

A METHOD AND SYSTEM FOR OBTAINING MULTIPLE VIEWS OF AN OBJECT FOR REAL-TIME VIDEO OUTPUT

Field of the Invention

The invention concerns a method and system for obtaining multiple views of an object for real-time video output.

Background of the Invention

A microscope provides a high magnification image of a narrow field of view through the device's optics and eyepiece(s). Some microscopes provide zoom and/or interchangeable lenses so that the user can gain a wider field of view to help scene visualization or check objects near to the region of interest ("ROI"). Some microscopes are equipped with a camera and display so that the operator is not restricted to viewing through the eyepieces and can simultaneously view the microscope display and surrounding equipment.

Micro-assembly systems have been proposed which provide multiple camera views so that the advantages of the high resolution microscope view are combined with the scene information available in broader fields of view. Certain systems combine two microscopes at orthogonal viewpoints; e.g. one from above (plan) and one from the side (lateral). Systems such as scanning electron microscopes (SEM) are sometimes provided with an access port through which a camera view of the sample stage is possible. The user can view the motion of the stage on a separate monitor.

The single view microscope is tiring to use and unproductive because there is often detail beyond the field of view that must be checked by combinations of lateral displacement of the sample stage and zooming out with the optics. A monitor type display allows the operator to view the scene alongside the microscope view but the regions of interest are too small to be seen by the unaided eye in most applications.

Multiple cameras may help but require the operator to switch attention between multiple screens or to manually select views. Multiple views can cause spatial disorientation and require some skill in mental rotation of images. With cluttered

scenes or restricted working distance optics, it may not be physically possible for a second camera to obtain a view of the ROI. A microscope has a very shallow depth of field and the ROI are essentially planar. In this case a second view, inclined obliquely to the plane, will not be able to attain good focus on the scene.

Therefore, there is a desire for an improved system to obtain multiple views of a region of interest.

Summary of the Invention

In a first preferred aspect, there is provided a method for obtaining multiple views of an object for real-time video output, the method comprising:

- obtaining a first view of the object;
- obtaining a second view of the object that is different from the first view;
- and
- warping the first view using a transformation between the first and second views such that the shape and position of the first view matches the second view;
- wherein the warped first view is overlaid on the second view by applying a copy using a mask calculated by the transformation.

In a second aspect, there is provided a system for obtaining multiple views of an object for real-time video output, the system comprising:

- a first image capture device to obtain a first view of the object;
- a second image capture device to obtain a second view of the object that is different from the first view; and
- a warping module to warp the first view using a homography between the first and second views such that the shape and position of the first view matches the second view;
- wherein the warped first view is overlaid on the second view by applying a conditional copy using a mask calculated by the homography.

The object may be at least partially occluded by another object.

The transformation may be obtained by solving the inter-image motion and constructing a mosaic. The transformation may be an homography.

The homography may be obtained by the equation: $H_{ZP}^U = H_s H_d H_P^M H_M^{mos} H_{mos}^U$, where H_{mos}^U is a simple translation to account for any offset of the reference frame origin in the mosaic, H_M^{mos} is an isometry or scaled Euclidean transformation and takes the mosaic coordinates to a model space, H_P^M is a projective transformation that projects the model onto the perspective view coordinate system, H_d is a translation to account for the change of origin in a digital zoom, H_s is the scaling factor of a digital zoom, and H_{ZP}^U is the matrix transformation that takes a feature in the first image to its correct position in the second image.

The inter-image motion may be solved by at least one predetermined inter-image tracking algorithm. The mosaic may be constructed with reference to a predetermined time instant, the method further comprising:

- determining a transformation that takes the mosaic to a model space;
- recovering a transformation from the model space to the second view; and
- determining a transformation from the first view to the second view for the predetermined time instant.

The transformations may be based on available image data. The model space may be a predetermined model or a model that is created from the second view. The transformation from the first view to the second view for the predetermined time instant may be determined by matrix composition. The second view may be zoomed. Coordinate transformation may be provided to handle the effect of zooming the second view. The second view may be obtained at an oblique angle from the first view. Either or both the first view and the second view may be taken at an oblique angle to the plane of the object.

The image capture device may be an image digitiser or frame grabber. The first image capture device may capture an image via a microscope. The copy may be a conditional copy. When the warping introduces a parallax disparity, a shadow of the object may be used to resolve the parallax disparity.

The present invention provides a method so that an operator of a microscope can visualize the view provided by a microscope with respect to the 3D surrounding structure and environment; e.g. the extended sample and application tools. The present invention fuses the visual data from multiple images taken from multiple viewpoints onto a single viewpoint. This enables the operator to view and control

the task efficiently in an intuitive setting without the stress of physical and mental switching of viewpoint. The technique of fusion allows recovery of lost or obscured data and provides the ability to see through objects such as peripheral tools and the main microscope which may obstruct a region of interest. The operator may view in-focus microscope images at high resolution as seen from a perspective viewpoint.

The present invention projects the view from one view onto the perspective of another view so that this projected view is augmented with visual information (images) that would normally be occluded in that view. The operator is provided with functions for digital zooming and panning so that images can be presented at the optimal or desired combinations of field of view and resolution of detail and assist the operator in resolving depth ambiguity.

The present invention allows an operator to view image content from multiple sources (real or synthetic) on a single view, and allows multiple resolution data to be combined on a single display. This combines the benefits of fine resolution with wide fields of view. The present invention assists hand-eye coordination and reduces fatigue of the operator. The present invention does not require the real-time modelling as proposed by the prior art. In contrast, it is able to solely use real images and transform (warp) these to reconstruct hidden detail. No prior knowledge, prior models or model reconstruction is required to achieve this. There is no dependence on a priori knowledge or rigid configuration.

Brief Description of the Drawings

An example of the invention will now be described with reference to the accompanying drawings, in which:

Figure 1 is a perspective view of a work zone (marked by the large 'X') that is partially occluded by microscope lens housing;

Figure 2 is a view of the work zone of Figure 1 seen through microscope optics;

Figure 3 illustrates a preferred embodiment of the present invention where a first image of Figure 1 is digitally zoomed and a second image of Figure 2 is projected onto its corresponding location in a new image, and the new image is overlaid on the first image to provide a view through the occluding structure of the microscope lens housing;

Figure 4 is a process flow diagram for image superimposition;

Figure 5 is an image of the final result obtained by the process of Figure 4;
Figure 6 is a series of images of a frame pair from an electronic package sequence;
Figure 7 illustrates two images viewing the same plane which are projections of details from this plane;
Figure 8 illustrates a mosaic image to establish correspondences between two images spaces at different scales of resolution;
Figures 9 and 10 are two additional examples of mosaic images formed by tracking microscope images and solving the motion parameters;
Figure 11 is a series of representations to illustrate the process of recovering motion from the images and then reconstructing the scene as a mosaic image; and
Figure 12 is a block diagram of homography transformation from a microscope view to its corresponding perspective view being composed of other transformations involving a mosaic space and a model space.

Detailed Description of the Drawings

Referring to Figures 1 and 2, an operator is attempting to view the location of tools/probes 5, 6 approaching a work zone (marked by the large 'X')/region of interest (ROI) in 3D space at the same time as viewing the ROI through a microscope 7. In Figure 1, a first image 20 is captured where it is not possible to see how the probes 5, 6 interact with the ROI. There is also insufficient resolution to determine how close the tip of probe 5 is to the ROI. In Figure 2, a second image 30 is captured where the foremost probe 5 is seen without any sense of its vertical displacement from the plane 4 and the second probe 6 cannot be seen nor can any other detail beyond the 'X'.

Preferably, the second image 30 from a microscope viewpoint is captured and projected onto the first image 20 taken from a different viewpoint. The first image 20 has a wider field of view and is at an oblique angle to the second image 30. The advantage of this projection is that the operator can then see the detail provided by the microscope 7 in the context of the wider scene which may also contain structures in 3D space, and structures that obscure the view to the ROI. A method for implementing this projection is provided. A method is also provided for solving the coordinate transformations which is required to produce an accurate projection of data, despite the restricted correspondence between two views 20, 30 of

differing resolution. A 2D transformation is found that takes the geometric features from the plane 4 of the sample under the microscope 7 to the projection of this plane 4 in the second view 20. Once this has been found, real-time images from the microscope 7 are transformed and mixed with the live images of the second view 20. The microscope 7 may be of any suitable size, shape, power or form and may include a scanning electron microscope, a zooming X-Ray apparatus, and the like. Either or both the first view 20 and the second view 30 may be taken at an oblique angle to the plane of the probe 5.

In Figure 3, a warped copy of the microscope image 30 is overlaid onto the plane 4 it occupies in the perspective view. This warping (or projection) should be precise if it is to be useful to the operator. The image 30 may be augmented by reconstructing lost detail with computer graphics. Image augmentation may include combining real-time images 20, 30, mosaic images 60 and graphically modelled data. This process is carried out in real-time. By real-time it is interpreted to mean that the operation is performed and the display is updated so that the images 20, 30 are temporally consistent, any delay is not perceivable to the operator, and an acceptable frame rate is maintained. The projection introduces a parallax disparity for objects out of the plane 4. Therefore the image of the probe 5, which occupies 3D space, appears in a different perspective. The amount of disparity is related to the distance of the object from the plane 4 and advantageously provides an operator with information that is lost in a single view. It is as though a shadow 9 of the object cast by a light source directly above it is being seen. The original perspective image 20 of the probe 5 is redrawn onto the new image 40. The shadow 9 may be used to resolve the parallax disparity as it provides an understanding of the depth.

Referring to Figure 4, images 20, 30 are captured in real-time by two image capturing devices 21, 31 such as, for example, digital cameras. A clock 50 may be used for clocking the digitisers 22, 32 of the cameras 21, 31. Although the images are preferably captured somewhat simultaneously, this is not essential. Alternatively, a frame grabber that is capable of capture from two or more image capturing sources, preferably substantially simultaneous capture, may be used. The perspective view 20 is subjected to an optional cropping and resizing to produce the effect of a digital zoom on the ROI. The ROI may be determined so that it is centered over the location of the microscope view 30. The microscope view 30 is transformed so that it matches the perspective view 20. The microscope

image 30 is warped (projected) with a transformation such as, for example, an homography H so that its perspective dimensions match the zoomed perspective view 23. Alternatively to an homography, a projective transformation, two-dimensional transformation or general linear transformation may be used. The warping may be performed by a warping module embodied as software or hardware. The warped view 33 is then superimposed on the perspective view 20 to produce a fused image 40. Specifically, both images 23, 33 are fused by applying a conditional copy using a mask calculated by the same warping homography H . The composite image 40 can then be copied to the display buffer for display 41.

Figure 5 is an actual fused image 40 generated by the process illustrated in Figure 4. In the centre of the new image 40, the high resolution image 23 from the microscope 31 is seen. The visual information is entirely 2D as it is taken from above. The rectangular warped view 33 projects to quadrilateral in this view 40. Surrounding warped view 33 is the low resolution image of the zoomed perspective view 23. The circular features on the plane 4 do match but the 3D pins structures exhibit parallax. The overlay of the warped microscope image 33 enables the operator to see through the occluding pins.

In Figure 6, the top two images 20, 30 are the raw input images received from the image digitiser (frame grabber), and the middle two images 23, 33 are the transformed images, and the bottom image is the fused image result 40. The perspective view image 20 at the top left of Figure 6 is zoomed into the region shown in the image 23 below. In this zoomed image 23, the vertical pins occlude some details of the ground plane 4. The microscope image 30 at the top right is warped into the correct shape and position in the perspective view, using a required transformation (homography, H). Finally, the warped image 33 is overlaid on the zoomed image 23 to generate the fused image 40. There are differences in image resolution on the fused image 40 between the warped image 33 and zoomed image 23. The operator is free to choose the amount of zoom in the fused view 40; it is performed almost instantaneously as the zooming is digital.

In Figure 7, two cameras 21, 31 are viewing a plane 4 from two different views 20, 30. There is a one-to-one correspondence between features on the plane 4 and features on the image plane 8. Therefore, there is a one-to-one correspondence between the features appearing in the two image planes 8. The physical plane 4 induces a homographic correspondence between the two images (a 2D

homography 'H'). There is a homography H_w^1 from each camera 21, 31 to the plane 4 in the world coordinate system and so by composition there is an homography from the first camera 21 to the second camera 31:

$$H_2^1 = H_w^1 H_2^w = H_w^1 (H_w^2)^{-1}$$

The required homography H between the two views 20, 30 must be determined. The indices of H_i^j are dropped for convenience when the meaning is unambiguous. If this homography is determined, images may be mapped from one view to the other. The useful visual information from the microscope 31 is essentially planar and so the properties of planar homographies may be used to project data from one image to another. An homography is invertible and can be chained by matrix composition. A 2D homography transforms geometric data from one planar space to another. For example, a 2D point $x_1 = \{x, y, 1\}$ in homogeneous coordinates transforms to: $x_1 \mapsto x_2 = H_2^1 x_1$. In this notation the (lower) index on x refers to the coordinate system and the transformation H goes from the upper index coordinate system to the lower index coordinate system.

The homography H between the two views may be calculated with knowledge of the system's physical parameters. But this would make the system very sensitive to errors and disturbances. The homography may also be solved by placing a calibration object so that it appears in both views and using a standard scheme to solve the algebraic coefficients of the transformation by minimization of a suitable error function. Some problems associated with this include:

- a) the calibration will have to be repeated every time there is a change in the physical set-up – a slight movement of the camera, zoom or even refocus;
- b) the system's parameters can drift with time or changes in ambient conditions;
- c) it may not be possible to see the calibration object in both views; and
- d) with large differences in scale between the two views, the calibration will be highly sensitive to measurement errors.

In a preferred embodiment, there is no reliance on explicit calibration. A robust implicit calibration is achieved by tracking motion of microscope images 30, solving the inter-image motion and creating a mosaic image 60. The use of the mosaic 60 enables correspondence to be found between views taken at different scales. This function is automatic and can be performed whilst the system is in normal and continuous use.

Referring to Figure 8, the mosaic 60 is shown in the top image and is composed of about thirty microscope images 30 stitched together. By fitting this mosaic 60 to the perspective view 20, the calibration of the inter-image homography is implicitly solved without any knowledge of either camera's intrinsic and extrinsic parameters. There are features on the mosaic 60 that are recognizable on the perspective view 20, and thus the transformation between these two images spaces can be solved directly. Establishing the correspondence for the transformations may be solved manually or automatically.

Figure 9 depicts that a complete loop has been made and the features at either end are correctly aligned. This loop closure problem is a major concern in the implementation of tracking and reconstruction algorithms where errors are liable to propagate in chains of transformations. Figure 10 depicts that the tracking algorithms may succeed with more challenging regions such as, for example, when crossing a central region with little or no recognisable features, and with a metallic surface.

Figure 11a illustrates synthetic solid planar shapes. The rectangles represent the various image frames at different instances of time in the sequence. Figure 11b shows how this sequence is captured as individual frames, and how each would appear on a monitor. It is not always possible to infer inter-frame motion from these simple shapes. For example, the amount of rotation is undetermined when the image only contains a circle (frame 5 to 6). Figure 11c illustrates recovery of motion parameters. Each frame is drawn in its correct location relative to a reference frame (frame 1). Figure 11d illustrates the reconstruction of the image by transforming the individual images to fit the frames recovered in the previous step. Figure 11e is the final mosaic 60. This process creates a new image with its own coordinate system.

General motion is allowable when forming a mosaic 60. This means that the motion parameters do not have to be controlled with high precision and that the sample could even be moved manually by the operator. The motion of the view frame relative to the sample is quite general. Inter-image tracking algorithms solve the motion parameters which then lead to a reconstruction of the mosaic 60.

In Figure 12, a hypothetical sequence of frame instants is denoted by six boxes in the microscope view (U), model space (M) and the perspective view (P). A mosaic 60 is constructed with reference to a chosen time instant. Frame instant 2 has been chosen as a reference instant. The mosaic 60 is composed with respect to the second microscope image (U2) and the model 65 is made with reference to the second perspective frame (P2). Once the homography from microscope frame 2 (U2) to perspective frame 2 (P2) has been found, it can be applied to all other frame pairs {Ui, Pi}.

Next, the transformation that takes the mosaic 60 to a model space 65 is determined. The model 65 is either known beforehand or can be created from one or more perspective views 20. Then, the transformation from model space 65 to the reference perspective view image 20 is recovered. By matrix composition, the transformation from the reference frame in the microscope view (U2) to its synchronized perspective view (P2) is determined. It is then assumed that this transformation remains valid for subsequent image pairs {Ui, Pi}. The final coordinate transformation is to take into account any zoom effect, but this is not depicted on the diagram. As these operations are on digital images, they are effects equivalent to cropping and resizing so they are known and can be represented by simple translation and scaling matrices. The homography from the microscope image (U2) to the zoomed perspective view (P2) is therefore composed in the following way:

$$H_{ZP}^U = H_s H_d H_P^M H_M^{mos} H_{mos}^U$$

where all matrices H may be, for example, 3x3. They may be, for example, 4x4 if required or desired.

H_{mos}^U is a simple translation to account for any offset of the reference frame origin in the mosaic image 60,

H_M^{mos} is an isometry (or scaled Euclidean transformation) and takes the mosaic coordinates to the model space 65,

H_P^M is a projective transformation that projects the model 65 onto the perspective view coordinate system,

H_d is a translation to account for the change of origin in the digital zoom,

H_s is the scaling factor of the digital zoom.

H_{ZP}^U is the matrix transformation that takes a feature in any microscope image 30 to its correct position in the corresponding zoomed perspective view 20.

The transformation from mosaic 60 to model 65 can be factored into scaling, rotation and translation. This offers the opportunity to parameterize H with an explicit microscope zoom. As such, H can be estimated for variable zoom without having to recalibrate all the component transformations. The efficacy of the transformation H (in full: H_{zp}^U) is dependent on the quality and extent of the mosaic image 60. Enough images have to be collected so that correspondences can be found with the model 65. Clearly, at higher microscope magnifications, more images 20, 30 are required before corresponding features are found. The quality of the mosaic 60 depends on finding good inter-image transformations. High quality mosaics 60 are achievable.

Advantageously, the present invention provides a view of a scene that would not normally be possible from many viewing angles. The method and system described do not require explicit calibration or precise calibration procedures, are self-maintaining, and do not require a skilled operator to provide regular adjustment or reconfiguration. The present invention enables real-time video output, and is capable of operating with changes in focus or optical zoom. The present invention may use existing equipment of a microscopy system and require minimal modification or addition of expensive components. The system can be easily reconfigured to suit operator preference or changes in operating procedures.

The method and system may be used for motion control as it is able to resolve motion transformations. For example, it can be used for input for navigation for touch screen systems, and mouse control, as image coordinates are sufficient.. If the microscope is a zooming X-ray apparatus, the method and system may be used for semiconductor device error detection.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the scope or spirit of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects illustrative and not restrictive.

THE CLAIM:

1. A method for obtaining multiple views of an object for real-time video output, the method comprising:
 - obtaining a first view of the object;
 - obtaining a second view of the object that is different from the first view;
 - and
 - warping the first view using a transformation between the first and second views such that the shape and position of the first view matches the second view;
 - wherein the warped first view is overlaid on the second view by applying a copy using a mask calculated by the transformation.
2. The method according to claim 1, wherein the object is at least partially occluded by another object.
3. The method according to claim 1, wherein the transformation is obtained by solving the inter-image motion and constructing a mosaic.
4. The method according to claim 1, wherein the transformation is an homography.
5. The method according to claim 4, wherein the homography is obtained by the equation: $H_{ZP}^U = H_s H_d H_P^M H_M^{mos} H_{mos}^U$, where H_{mos}^U is a simple translation to account for any offset of the reference frame origin in the mosaic, H_M^{mos} is an isometry or scaled Euclidean transformation and takes the mosaic coordinates to a model space, H_P^M is a projective transformation that projects the model onto the perspective view coordinate system, H_d is a translation to account for the change of origin in a digital zoom, H_s is the scaling factor of a digital zoom, and H_{ZP}^U is the matrix transformation that takes a feature in the first image to its correct position in the second image.
6. The method according to claim 3, wherein the inter-image motion is solved by using at least one predetermined inter-image tracking algorithm.
7. The method according to claim 3, wherein the mosaic is constructed with reference to a predetermined time instant, the method further comprising:
 - determining a transformation that takes the mosaic to a model space;

recovering a transformation from the model space to the second view; and
determining a transformation from the first view to the second view for the
predetermined time instant.

8. The method according to claim 7, wherein the transformations are based
on available image data.

9. The method according to claim 7, wherein the model space is one of: a
predetermined model and a model that is created from the second view.

10. The method according to claim 7, wherein the transformation from the first
view to the second view for the predetermined time instant is determined by matrix
composition.

11. The method according to claim 1, wherein the second view is zoomed.

12. The method according to claim 11, further comprising coordinate
transformation to handle the effect of zooming the second view.

13. The method according to claim 1, wherein the second view is obtained at
an oblique angle from the first view.

14. The method accordingly to claim 1, wherein the copy is a conditional copy.

15. The method according to claim 1, wherein when the warping introduces a
parallax disparity, a shadow of the object is used for resolving the parallax
disparity.

16. A system for obtaining multiple views of an object for real-time video
output, the system comprising:

a first image capture device for obtaining a first view of the object;

a second image capture device for obtaining a second view of the object,
the second view being different to the first view; and

a warping module for warping the first view using a transformation between
the first and second views such that the shape and position of the first view
matches the second view;

wherein the warped first view is able to be overlaid on the second view by applying a copy using a mask calculated by the transformation.

17. The system according to claim 16, wherein the image capture device is an image digitiser or frame grabber.
18. The system according to claim 16, wherein the object is at least partially occluded by another object.
19. The system according to claim 16, wherein the transformation is obtained by solving the inter-image motion and constructing a mosaic.
20. The system according to claim 16, wherein the transformation is an homography.
21. The system according to claim 20, wherein the homography is obtained by the equation: $H_{ZP}^U = H_s H_d H_P^M H_M^{mos} H_{mos}^U$, where H_{mos}^U is a simple translation to account for any offset of the reference frame origin in the mosaic, H_M^{mos} is an isometry or scaled Euclidean transformation and takes the mosaic coordinates to a model space, H_P^M is a projective transformation that projects the model onto the perspective view coordinate system, H_d is a translation to account for the change of origin in a digital zoom, H_s is the scaling factor of a digital zoom, and H_{ZP}^U is the matrix transformation that takes a feature in the first image to its correct position in the second image.
22. The system according to claim 19, wherein the inter-image motion is solved by using at least one predetermined inter-image tracking algorithm.
23. The system according to claim 19, wherein the mosaic is constructed with reference to a predetermined time instant, the system further comprising:
 - determining the transformation that takes the mosaic to a model space;
 - recovering the transformation from the model space to the second view;
 - and
 - determining the transformation from the first view to the second view for the predetermined time instant.

24. The system according to claim 23, wherein the model space is a predetermined model or a model that is created from the second view.
25. The system according to claim 23, wherein the transformation from the first view to the second view for the predetermined time instant is determined by matrix composition.
26. The system according to claim 16, wherein the second view is zoomed.
27. The system according to claim 26, further comprising coordinate transformation to handle the effect of zooming the second view.
28. The system according to claim 15, wherein the second view is obtained at an oblique angle from the first view.
29. The system according to claim 16, wherein the first image capture device captures an image via a microscope.
30. The system according to claim 16, wherein the copy is a conditional copy.
31. The system according to claim 16, wherein when the warping introduces a parallax disparity, a shadow of the object is used for resolving the parallax disparity.
32. The system according to claim 28, wherein at least one of the first view and the second view is at an oblique angle to the object plane.