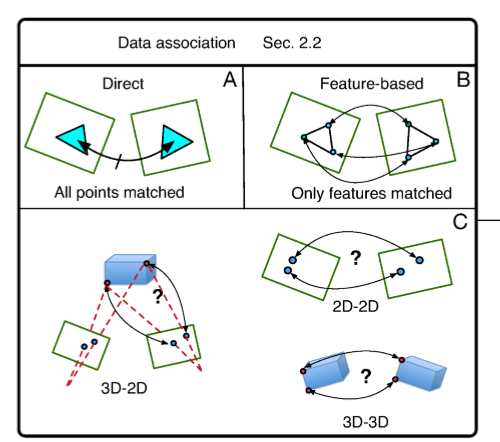
uses previous poses to guess poses for new frames

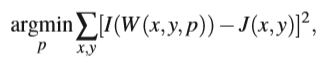


Data Association Choices – the data association method can be direct, feature-based, or a mix of the two. Direct methods use every pixel, whereas semi-dense methods only use pixels at which the gradient of image brightness is signiﬁcant.

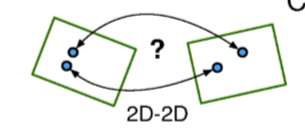
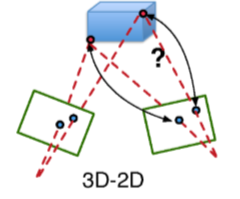
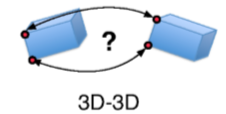
The basic underlying principle for all direct methods is known as the brightness consistency constraint and is best described as:

J(x,y,t) = I(x+u(x,y),y+v(x,y),t+1), (1)

where x and y are pixel coordinates; u and v denote displacement functions of the pixel (x,y) between two images I and J of the same scene taken at time t and t +1 respectively. The brightness consistency constraint is based on the assumption that a point from the world’s surface, observed in image I, will have the same intensity when observed in a subsequent image J. To render equation (1) solvable, [7] suggested, in what they referred to as Forward Additive Image Alignment (FAIA), to replace all the individual pixel displacements u and v by a single general motion model, in which the number of parameters is dependent on the implied type of motion. FAIA iteratively minimizes the squared pixel-intensity difference between the two images over the transformation parameters p:

where W(.,., p) is a warping transform that encodes the relationship relating the two images and p corresponds to the parameters of the transform. Equation (2) is non-linear and requires an iterative non-linear optimization process, with a computational complexity of O(n2N +n3) per iteration, where n is the number of parameters in p and N is the number of pixels in the image. Since 1981, other variants of the FAIA were suggested such as FCIA (Forward Compositional Image Alignment), ICIA (Inverse Compositional Image Alignment) and IAIA (Inverse Additive Image Alignment) each with different computational complexities. A detailed comparison between these variations can be found in [8].

Data Association Types –

* 2D-2D
* 3D-2D
* 3D-3D

Feature-based methods were introduced to reduce the computational complexity of processing each pixel; this is done by matching only salient image locations, referred to as features, or keypoints. An example of feature-based matching is shown in Fig. 3-B. A descriptor is associated to each feature, which is used to provide a quantitative measure of similarity to other keypoints. On one hand, features are expected to be distinctive, invariant to viewpoint and illumination changes, as well as resilient to blur and noise; on the other hand, it is desirable for feature extractors to be computationally efﬁcient and fast. Unfortunately, such objectives are hard to achieve at the same time causing a trade-off between computational speed and feature quality. The computer vision community has developed, over decades of research, many different feature extractors and descriptors, each exhibiting varying performances in terms of ro

tation and scale invariance, as well as speed [9]. The selection of an appropriate feature detector depends on the platform’s computational power and the type of environment. Feature detector examplesinclude the Hessian corner detector [10], Harris detector [11], Shi-Tomasi corners [12], Laplacian of Gaussian detector [13], MSER [14], Difference of Gaussian [15] and the accelerated segment test family of detectors (FAST, AGAST, OAST) [16].

Feature descriptors include, and are not limited to, BRIEF [17], BRISK [18], SURF [19], SIFT [20], HoG [21], FREAK [22], ORB [23], and a low level local patch of pixels. Further information regarding feature extractors and descriptors is outside the scope of this work, but the reader can refer to [24], [25], [26], or [27] for comparisons.

