Chapter 16

Multi-Photo Fusion through Projective Geometry

Luigi Barazzetti
Department of Building Environmental Science and Technology, Politecnico di Milano, Italy

ABSTRACT

Multiple images of the same object can be registered and fused after identifying a proper mathematical model. This chapter illustrates the contribution of projective geometry and automated matching techniques to several photographic applications: high dynamic range, multi-focus, and panoramic photography. The method relies on the estimation of a set of planar transformations that make the data consistent. This provides pixel correspondence and allows the photographer to obtain new digital products in an almost fully automated way.

INTRODUCTION

Image registration allows merging multiple shots acquired with digital cameras. Nowadays, most digital and film-based photographs can be indentified as central perspectives in space. The geometric process behind the creation of a photo requires the alignment between the (i) perspective centre of the camera, the (ii) object point and its (iii) projection on the image plane. Most cameras follow this condition (called collinearity principle) and are therefore termed pinhole cameras (Hartley & Zisserman, 2004). For the sake of completeness, it is important to remember that there are other camera models. For instance, a standard SRL camera with a fisheye lens is no longer a pinhole camera, whereas pushbroom sensors are often employed in the field of aerial and satellite photogrammetry. In some cases, special line-based panoramic cameras are used in terrestrial applications (Luhmann & Tecklenburg, 2002).

Digital pinhole cameras are today very popular because they are simple, quick and reliable recording tools. The market seems to be growing.

DOI: 10.4018/978-1-4666-4490-8.ch016
(the global digital camera shipment was about 130 million unites in 2009, 141 million units in 2010, and it is expected to reach 145 million units in 2011 (Global and China Digital Still Camera –DSC, 2011). and the development of digital camera technology provides sensors with high radiometric and geometric resolutions and good storage capacity. Data can be easily shared, modified, enhanced, and copied. If photography has been a chemical process for over 150 years, nowadays films, chemicals or dark rooms are a thing of the past.

As most consumer cameras produce pinhole images, this chapter focuses on this particular kind of data, although the collinearity principle is satisfied only in the case of ideal cameras. A real camera equipped with real lenses produces distorted images because of the effect of lens distortion. This effect is strongly reduced if the camera is calibrated (Remondino & Fraser, 2006), meaning that distortion coefficients are known and a new distortion-free image can be created.

One of the most remarkable advantages of digital data is the opportunity to run processing algorithms able to modify, enhance, and register multiple shots. In this chapter three products obtainable with pinhole cameras are illustrated and discussed. They include High Dynamic Range (HDR) images (Debevec & Malik, 1997), multi-focus images (Haebeler, 1994), and panoramas (Brown & Lowe, 2003). Although very different, the processing algorithm relies on a similar workflow, where a set of projective transformations is used to align different images. All these products can be created in an almost fully automated way, allowing expert operators or normal tourists to overcome some drawbacks of single shots. After a brief review of the most common 2D transformations, it will be demonstrated that a homography (Liebowitz & Zisserman, 1998) can be successfully used to register images and create HDR, multi-focus, and panoramic images.

**IMAGE REGISTRATION THROUGH PROJECTIVE TRANSFORMATIONS**

Projective geometry allows us to deal with geometric transformations and offers more powerful methods than conventional Euclidean geometry. For instance, in Euclidean geometry any two distinct points in a plane define a line. On the other hand, any two distinct lines define a point, unless the lines are parallel. Projective geometry overcomes this limitation in a symmetric way: any two distinct points determine a unique line and any two distinct lines intersect at a point. When we consider the image formation process by means of a pinhole camera, it is quite evident that Euclidean geometry is not sufficient. Properties like lengths, angles, areas, and parallelism are not preserved during the imaging process: parallel lines may intersect.

We get to projective geometry by taking Euclidean geometry and adding an extra dimension. A point in an n-dimensional Euclidean space is represented as a point in an (n+1)-dimensional projective space. Suppose we have a point \((x, y)\) in the Euclidean space, the *homogenous coordinates* of a point can be obtained by adding an extra coordinate to the pair as \((\lambda x, \lambda y, \lambda)\). We say that this 3-vector is the same point in homogeneous coordinates (for any non-zero value \(\lambda\)). An arbitrary homogeneous vector \(\mathbf{x} = (x_1, x_2, x_3)\) represents the point \(\mathbf{x} = (x_1 / x_2, x_3 / x_3)\) in \(\mathbb{R}^2\).

As a line in the plane is represented by the equation \(ax + by + c = 0\), a similar notation can be used to identify a line through a vector \(\mathbf{l} = (a, b, c)^T\).

A point \(\mathbf{x} = (x_1, x_2, x_3)^T\) lies on the line \(\mathbf{l} = (a, b, c)^T\) if and only if \(ax + by + c = 0\). This may be written in terms of a vector product as \(\mathbf{x}^T \mathbf{l} = 0\). The intersection of two lines \(\mathbf{l} = (a, b, c)^T\) and \(\mathbf{l}' = (a', b', c')^T\) is the point \(\mathbf{x} = \mathbf{l} \times \mathbf{l}'\). As can be easily seen, the role of points and lines can be interchanged: \(\mathbf{x}^T \mathbf{l} = 0\) implies \(\mathbf{l}^T \mathbf{x} = 0\). This