Appendix A
Imaging Geometry Issues

A primary goal of UVIS is to produce an image, or set of images, of the underside of a vehicle that can be readily interpreted by an operator and will facilitate detection of anomalous objects. The image set must provide complete coverage of the vehicle undercarriage and, ideally, should simulate different viewpoints to enable seeing into and around undervehicle features. System requirements dictate that the system have a small footprint to increase portability, ease of use and cost efficiency.
The resulting set of requirements pose a challenge for image processing, especially since some requirements tend to drive system design in opposite directions, therefore, an engineering design that respects and balances these requirements must be developed.

One inevitable element of the system design is that multiple cameras (real or virtual) are required. Since the entire undercarriage must be imaged and the requirement for a small footprint dictates limited camera standoff, multiple cameras are required to image the entire vehicle (see Figure A.1).

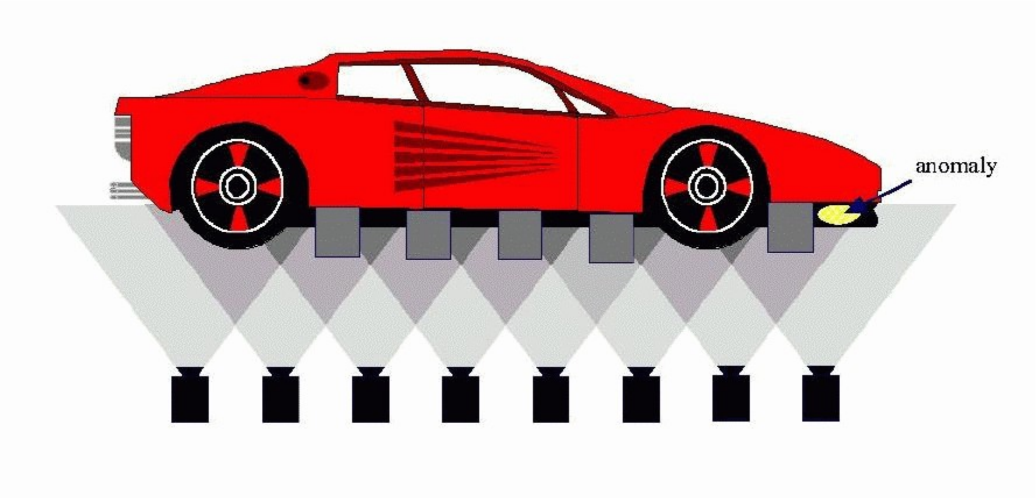


Figure A.1 Multiple Cameras are Required

Given that multiple cameras are used, the different views from these cameras must be merged together into a composite image that represents the entire viewing area. The resulting composite image is referred to as a “mosaic” and the merging process is known as “mosaicing” (see Figure A.2).

There are a range of different approaches that can be used to accomplish mosaicing. These approaches can generally be characterized based on how they use image information from the different cameras. At one end of this range are approaches that recover and exploit 3D information, i.e. the 2D planar information inherent in an image is augmented by depth information perpendicular to the image plane. These approaches maintain a high level of geometric integrity.

Alternatively, mosaicing can be performed by merging images along a seam that tends to minimize disruption from one image to the next. Since 2D image information from different views is used directly without appealing to depth information, geometric integrity is not necessarily maintained.

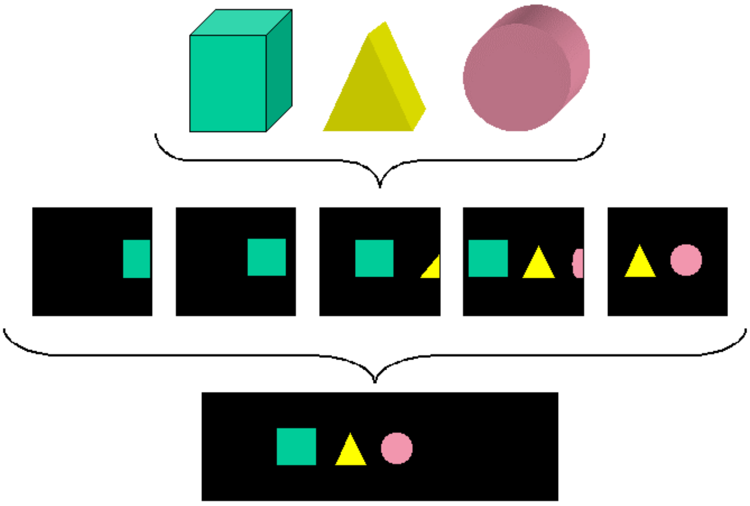


Figure A.2. Multiple View Imaging - Mosaicing

Mosaicing approaches that recover and use depth information rely on the relationship between depth and the corresponding shift (also know as “disparity” - displacements in the X and Y directions) induced between two images that view a common area from different viewpoints (see Figure A.3). The closer a scene component is to viewing cameras, the more the corresponding image component will shift from one view to the next. Depth information relates directly to image shifts between views and, symmetrically, image shifts between views relates directly to depth information. Therefore, if corresponding components (e.g. features, regions, pixels ...) in different images can be identified, the amount the components shift from one image to the next determines the depth to the associated scene component. If depth information can be recovered over almost all of a viewing area, the information is considered to be “dense”. Alternatively, if only some depth information can be recovered, the information is considered to be “sparse” and depth for the remaining areas is not directly known.

If dense depth information can be recovered, either directly or by estimation, a 3D model of the imaged scene can be constructed and synthetic views from different viewpoints can be rendered, providing high quality images of the entire viewing area that maintain geometric integrity. In some cases, the recovered depth information may provide an additional discrimination component for enhanced change detection. While the recovered depth information can facilitate both operator interpretation and change detection, the process requires reliable identification of corresponding image components in different images (as described above). The identification of corresponding image components is commonly referred to as “matching” and is critical to the success of depth recovery and the use of 3D information to create image mosaics.

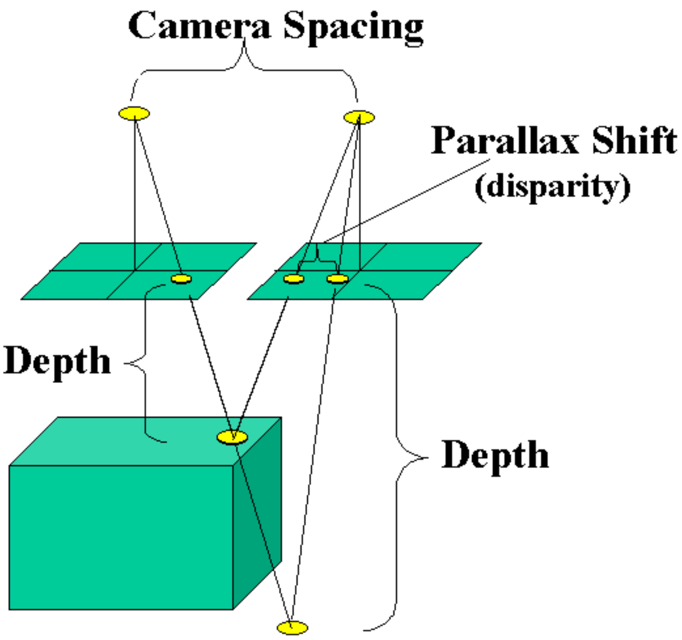


Figure A.3 Relation between image disparity and scene depth

SHIFT <> DEPTH

As mentioned earlier, mosaicing can be performed without explicitly recovering 3D information. Instead, images may be cut along minimally disruptive seams and then the image segments can be pasted together into a new composite image, which forms the basis of a mosaic image. Matching must be performed, along potential seams at least, to determine where to cut and paste image segments. Here, image segments derived from different views may be used directly, which will not necessarily maintain image integrity. Further, if a chosen seam is not optimum, it may be necessary to warp or "morph" objects along the seam to achieve a smooth transition.

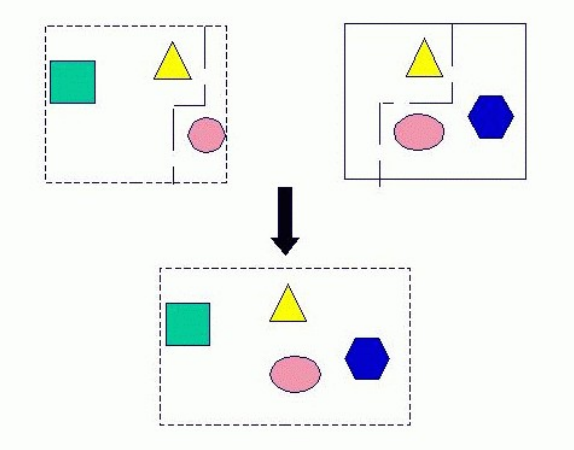


Figure A.4. Merging along a minimally disruptive seam using 2D image data. Note the different locations and shapes of objects in each image

.

The result is images that may be difficult for an operator to interpret and may complicate change detection processing. However, since this approach may not require as dense matching information, it will not necessarily require as powerful a matching process as a 3D approach that provides higher geometric integrity. It is important to note that, while the power of a matching process required may vary, **all mosaicing processes under consideration are critically dependent on matching**. Quality matching information is necessary, but not sufficient, to produce quality image mosaics.

The requirement for quality image matching across adjacent views is an important driver of engineering considerations and system design. This can be seen by noting that matching can only be performed effectively when images that need to be matched are “similar” in appearance. It should be further noted that, fundamentally, this is a consequence of information content and not the strengths or weaknesses of any given process. Reliable matching in the image domain can only be performed when the patterns to be matched are similar.

As differences between viewing geometry for adjacent views increase, so too does the appearance of the corresponding images. Image components will undergo increasing distortion and, eventually, some components may become occluded, making matching difficult or even impossible (see Figure A.5).

Images will tend to appear “similar” if the light rays that produce the images are “nearly parallel”. Quantitatively, this occurs when the distance between adjacent cameras (camera spacing, S) is small with respect to the standoff distance from the scene (depth, D). Therefore, the condition for similar image appearance and, consequently, reliable matching is **S << D or S/D << 1** (see Figures A.6 and A.7).

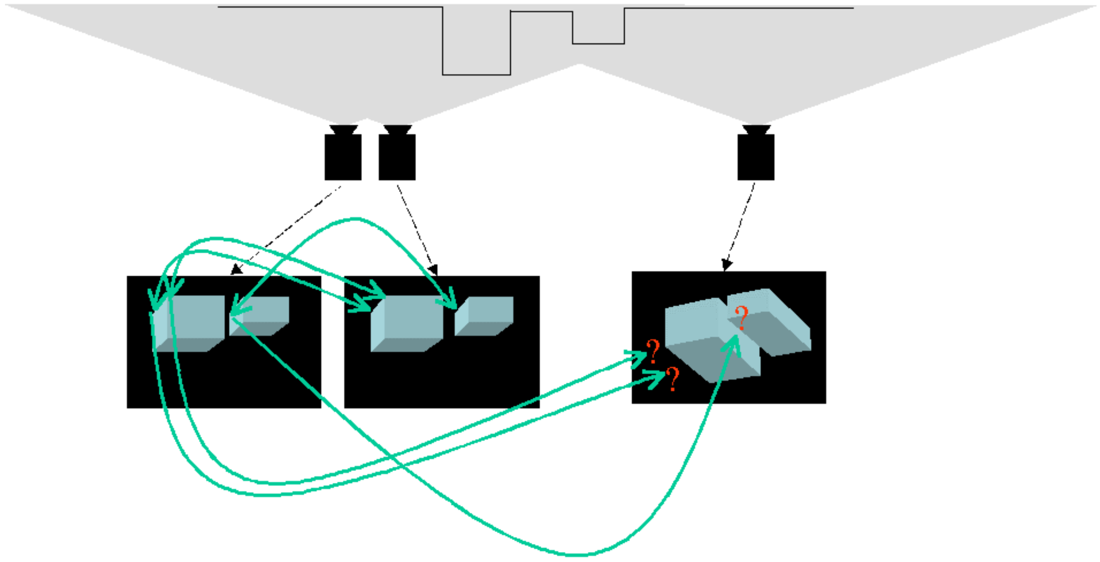
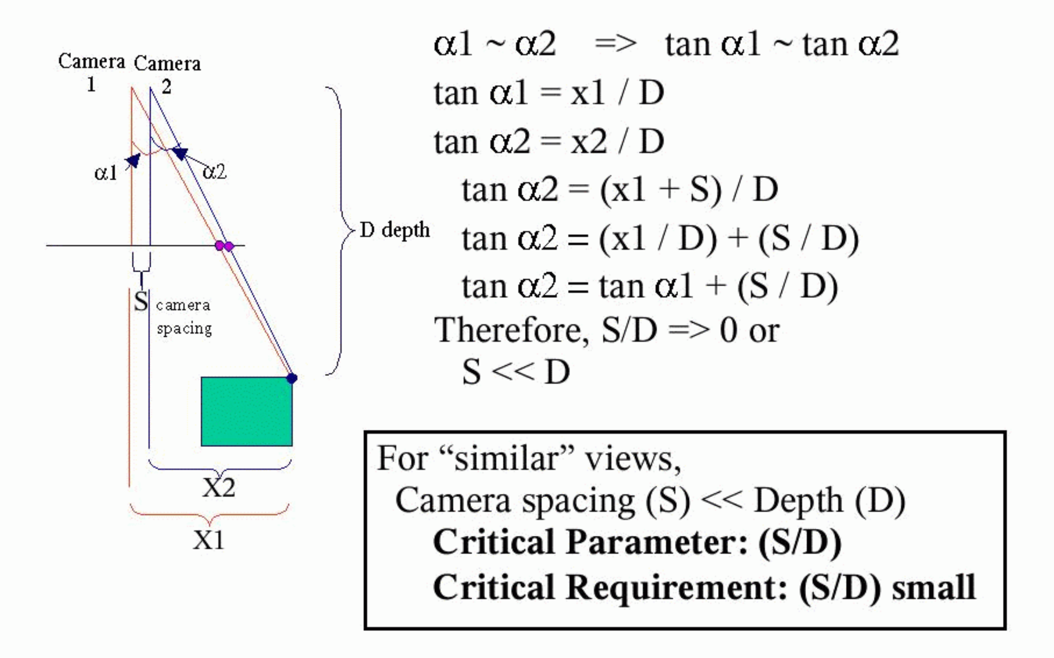


Figure A.5 Effect of viewing geometry variations on image appearance and matching. Image features can be effectively matched between the first and second images (from the left). However features in the third image are distorted or missing due to occlusion.

Since the standoff distance must be relatively small due to the requirement for a low profile system, the spacing between views must also be small. Consequently, there is a need for a large number of views. The number of views required can be reduced if the matching process used can tolerate some distortion between images – the more distortion that can be tolerated, the less views are required, which has practical consequences, such as affecting the number of cameras that are needed. However, many of the existing techniques for matching and mosaicing were developed for domains where S/D values are typically far smaller than in the UVIS system (e.g. - aerial imagery). Therefore, the quality of image matching versus image geometry is a critical issue in UVIS that drives many of the engineering decisions that must be made.



$$\alpha 1 \sim \alpha 2 \quad \Rightarrow \quad \tan \alpha 1 \sim \tan \alpha 2$$
$$\tan \alpha 1 = x1 / D$$
$$\tan \alpha 2 = x2 / D$$
$$\tan \alpha 2 = (x1 + S) / D$$
$$\tan \alpha 2 = (x1 / D) + (S / D)$$
$$\tan \alpha 2 = \tan \alpha 1 + (S / D)$$

Therefore, S/D => 0 or
$S << D$

For “similar” views,
Camera spacing (S) << Depth (D)
Critical Parameter: (S/D)
Critical Requirement: (S/D) small

Figure 2.6 Geometric condition for similar viewpoints

The images above were acquired right to left, as indicated and the right most image serves as reference. The effect of similarity can be seen as S/D increases. The first two (S/D = 0.06 and 0.3125) are reasonably similar to the reference image, however, the remaining images exhibit significant distortion and probably can not be successfully matched with the reference.

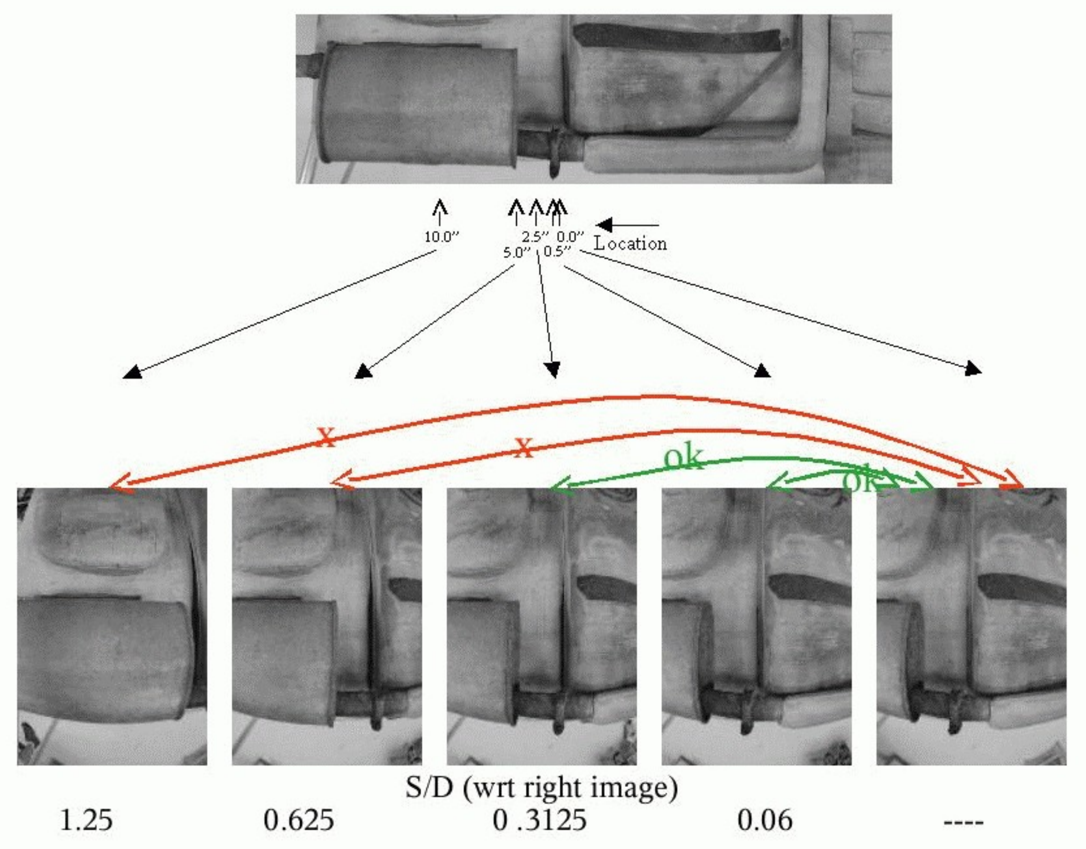


Figure A.7. Example of impact on S/D on image similarity**.**