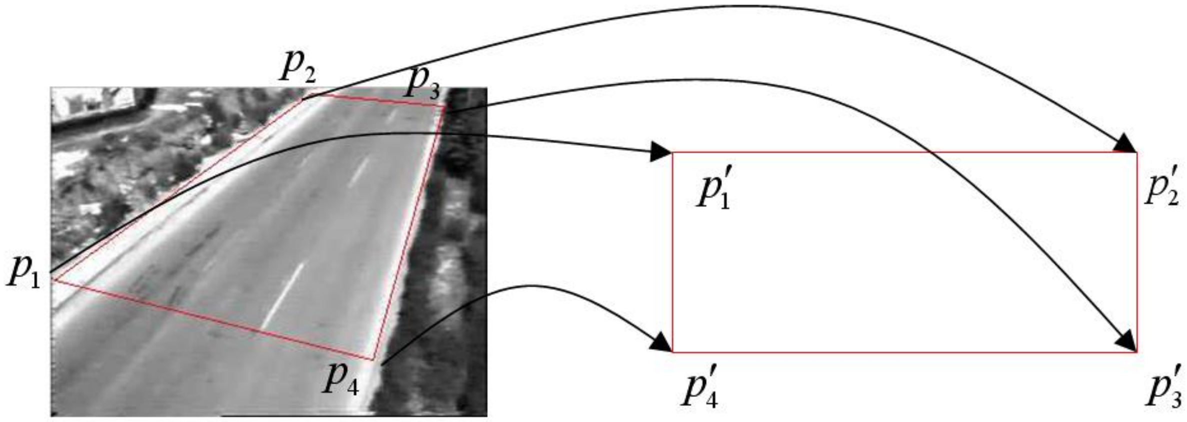


Figure 1. Graphical illustration of the use of projective geometry to derive a homography-based transformation that transforms an oblique view of a planar surface to a top view.

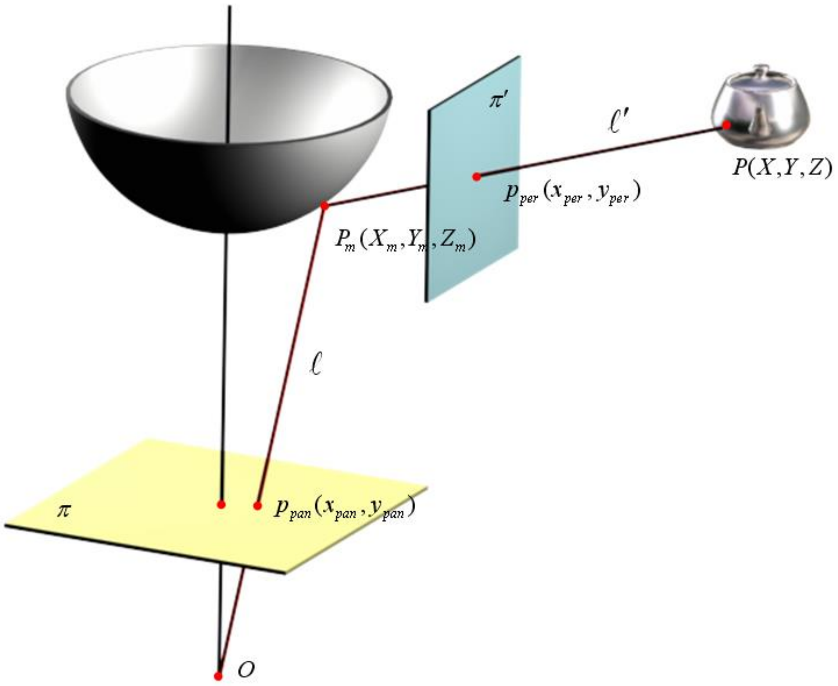
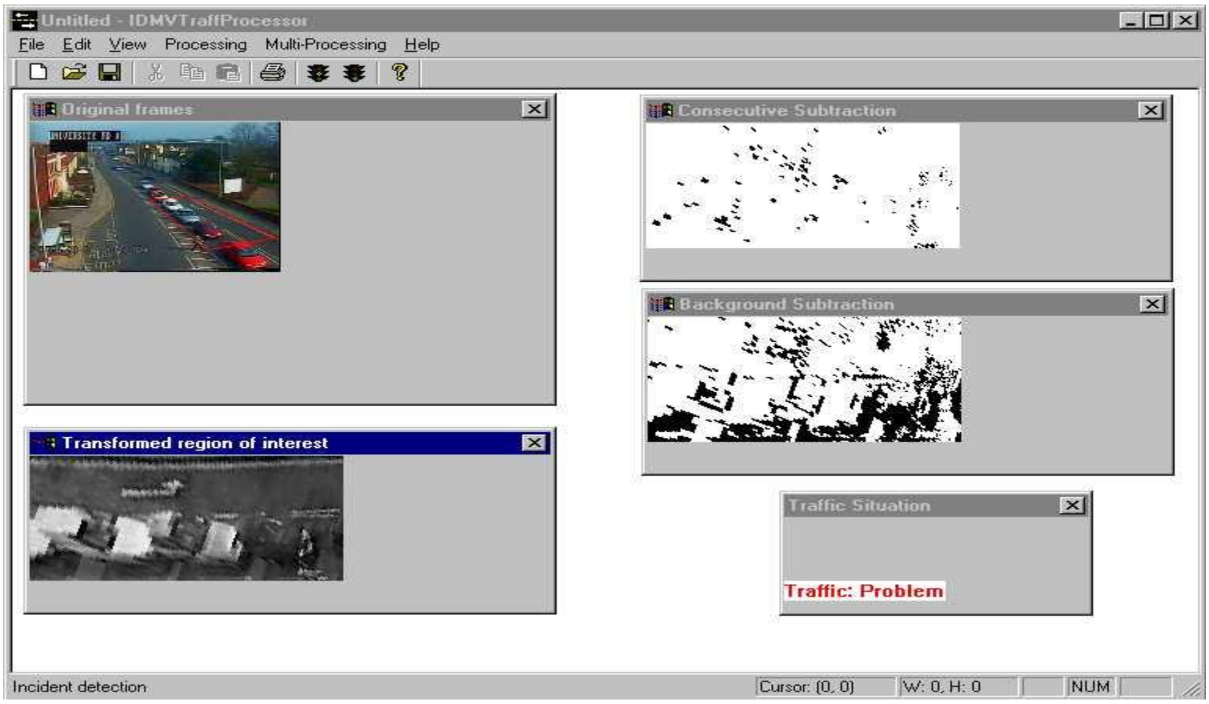


Figure 2. Graphical illustration of geometry of panoramic camera with paraboloidal mirror.



Traffic: Problem

Figure 3. View of the integrated detection during operation, Southampton

Figure 4. View from detection on Kymis Str., Athens – Input frame

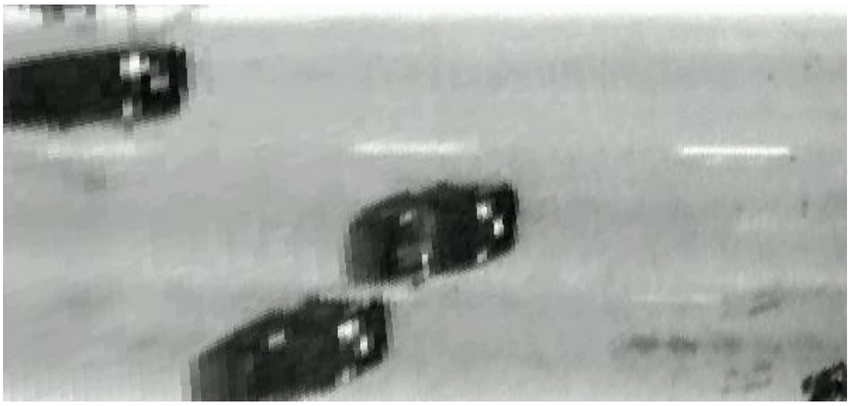


Figure 5. View from detection on Kymis Str., Athens -- Resulting homography-based transformation for image of Figure 4

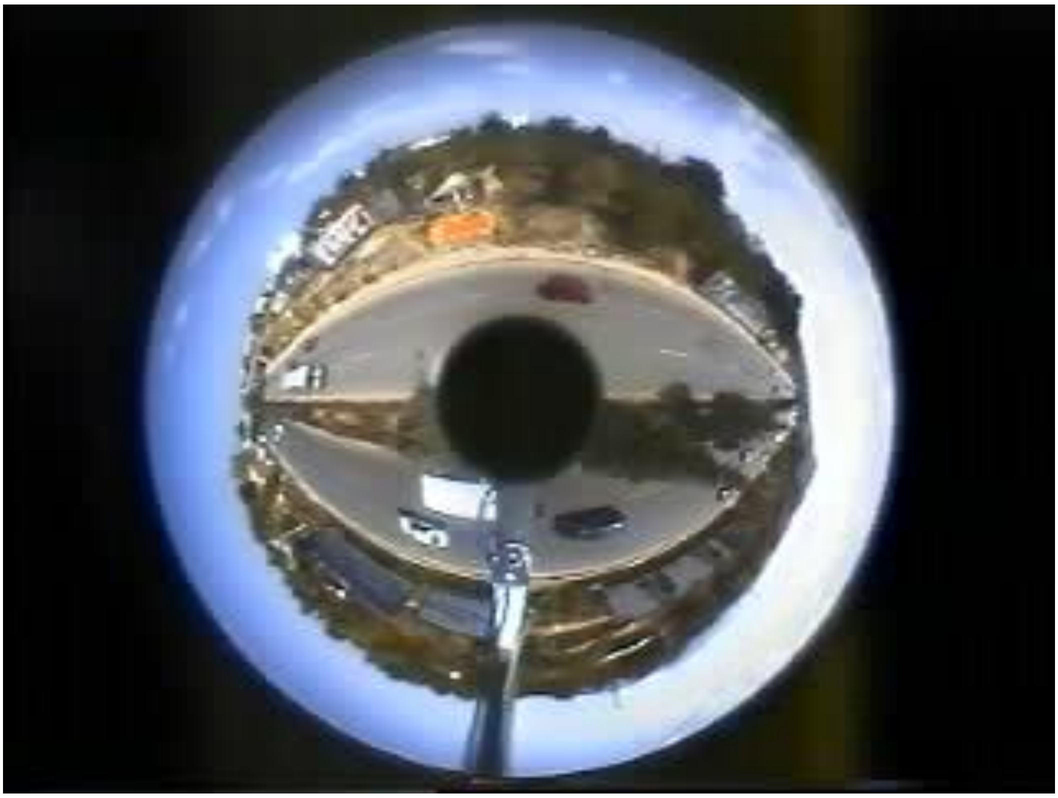


Figure 6. Original panoramic image

Figure 7. Perspective correction of a portion of the panoramic image of Figure 6. The road is not distorted in the new image and the road edges in the background appear to be formed by straight-line segments.